TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 364, Number 7, July 2012, Pages 3449–3478
S 0002-9947(2012)05380-3
Article electronically published on February 20, 2012

THE SOLUTION OF THE KATO PROBLEM FOR DEGENERATE ELLIPTIC OPERATORS WITH GAUSSIAN BOUNDS

DAVID CRUZ-URIBE, SFO AND CRISTIAN RIOS

Abstract. We prove the Kato conjecture for degenerate elliptic operators on $\mathbb{R}^n$. More precisely, we consider the divergence form operator $\mathcal{L}_w = -w^{-1}\mathrm{div}\mathbf{A}\nabla$, where w is a Muckenhoupt A_2 weight and $\mathbf{A}$ is a complex-valued $n\times n$ matrix such that $w^{-1}\mathbf{A}$ is bounded and uniformly elliptic. We show that if the heat kernel of the associated semigroup $e^{-t\mathcal{L}_w}$ satisfies Gaussian bounds, then the weighted Kato square root estimate, $\|\mathcal{L}_w^{1/2}f\|_{L^2(w)} \approx \|\nabla f\|_{L^2(w)}$, holds.

1. Introduction

The purpose of this work is to give a positive answer to the Kato square root problem for a class of degenerate elliptic operators, under the assumption that the associated heat kernel satisfies classic Gaussian upper bounds.

Before stating our results, we briefly sketch the background. Given a uniformly elliptic, $n \times n$ complex matrix $\mathbf{A}$, define the second-order elliptic operator $\mathcal{L} = -\mathrm{div}\mathbf{A}\nabla$. Then the square root $\mathcal{L}^{1/2}$ can be defined using the functional calculus. The original Kato problem was to show that for all f in the Sobolev space H^1, $\|\mathcal{L}^{1/2}f\|_{L^2} \approx \|\nabla f\|_{L^2}$. This was first posed by Kato [25] in 1961, but only solved in the past decade in a series of remarkable papers by Auscher, *et al.* [4, 5, 22]. Initially, they solved the problem given the additional assumption that the heat kernel of the semigroup $e^{-t\mathcal{L}}$ satisfied Gaussian bounds. Such estimates were known to be true in the case $\mathbf{A}$ was real symmetric, but it had been shown that they need not hold for complex matrices in higher dimensions [3]. The final proof omitted this hypothesis. For a more complete history of this problem, we refer the reader to the above papers or to the review by Kenig [27].

We have extended this result to the case of degenerate elliptic operators, where the degeneracy is controlled by a weight in the Muckenhoupt class A_2. We say that a weight w (i.e., a non-negative, locally integrable function) is in A_2 if

$$\sup_Q \left(\fint_Q w(x)\,dx\right)\left(\fint_Q w(x)\,dx\right) = [w]_{A_2}^2 < \infty,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^n$. Given $w \in A_2$ and constants λ, Λ, $0 < \lambda \le \Lambda < \infty$, let $\mathcal{E}_n(w,\lambda,\Lambda)$ denote the class of $n \times n$ matrices $\mathbf{A} = (A_{ij}(x))_{i,j=1}^n$ of complex-valued, measurable functions satisfying the degenerate

Received by the editors July 23, 2009 and, in revised form, May 18, 2010.
2010 *Mathematics Subject Classification.* Primary 35J70, 35K45, 35K65, 35C15; Secondary 47D06, 47N20.

ellipticity condition

$$\begin{cases} \lambda w\left(x\right)|\xi|^2 \leq \mathrm{Re}\left\langle \mathbf{A}\xi,\xi\right\rangle = \mathrm{Re}\sum_{i,j=1}^n A_{ij}\left(x\right)\xi_j\bar{\xi}_i, \\ |\left\langle \mathbf{A}\xi,\eta\right\rangle| \leq \Lambda w(x)|\xi||\eta| \end{cases} \tag{1.1}$$

for all ξ, $\eta \in \mathbb{C}^n$.

Given $\mathbf{A} \in \mathcal{E}_n(w,\lambda,\Lambda)$, define the degenerate elliptic operator in divergence form $\mathcal{L}_w = -w^{-1}\mathrm{div}\mathbf{A}\nabla$. Such operators were first considered by Fabes, Kenig and Serapioni [17] and have been considered by a number of other authors since. (See, for example, [8, 9, 10, 11, 18, 28, 29].) It is a natural question to extend the Kato problem to these operators: that is, to show that

$$\|\mathcal{L}_w^{1/2}\|_{L^2(w)} \approx \|\nabla f\|_{L^2(w)}$$

for all f in the weighted Sobolev space $H^1(w)$. (Exact definitions will be given below.) We consider this problem in the special case that the heat kernel of the semigroup $e^{-t\mathcal{L}_w}$ satisfies Gaussian bounds. More precisely, we assume there exists a heat kernel $W_t(x,y)$ associated to the operator $e^{-t\mathcal{L}_w}$ such that for all $f \in C_c^\infty$,

$$e^{-t\mathcal{L}_w}f(x) = \int_{\mathbb{R}^n} W_t(x,y)f(y)\,dy. \tag{1.2}$$

Furthermore, for all $t > 0$ and x, $y \in \mathbb{R}^n$, the kernel W_t satisfies the Gaussian bounds

$$|W_t(x,y)| \leq \frac{C_1}{t^{n/2}}\exp\left(-C_2\frac{|x-y|^2}{t}\right), \tag{1.3}$$

and the Hölder continuity estimates

$$\begin{aligned} &|W_t\left(x+h,y\right) - W_t\left(x,y\right)| + |W_t\left(x,y+h\right) - W_t\left(x,y\right)| \\ &\leq \frac{C_1}{t^{n/2}}\left(\frac{|h|}{t^{1/2}+|x-y|}\right)^\mu \exp\left(-C_2\frac{|x-y|^2}{t}\right), \end{aligned} \tag{1.4}$$

where $h \in \mathbb{R}^n$ is such that $2|h| \leq t^{1/2}+|x-y|$. The constants C_1, C_2 and μ depend only on n, w, λ, and Λ. If these three properties hold we will say that $e^{-t\mathcal{L}_w}$ satisfies Condition (G).

Our main result is the following theorem.

Theorem 1.1. *Given $w \in A_2$ and $\mathbf{A} \in \mathcal{E}_n(w,\lambda,\Lambda)$, suppose that $e^{-t\mathcal{L}_w}$ satisfies Condition (G). Then there exists a positive constant $C = C\left(n,\lambda,\Lambda,w\right)$ such that for all f in $H_0^1(w)$,*

$$\|\nabla f\|_{L^2(w)}C^{-1} \leq \|\mathcal{L}_w^{1/2}f\|_{L^2(w)} \leq C\|\nabla f\|_{L^2(w)}. \tag{1.5}$$

We must remark that Condition (G) in Theorem 1.1 is indeed the appropriate type of Gaussian decay to be expected for the semigroup generated by the operator $\mathcal{L}_w$. As shown in [14], if the matrix $\mathbf{A} \in \mathcal{E}\left(n,\lambda,\Lambda,w\right)$ is real and symmetric, then the semigroup $e^{-t\mathcal{L}_w}$ is given by a heat kernel satysfying (1.2), (1.3), and (1.4). This is somewhat striking as the weight w does not intervene in the geometry of the kernel, and the semigroup behaves just like that of the classic heat equation in $\mathbb{R}^n$.

The natural question arising is whether there are other matrices than real and symmetric in the class $\mathcal{E}\left(n,\lambda,\Lambda,w\right)$ such that their associated heat semigroup is given by a kernel satisfying Condition (G). The answer to this question is, not surprisingly, positive. This is not hard to check for a class of smooth perturbations.

However, we conjecture that, as in the unweighted situation (see Theorems 1.6 and 1.26 in [6]), Gaussian bounds are stable by small L^∞ perturbations of the matrix $\mathbf{A}/w$.

Finally, let us note that, as in the non-weighted case, to prove Theorem 1.1, it actually suffices to prove the second inequality in (1.5):

$$\|\mathcal{L}_w^{1/2} f\|_{L^2(w)} \leq C\|\nabla f\|_{L^2(w)}. \tag{1.6}$$

For suppose this inequality holds. Since $\mathbf{A}\in\mathcal{E}\,(n,\lambda,\Lambda,w)$ implies $\mathbf{A}^*\in\mathcal{E}\,(n,\lambda,\Lambda,w)$, (1.6) holds for $(\mathcal{L}_w^{1/2})^* = (\mathcal{L}_w^*)^{1/2}$. (This operator identity follows from the functional calculus, for instance, from (6.2).) Therefore, by the ellipticity conditions,

$$\begin{aligned}
\|\nabla f\|_{L^2(w)}^2 &= \int_{\mathbb{R}^n} |\nabla f(x)|^2 w(x)\,dx \\
&\leq \lambda^{-1}\mathrm{Re}\int_{\mathbb{R}^n} \mathbf{A}\nabla f(x)\cdot\overline{\nabla f(x)}\,dx \\
&= \lambda^{-1}\mathrm{Re}\int_{\mathbb{R}^n} \mathcal{L}_w f(x)\overline{f(x)}w(x)\,dx \\
&= \lambda^{-1}\mathrm{Re}\int_{\mathbb{R}^n} \mathcal{L}_w^{1/2} f(x)\overline{(\mathcal{L}_w^{1/2})^* f(x)}w(x)\,dx \\
&\leq \lambda^{-1}\|\mathcal{L}_w^{1/2} f\|_{L^2(w)}\|(\mathcal{L}_w^*)^{1/2} f\|_{L^2(w)} \\
&\leq C\lambda^{-1}\|\mathcal{L}_w^{1/2} f\|_{L^2(w)}\|\nabla f\|_{L^2(w)}.
\end{aligned}$$

Our proof of inequality (1.6) follows the outline of the proof of the classical Kato problem with Gaussian bounds in [22]. (See also the expository treatment in [21].) There are four main steps: in Section 6 we reduce (1.6) to a square function inequality; in Section 7 we show that this inequality is a consequence of a Carleson measure estimate; in Section 8 we prove a weighted Tb theorem for square roots; finally, in Section 9 we construct the family of test functions needed to use the Tb theorem to prove the Carleson measure estimate. Prior to the proof itself, in Section 2 we motivate our choice of operator $\mathcal{L}_w$; in Sections 3 and 4 we give some preliminary results about degenerate elliptic operators, Gaussian bounds and weighted norm inequalities. In Section 5 we prove two weighted square function inequalities needed in our proof. The first inequality in particular is central to our purposes, since it is the substitute for the much simpler Fourier transform approach possible in the unweighted case.

Throughout, all notation is standard or will be defined as needed. The letters C, c will denote constants whose value may change at each appearance. Given a function f and $t>0$, define $f_t(x) = t^{-n}f(x/t)$. Given an operator T on a Banach space X, let $\|T\|_{\mathcal{B}(X)}$ denote the operator norm of T.

2. About the choice of the operator $\mathcal{L}_w$

To motivate the choice of the operator $\mathcal{L}_w$ as the appropriate one for the treatment of the particular weighted Kato problem considered, we may rely on the first order approach to the Kato problem given by Auscher, Axelsson, and McIntosh [2]. In the classic Kato problem for divergence form elliptic operators, the differential operator considered is $\mathcal{L} = -\mathrm{div}\mathcal{A}\,(x)\,\nabla$, where the matrix $\mathcal{A}$ is complex elliptic:

$$\mathrm{Re}\,\langle\mathcal{A}\xi,\xi\rangle \geq \lambda\,|\xi|^2\,, \qquad |\langle\mathcal{A}\xi,\eta\rangle| \leq \Lambda\,|\xi|\,|\eta|\,, \qquad \xi,\eta\in\mathbb{C}^n.$$

The operator $\mathcal{L}$ is seen as the composition of the gradient operator with a multiplication operator (by the matrix $\mathcal{A}$), and then a further composition with the divergence operator $\operatorname{div} = (\nabla)^*$, which is the adjoint operator to the gradient. Algebraically, this is achieved by working in the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^n, \mathbb{C}) \times L^2(\mathbb{R}^n, \mathbb{C}^n)$ (the product of the space where "functions live", times the space where "gradients live") and considering a differential operator D and a multiplication operator B, where

$$D = \begin{pmatrix} 0 & \operatorname{div} \\ -\nabla & 0 \end{pmatrix} \qquad \text{and} \qquad B(x) = \begin{pmatrix} 1 & 0 \\ 0 & \mathcal{A}(x) \end{pmatrix}.$$

These are $(n+1) \times (n+1)$ matrices. Of course, the domain of D is not in general the whole space, but rather a dense subset of it. It follows that

$$(BD)^2 = \begin{pmatrix} \mathcal{L} & 0 \\ 0 & -\mathcal{A}(x)\nabla\operatorname{div} \end{pmatrix}.$$

In [2] it was shown that if the operator BD satisfies certain quadratic estimates, then the functional calculus formula $\left\|\sqrt{(BD)^2}\right\| \approx \|BD\|$ is satisfied; in particular, applying this to functions

$$\left\|\sqrt{\mathcal{L}}f\right\| = \left\|\sqrt{(BD)^2}\begin{pmatrix} f \\ 0 \end{pmatrix}\right\| \approx \left\|BD\begin{pmatrix} f \\ 0 \end{pmatrix}\right\| \approx \left\|D\begin{pmatrix} f \\ 0 \end{pmatrix}\right\| = \|\nabla f\|,$$

which is Kato's equivalence.

In the weighted Kato problem considered in this paper, functions lie in the space $L^2(w) = L^2_w(\mathbb{R}^n, \mathbb{C})$ and gradients in $L^2_w(\mathbb{R}^n, \mathbb{C}^n)$. That is, the gradient operator ∇ has domain contained in $L^2_w(\mathbb{R}^n, \mathbb{C})$ and range in $L^2_w(\mathbb{R}^n, \mathbb{C}^n)$. It follows that the adjoint operator is given by $(\nabla)^* = -w^{-1}\operatorname{div}(w\cdot)$. Then we define the differential operator

$$D_w = \begin{pmatrix} 0 & w^{-1}\operatorname{div}(w\cdot) \\ -\nabla & 0 \end{pmatrix}$$

on the product space; and for B as above, we have

$$(BD_w)^2 = \begin{pmatrix} \mathcal{L}_w & 0 \\ 0 & -w^{-1}\mathbf{A}(x)\nabla\left[w^{-1}\operatorname{div}(w\cdot)\right] \end{pmatrix},$$

where $\mathbf{A}(x) = w(x)\mathcal{A}(x) \in \mathcal{E}_n(w, \lambda, \Lambda)$. This shows that the operator $\mathcal{L}_w$ is indeed the right choice to work on the weighted spaces L^2_w and $L^2_w(\mathbb{R}^n, \mathbb{C}^n)$. We conjecture that by adapting the approach in [2] to the weighted case, it might be possible to obtain the general Kato estimate, independent of the Gaussian bounds assumption. This matter is left for future work.

The above method also suggests how to treat other weighted operators, or rather, it sheds light on what spaces should functions lie on. For example, to treat the operator $\tilde{\mathcal{L}}_w = -\operatorname{div}\mathbf{A}\nabla$, with $\mathbf{A} \in \mathcal{E}_n(w, \lambda, \Lambda)$, the gradient should be considered as a mapping from $L^2(\mathbb{R}^n, \mathbb{C})$ into $L^2_w(\mathbb{R}^n, \mathbb{C}^n)$.

The second author thanks Alan McIntosh for introducing this method during discussions at the recent workshop at the Australian National University, Canberra.

The authors thank Steven Hofmann for his suggestion to work on such a beautiful problem, and for all his support during the process of this research.

Finally, the authors thank the referee for the careful editing and suggestions given.

3. Degenerate elliptic operators

The properties of the degenerate elliptic operator $\mathcal{L}_w$ and the associated semigroup $e^{-t\mathcal{L}_w}$ were developed in detail in [14] and we refer the reader there for complete details. Here we state the key ideas.

Given a weight $w \in A_2$, the space $H^1(w)$ is the weighted Sobolev space that is the completion of C_c^∞ with respect to the norm

$$\|f\|_{H^1(w)} = \left(\int_{\mathbb{R}^n} \left(|f(x)|^2 + |\nabla f(x)|^2 \right) w(x)\, dx \right)^{1/2}.$$

Given a matrix $\mathbf{A} \in \mathcal{E}_n(w, \lambda, \Lambda)$, define $\mathfrak{a}(f, g)$ to be the sesquilinear form

$$\mathfrak{a}(f, g) = \int_{\mathbb{R}^n} \mathbf{A}(x) \nabla f(x) \cdot \overline{\nabla g(x)}\, dx. \tag{3.1}$$

Since $w \in A_2$ and $\mathbf{A}$ satisfies (1.1), $\mathfrak{a}$ is a closed, maximally accretive, continuous sesquilinear form. Therefore, there exists a densely defined operator $\mathcal{L}_w$ on $L^2(w)$ such that for every f in the domain of $\mathcal{L}_w$ and every $g \in L^2(w)$,

$$\mathfrak{a}(f, g) = \langle \mathcal{L}_w f, g \rangle_w = \int_{\mathbb{R}^n} \mathcal{L}_w f(x) \overline{g(x)} w(x)\, dx. \tag{3.2}$$

If f, g are in $\mathcal{C}_c^\infty$ and in the domain of $\mathcal{L}_w$, then integration by parts yields

$$\langle \mathcal{L}_w f, g \rangle_w = \mathfrak{a}(f, g) = \langle \mathcal{L} f, g \rangle,$$

where $\langle\, ,\, \rangle$ is the standard complex inner product on L^2. Thus, at least formally, $\mathcal{L}_w = w^{-1}\mathcal{L} = -w^{-1}\mathrm{div}\mathbf{A}\nabla$.

Further, the properties of the sesquilinear form $\mathfrak{a}$ guarantee that the semigroup $e^{-t\mathcal{L}_w}$ exists. In the special case when $\mathbf{A}$ is real and symmetric, then the heat kernel of $e^{-t\mathcal{L}_w}$ satisfies Condition (G).

Finally, in [14], we proved the following results, which will be needed in our proof.

Lemma 3.1. *Given a matrix $\mathbf{A} \in \mathcal{E}_n(w, \lambda, \Lambda)$, suppose that the heat kernel of the associated semigroup $e^{-t\mathcal{L}_w}$ satisfies Condition (G). Then for all $t > 0$, $e^{-t\mathcal{L}_w} 1 = 1$: that is, for all $x \in \mathbb{R}^n$,*

$$\int_{\mathbb{R}^n} W_t(x, y)\, dy = 1.$$

If $e^{-t\mathcal{L}_w}$ satisfies Gaussian bounds, then its derivative satisfies similar bounds. More precisely, let $V_t = -2t\mathcal{L}_w e^{-t^2\mathcal{L}_w} = \frac{d}{dt} e^{-t^2\mathcal{L}_w}$. Then the following result holds.

Lemma 3.2. *The operator V_t has a kernel $V_t(x, y)$ with the following properties:*

For all x, $y \in \mathbb{R}^n$ and $t > 0$,

$$|V_t(x, y)| \le \frac{C_1}{t^{n+1}} \exp\left(-C_2 \frac{|x - y|^2}{t^2} \right). \tag{3.3}$$

For almost every $x \in \mathbb{R}^n$, $V_t 1 = 0$, that is, for all $x \in \mathbb{R}^n$,

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

There exists $\alpha = \alpha(n, \lambda, \Lambda, w) > 0$ such that for almost every x, y, $\in \mathbb{R}^n$, $2|h| < t + |x|$,

$$|V_t(x,y) - V_t(x,y+h)| + |V_t(x+h,y) - V_t(x,y)| \le \frac{C_1}{t^{n+1}} \left(\frac{|h|}{t+|x|} \right)^\alpha \exp\left(-C_2 \frac{|x-y|^2}{t^2} \right).$$

4. Weighted norm inequalities

Central to our proof is the theory of weighted norm inequalities for classical operators, particularly singular integrals and square functions. In this section we state the results we need; the standard ones are given without proof and we refer the reader to Duoandikoetxea [15], García-Cuerva and Rubio de Francia [19], and Grafakos [20] for complete information.

We begin with the weighted norm inequalities for the Hardy-Littlewood maximal operator, for convolution operators, and for singular integrals.

Lemma 4.1. *Let $w \in A_2$. Then M is bounded on $L^2(w)$ and $\|M\|_{\mathbf{B}(L^2(w))} \leq C(n,[w]_{A_2})$. Furthermore, suppose ϕ and Φ are such that for all x, $|\phi(x)| \leq \Phi(x)$, and Φ is radial, decreasing and integrable. Then the operators $\phi_t * f$ are uniformly bounded on $L^2(w)$; in fact,*

$$\sup_{t>0} |\phi_t * f(x)| \leq C(n)\|\Phi\|_1 Mf(x).$$

Remark 4.2. If ϕ is a Schwartz function, then such a function Φ always exists.

In what follows we denote by $\widehat{f}$ the Fourier transform of f,

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

Lemma 4.3. *Let $K : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ be a locally integrable function such that $\widehat{K} \in L^\infty$ and*

$$|\nabla K(x)| \leq \frac{C}{|x|^{n+1}}, \quad x \neq 0.$$

Then the singular integral

$$Tf(x) = \int_{\mathbb{R}^n} K(x-y)f(y)\,dy$$

and the maximal singular integral

$$T^*f(x) = \sup_{\epsilon>0}\left|\int_{|x-y|>\epsilon} K(x-y)f(y)\,dy\right|$$

are bounded on $L^2(w)$, and $\|T\|_{\mathbf{B}(L^2(w))}$, $\|T^*\|_{\mathbf{B}(L^2(w))} \leq C(n,[w]_{A_2},K)$.

Remark 4.4. An important example of singular integrals are the Riesz transforms:

$$R_jf(x) = c_n\int_{\mathbb{R}^n} \frac{x_j-y_j}{|x-y|^{n+1}} f(y)\,dy, \quad 1 \leq j \leq n,$$

where the constant c_n is chosen so that

$$\widehat{R_jf}(\xi) = -i\frac{\xi_j}{|\xi|}\widehat{f}(\xi).$$

Our next two results are square function inequalities. The first is a weighted version of Carleson's theorem due to Journé [23]. Define the weighted Carleson measure norm of a function γ_t by

$$\|\gamma_t\|_{C,w} = \sup_Q \frac{1}{w(Q)}\int_Q\int_0^{\ell(Q)} |\gamma_t(x)|^2\,\frac{dt}{t}\,w(x)\,dx. \tag{4.1}$$

Lemma 4.5. *Let $w \in A_2$ and suppose γ_t is such that $\|\gamma_t\|_{C,w} < \infty$. Let $p \in C_c^\infty(\mathbb{R}^n)$ be such that p is a non-negative, radial, decreasing function,* $\operatorname{supp}(p) \subset B_1(0)$, *and $\|p\|_1 = 1$. Then for all $f \in L^2(w)$,*

$$\int_0^\infty \int_{\mathbb{R}^n} |(p_t * f)(x)|^2 |\gamma_t(x)|^2 \, w(x)\, dx \, \frac{dt}{t} \le C\|\gamma_t\|_{C,w} \|f\|^2_{L^2(w)}.$$

The second result is a Littlewood-Paley type inequality.

Lemma 4.6. *Given $w \in A_2$, let ψ be a Schwartz function such that $\widehat{\psi}(0) = 0$. Then for all $f \in L^2(w)$,*

$$\int_{\mathbb{R}^n} \int_0^\infty |\psi_t * f(x)|^2 \, \frac{dt}{t} \, w(x)\, dx \le C(n, \psi, [w]_{A_2}) \|f\|^2_{L^2(w)}. \tag{4.2}$$

Proof. A direct proof of Lemma 4.6 is given by Wilson [33]. Here we sketch a proof that is implicitly based on the idea that

$$g_\psi(f)(x) = \left(\int_0^\infty |\psi_t * f(x)|^2 \, \frac{dt}{t} \right)^{1/2}$$

can be regarded as a vector-valued singular integral.

By a standard argument in the theory of weighted norm inequalities, it will suffice to prove that for all $0 < \delta < 1$ there exists a constant C_δ such that for $x \in \mathbb{R}^n$,

$$M^\#(g_\psi(f)^\delta)(x) \le C_\delta M f(x)^\delta, \tag{4.3}$$

where M is the Hardy-Littlewood maximal operator and $M^\#$ is the sharp maximal operator of Fefferman and Stein.

The proof of inequality (4.3) is readily obtained by adapting the argument in Cruz-Uribe and Pérez [13, Lemma 1.6] for the g_λ^* operator. The changes are straightforward, so here we only indicate the key steps. (Also see Álvarez and Pérez [1].) Our assumptions on ψ guarantee that g_ψ is bounded on L^2 and is weak $(1,1)$. (See [19, p. 505].) Therefore, we only have to prove that if $|x| > |h|/2$, then

$$\left(\int_0^\infty |\psi_t(x+h) - \psi_t(x)|^2 \frac{dt}{t} \right)^{1/2} \le C \frac{|h|^{1/2}}{|x|^{n+1/2}}. \tag{4.4}$$

This is the vector-valued analog of the gradient condition in [13, Lemma 1.6]. To prove (4.4), note that since ψ is a Schwartz function, it is bounded. Hence, by the mean value theorem, for each x and t there exists θ, $0 < \theta < 1$, such that

$$\begin{aligned}
\int_0^\infty |\psi_t(x+h) - \psi_t(x)|^2 \frac{dt}{t} &= \int_0^\infty |\psi\left(\frac{x+h}{t}\right) - \psi\left(\frac{x}{t}\right)|^2 \frac{dt}{t^{2n+1}} \\
&\le C \int_0^\infty |\psi\left(\frac{x+h}{t}\right) - \psi\left(\frac{x}{t}\right)| \frac{dt}{t^{2n+1}} \\
&\le C|h| \int_0^\infty |\nabla\psi\left(\frac{x+\theta h}{t}\right)| \frac{dt}{t^{2n+2}} \\
&= C|h| \int_0^{|x|} + C|h| \int_{|x|}^\infty .
\end{aligned}$$

Since $|x + \theta h| > |x|/2$ and $|\nabla\psi(y)| \le C|y|^{-2n-2}$,

$$|h| \int_0^{|x|} |\nabla\psi\left(\frac{x+\theta h}{t}\right)| \frac{dt}{t^{2n+2}} \le C|h| \int_0^{|x|} |x|^{-2n-2}\, dt = C \frac{|h|}{|x|^{2n+1}}.$$

We estimate the second integral in the same way, using that $|\nabla\psi(x)| \leq C|x|^{-2n}$. Taking the square root we get (4.4). □

The next proposition is a key estimate in our proof of Theorem 1.1. It yields a square function estimate given the size and regularity assumptions on the kernel of the operator. In the unweighted case, this result can be found in, for example, Grafakos [20, p. 643] or Hofmann [21]; in a somewhat different form it can be found in Auscher and Tchamitchian [6].

Proposition 4.7. *Let $w \in A_2$ and let ψ be a radial Schwartz function such that $\widehat{\psi}(0) = 0$ and*

$$\int_0^\infty \widehat{\psi}(t)^2 \frac{dt}{t} = 1. \tag{4.5}$$

Let $Q_t f(x) = \psi_t * f(x)$. Given a family of sublinear operators $\{R_t\}$, suppose that each R_t is bounded on $L^2(w)$, and for all t, $s > 0$ the composition $R_t Q_s$ is bounded on $L^2(w)$ and for some $\alpha > 0$,

$$\|R_t Q_s\|_{\mathcal{B}(L^2(w))} \leq K \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha. \tag{4.6}$$

Then the family $\{R_t\}$ satisfies the square function estimate

$$\int_0^\infty \int_{\mathbb{R}^n} |R_t f(x)|^2 w(x)\, dx \frac{dt}{t} \leq K \cdot C(n, \psi, \alpha, [w]_{A_2}) \|f\|^2_{L^2(w)}. \tag{4.7}$$

The proof of Proposition 4.7 requires a weighted version of the Calderón reproducing formula given by Wilson [33].

Lemma 4.8. *For all $w \in A_2$ and $f \in L^2(w)$,*

$$\int_0^\infty Q_t^2 f(x) \frac{dt}{t} = f(x),$$

where this equality is understood as follows: for each $j > 1$, let B_j be the ball centered at 0 of radius j, and define the function

$$f_j(x) = \int_{1/j}^j Q_t(\chi_{B_j} Q_t f)(x) \frac{dt}{t}.$$

Then for each j, $f_j \in L^2(w)$ and $\{f_j\}$ converges to f in $L^p(w)$.

Proof of Proposition 4.7. Fix $f \in L^2(w)$ and let f_j be as in Lemma 4.8. Since R_t is bounded on $L^2(w)$, we have that for each $t > 0$,

$$\int_{\mathbb{R}^n} |R_t f(x)|^2 w(x)\, dx = \lim_{j\to\infty} \int_{\mathbb{R}^n} |R_t f_j(x)|^2 w(x)\, dx.$$

Since each R_t is sublinear, we have that

$$|R_t f_j(x)| \leq \int_{1/j}^j |R_t Q_s(\chi_{B_j} Q_s f)(x)| \frac{ds}{s}.$$

Therefore, by Fatou's lemma, Minkowski's inequality, and (4.6),

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}^n} |R_t f(x)|^2 w(x)\,dx\,\frac{dt}{t} \\
&\quad \le \liminf_{j\to\infty} \int_0^\infty \int_{\mathbb{R}^n} |R_t f_j(x)|^2 w(x)\,dx\,\frac{dt}{t} \\
&\quad \le \liminf_{j\to\infty} \int_0^\infty \int_{\mathbb{R}^n} \left(\int_{1/j}^j |R_t Q_s(\chi_{B_j} Q_s f)(x)| \frac{ds}{s} \right)^2 w(x)\,dx\,\frac{dt}{t} \\
&\quad \le \liminf_{j\to\infty} \int_0^\infty \left(\int_{1/j}^j \left(\int_{\mathbb{R}^n} |R_t Q_s(\chi_{B_j} Q_s f)(x)|^2 w(x)\,dx \right)^{1/2} \frac{ds}{s} \right)^2 \frac{dt}{t} \\
&\quad \le \liminf_{j\to\infty} K \int_0^\infty \left(\int_{1/j}^j \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \|\chi_{B_j} Q_s f\|_{L^2(w)} \frac{ds}{s} \right)^2 \frac{dt}{t} \\
&\quad \le K \int_0^\infty \left(\int_0^\infty \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \|Q_s f\|_{L^2(w)} \frac{ds}{s} \right)^2 \frac{dt}{t}.
\end{aligned}$$

For all $s > 0$,

$$\int_0^\infty \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \frac{dt}{t} = \int_0^\infty \min\left(u, \frac{1}{u}\right)^\alpha \frac{du}{u} = C(\alpha) < \infty,$$

and the same is true if we reverse the roles of s and t. Therefore, if we apply Schwartz' inequality, Fubini's theorem and Lemma 4.6 we get that

$$\begin{aligned}
&\int_0^\infty \left(\int_0^\infty \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \|Q_s f\|_{L^2(w)} \frac{ds}{s} \right)^2 \frac{dt}{t} \\
&\quad \le C(\alpha) \int_0^\infty \left(\int_0^\infty \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \frac{ds}{s} \right) \left(\int_0^\infty \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \|Q_s f\|^2_{L^2(w)} \frac{ds}{s} \right) \frac{dt}{t} \\
&\quad \le C(\alpha) \int_0^\infty \left(\int_0^\infty \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha \frac{dt}{t} \right) \|Q_s f\|^2_{L^2(w)} \frac{ds}{s} \\
&\quad \le C(\alpha, \psi, n, [w]_{A_2}) \|f\|^2_{L^2(w)}.
\end{aligned}$$

□

Operator norm bounds such as those in (4.6) can generally be deduced in the unweighted case using the Fourier transform or kernel estimates. The following result is related to the square function estimates techniques in [12] (see also [30]); it can be found in Grafakos [20, proof of Theorem 8.6.3] in its present form.

Lemma 4.9. *Let $\{T_t\}$, $t > 0$ be a family of integral operators such that $T_t 1 = 0$ and such that the kernels K_t satisfy*

$$|K_t(x,y)| \le \frac{C}{t^n (1 + t^{-1}|x-y|)^{n+1}}, \tag{4.8}$$

$$|K_t(x,y) - K_t(x,y')| \le \frac{C|y-y'|}{t^{n+1}}. \tag{4.9}$$

Then for some $\alpha > 0$,

$$\|T_t Q_s\|_{\mathbf{B}(L^2)} \le C \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha.$$

In order to get this estimate on $L^2(w)$, $w \in A_2$, we use the following clever application of interpolation due to Duoandikoetxea and Rubio de Francia [16].

Lemma 4.10. *Suppose that a sublinear operator T is bounded on $L^2(w)$ for all $w \in A_2$, with $\|T\|_{\mathcal{B}(L^2(w))}$ depending only on $[w]_{A_2}$ and the dimension n. Then for any $w \in A_2$, there exists θ, $0 < \theta < 1$, that depends on $[w]_{A_2}$ such that*

$$\|T\|_{\mathcal{B}(L^2(w))} \leq C(n, [w]_{A_2})\|T\|^{\theta}_{\mathcal{B}(L^2)}.$$

Proof. This is a consequence of the structural properties of A_2 weights and the theory of interpolation with change of measure due to Stein and Weiss [32] (see also Bergh and Löfström [7]). Given $w \in A_2$, there exists $s > 1$, depending only on $[w]_{A_2}$, such that $w^s \in A_2$ with $[w^s]_{A_2}$ depending only on $[w]_{A_2}$. Hence, by hypothesis, $\|T\|_{\mathcal{B}(L^2(w^s))}$ is bounded by a constant that depends only on $[w]_{A_2}$ and n. Choose θ such that $1 - \theta = 1/s$. Then $w = 1^{\theta} w^{s(1-\theta)}$, so by interpolation with change of measure,

$$\|T\|_{\mathcal{B}(L^2(w))} \leq \|T\|^{1-\theta}_{\mathcal{B}(L^2(w^s))}\|T\|^{\theta}_{\mathcal{B}(L^2)} \leq C(n, [w]_{A_2})\|T\|^{\theta}_{\mathcal{B}(L^2)}. \qquad \square$$

5. Two square function inequalities

In this section we prove two weighted square function inequalities. The first is for the operator $V_t = \frac{d}{dt}e^{-t^2\mathcal{L}_w}$ and is used in Sections 7 and 8 below.

Lemma 5.1. *Let $p \in C_c^{\infty}$ be a non-negative, radial, decreasing function such that $\|p\|_1 = 1$ and $\operatorname{supp}(p) \subset B_1(0)$, and for $f \in H^1(w)$ let*

$$G_t(x, y) = f(y) - f(x) - (y - x) \cdot (p_t * \nabla f)(x). \tag{5.1}$$

Then there exists $C > 0$ depending only on n, w and the constants in the Gaussian estimates, such that

$$\int_0^{\infty} \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} V_t(x, y) G_t(x, y)\, dy \right|^2 w(x)\, dx\, \frac{dt}{t} \leq C\|\nabla f\|^2_{L^2(w)}. \tag{5.2}$$

The second square function inequality is needed in the last step of the proof in Section 9. In the unweighted case this inequality is due to Journé [24]; our proof is adapted from that of Auscher and Tchamitchian [6] as explicated by Grafakos [20].

Define the averaging operator A_t by

$$A_t f(x) = \fint_{Q_t(x)} f(y)\, dy,$$

where $Q_t(x)$ is the unique dyadic cube containing x such that $t \leq \ell(Q_t(x)) < 2t$. If $2^{-k-1} < t \leq 2^{-k}$, and Q_j^k denotes the collection of dyadic cubes of side length 2^{-k}, then the kernel of A_t is

$$A_t(x, y) = \sum_j |Q_j^k|^{-1} \chi_{Q_j^k}(x) \chi_{Q_j^k}(y).$$

It follows immediately from the definition of A_t that $|A_t f(x)| \leq Mf(x)$, and so by Lemma 4.1, A_t is bounded on $L^2(w)$ for all $w \in A_2$ with a constant independent of t.

Define the operator $P_t f(x) = p_t * f(x)$, where $p_t(x) = t^{-n}p(x/t)$ and p is a non-negative, radial function such that $\operatorname{supp}(p) \subset B_1(0)$ and $\|p\|_1 = 1$. Then again by Lemma 4.1, $|P_t f(x)| \leq C(n)Mf(x)$ and P_t is uniformly bounded on $L^2(w)$, $w \in A_2$.

Lemma 5.2. *Given $w \in A_2$, there exists a constant C such that for all $f \in L^2(w)$,*

$$\int_0^\infty \int_{\mathbb{R}^n} |(P_t - A_t)f(x)|^2 w(x)\,dx \frac{dt}{t} \le C \int_{\mathbb{R}^n} |f(x)|^2 w(x)\,dx. \tag{5.3}$$

Remark 5.3. Though Lemma 5.2 is stated in terms of the dyadic grid, it will be clear from the proof that it is true if we replace the dyadic grid by the "dyadic" grid relative to a fixed cube Q: that is, the collection of cubes obtained as in the construction of the standard dyadic grid, but starting with Q instead of $[0,1)^n$.

Proof of Lemma 5.1. The proof requires several steps. First, we will show that there exists a family of sublinear operators $\{R_t^k\}$, $k \ge 0$, such that (5.2) holds provided that there exists $A > 1$ such that

$$\int_0^\infty \int_{\mathbb{R}^n} |R_t^k F(x)|^2\, w(x)\,dx\, \frac{dt}{t} \le CA^k \|F\|^2_{L^2(w)}, \tag{5.4}$$

where

$$F(x) = \vec{R} \cdot \nabla f(x) := \sum_{j=1}^n R_j \left(\frac{\partial f}{\partial x_j}\right)(x); \tag{5.5}$$

here $\vec{R} = (R_1, \dots, R_n)$, where R_j are the Riesz transforms. This square function estimate will follow from Proposition 4.7 if we can prove that the operators R_t^k are uniformly bounded on $L^2(w)$ and satisfy two operator norm estimates. Let ψ be a radial Schwartz function such that $\widehat{\psi}(0) = 0$ and such that (4.5) holds, and let $Q_s f = \psi_s * f$. We will first show that for all s, $t > 0$,

$$\|R_t^k Q_s\|_{\mathbf{B}(L^2(w))} \le C(n, [w]_{A_2}, \psi) \tag{5.6}$$

and then show the stronger estimate

$$\|R_t^k Q_s\|_{\mathbf{B}(L^2(w))} \le C(n, [w]_{A_2}, p, \psi) A^k \min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha. \tag{5.7}$$

Reduction to (5.4)**.** For G_t given by (5.1), Lemma 3.2 yields

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n} V_t(x,y) G_t(x,y)\,dy\right| \\
&\quad \le C_1 t^{-n-1} \int_{\mathbb{R}^n} \exp\left(-C_2 \frac{|x-y|^2}{t^2}\right) |G_t(x,y)|\,dy \\
&\quad = C_1 t^{-n-1} \int_{|x-y|<t} \exp\left(-C_2 \frac{|x-y|^2}{t^2}\right) |G_t(x,y)|\,dy \\
&\qquad + C_1 t^{-n-1} \sum_{k=1}^\infty \int_{2^{k-1}t \le |x-y| < 2^k t} \exp\left(-C_2 \frac{|x-y|^2}{t^2}\right) |G_t(x,y)|\,dy.
\end{aligned}$$

In the first integral, make the change of variables $h = (y-x)/t$; in the integrals in the sum, make the change of variables $h = (y-x)/(2^k t)$. Then there exist positive

constants B_1 and B_2 such that

$$\left|\int_{\mathbb{R}^n} V_t(x,y)G_t(x,y)\,dy\right| \leq B_1 t^{-1}\sum_{k=0}^{\infty} 2^{nk}\exp(-B_2 4^k)\int_{|h|<1}|G_t(x,x+2^k ht)|\,dh.$$

Note that for any $B_2 > 0$,

$$\sum_{k=0}^{\infty} 2^{nk}\exp(-B_2 4^k) < \infty.$$

Therefore, by Hölder's inequality we get the following estimate:

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n} V_t(x,y)G_t(x,y)\,dy\right|^2 w(x)\,dx\,\frac{dt}{t}\\
&\quad\leq C\int_0^\infty\int_{\mathbb{R}^n}\left(t^{-1}\sum_{k=0}^{\infty}2^{nk}\exp(-B_2 4^k)\int_{|h|<1}|G_t(x,x+2^k ht)|\,dh\right)^2 w(x)\,dx\,\frac{dt}{t}\\
&\quad\leq C\sum_{k=0}^{\infty}2^{nk}\exp(-B_2 4^k)\int_0^\infty\int_{\mathbb{R}^n}\left(t^{-1}\int_{|h|<1}|G_t(x,x+2^k ht)|\,dh\right)^2 w(x)\,dx\,\frac{dt}{t};
\end{aligned}$$

if we make the change of variables $t \mapsto t_k := 2^{-k}t$, we get

$$\begin{aligned}
&\leq C\sum_{k=0}^{\infty}2^{(n+2)k}\exp(-B_2 4^k)\\
&\qquad\times\int_0^\infty\int_{\mathbb{R}^n}\left(t^{-1}\int_{|h|<1}|G_{t_k}(x,x+ht)|\,dh\right)^2 w(x)\,dx\,\frac{dt}{t}.
\end{aligned}\tag{5.8}$$

Assume for the moment that for each k there exists a family of sublinear operators $\{R_t^k\}$ such that

$$R_t^k F(x) = t^{-1}\int_{|h|<1}|G_{t_k}(x,x+ht)|\,dh,\tag{5.9}$$

where $F = \vec{R}\cdot\nabla f$ is given by (5.5), and that there exists $A > 1$ such that (5.4) holds. By Lemma 4.3, $\|F\|_{L^2(w)} \leq C(n,[w]_{A_2})\|\nabla f\|_{L^2(w)}$, and so if we combine (5.8) and (5.4) we get inequality (5.2).

To construct the operators R_t^k and show that (5.9) holds, recall that the Riesz potential I_1 is the convolution operator with kernel $c_n|x|^{1-n}$, where the constant c_n is chosen so that

$$\widehat{I_1 f}(\xi) := c_n\widehat{|x|^{1-n}} * f = (2\pi|\xi|)^{-1}\widehat{f}(\xi).\tag{5.10}$$

The following identities are well known (see, for instance, Stein [31]): if f is a Schwartz function and $F = \vec{R}\cdot\nabla f$, then $f = I_1 F$. Further, given any $h \in \mathbb{R}^n$, $h\cdot\nabla f = -\vec{R}\cdot(hF)$.

Define the convolution kernel

$$J_t^h(x) = c_n t^{-1}\left(|x+ht|^{1-n} - |x|^{1-n}\right);\tag{5.11}$$

then we have that

$$\begin{aligned} t^{-1}G_{t_k}(x, x+ht) &= \frac{f(x+ht)-f(x)}{t} - p_{t_k} * (h\cdot\nabla f)(x) \\ &= \left(J_t^h * F\right)(x) + p_{t_k} * (\vec{R}\cdot(hF))(x). \end{aligned} \tag{5.12}$$

Therefore, if we define

$$R_t^k F(x) = \int_{|h|<1} \left|\left(J_t^h * F\right)(x) + p_{t_k} * (\vec{R}\cdot(hF))(x)F(x)\right| dh,$$

then (5.9) holds. Since the operators inside the absolute value in the integrand above are linear, each R_t^k is sublinear.

Proof of inequality (5.6). Since by Lemma 4.1 the operators Q_s are uniformly bounded on $L^2(w)$, to prove (5.6) it will suffice to prove that for all t and k,

$$\|R_t^k\|_{\mathbf{B}(L^2(w))} \le C(n, [w]_{A_2}, \psi). \tag{5.13}$$

By definition, for all Schwartz functions g we have that

$$|R_t^k g(x)| \le \int_{|h|<1} |(J_t^h * g)(x)|\, dh + \int_{|h|<1} |p_{t_k} * (\vec{R}\cdot(hg))(x)|\, dh,$$

where J_t^h is given by (5.11), and $\vec{R}$ is defined in (5.5). To prove (5.13) we will prove that each term on the right-hand side is uniformly bounded on $L^2(w)$. The boundedness of the second term follows immediately from Lemmas 4.1 and 4.3: since $\|p\|_1 = 1$, for all k and t we have that

$$\begin{aligned} &\int_{\mathbb{R}^n} \left(\int_{|h|<1} |p_{t_k} * (\vec{R}\cdot(hg))(x)|\, dh\right)^2 w(x)\, dx \\ &\qquad \le C(n) \int_{|h|<1}\int_{\mathbb{R}^n} |p_{t_k} * (\vec{R}\cdot(hg))(x)|^2 w(x)\, dx\, dh \\ &\qquad \le C(n, [w]_{A_2}) \int_{|h|<1}\int_{\mathbb{R}^n} |\vec{R}\cdot(hg)(x)|^2 w(x)\, dx\, dh \\ &\qquad \le C(n, [w]_{A_2}) \int_{|h|<1}\int_{\mathbb{R}^n} |hg(x)|^2 w(x)\, dx\, dh \\ &\qquad \le C(n, [w]_{A_2}) \int_{\mathbb{R}^n} |g(x)|^2 w(x)\, dx. \end{aligned}$$

We will now prove that the first term is bounded on $L^2(w)$:

$$\int_{\mathbb{R}^n} \left(\int_{|h|<1} |J_t^h * g(x)|\, dh\right)^2 w(x)\, dx \le C(n, [w]_{A_2}) \|g\|_{L^2(w)}^2. \tag{5.14}$$

To prove this we first estimate the inner integral on the left-hand side:

$$\begin{aligned}\int_{|h|<1} |J_t^h * g(x)|\,dh &\le \int_{|h|<1} \left| \left(J_t^h + c_n(n-1)\frac{h\cdot x}{|x|^{n+1}}\chi_{\{|x/t|>2\}}\right) * g(x)\right| dh \\ &\quad + \int_{|h|<1} \left| c_n(n-1)\frac{h\cdot x}{|x|^{n+1}}\chi_{\{|x/t|>2\}} * g(x)\right| dh \\ &\le \left(\int_{|h|<1} \left| J_t^h + c_n(n-1)\frac{h\cdot x}{|x|^{n+1}}\chi_{\{|x/t|>2\}}\right| dh\right) * |g|(x) \\ &\quad + C_n \sum_{i=1}^n \left| \frac{x_i}{|x|^{n+1}}\chi_{\{|x/t|>2\}} * g(x)\right| \\ &\le (L_t * |g|)(x) + C(n)\sum_{i=1}^n R_i^* g(x),\end{aligned}$$

where

$$L(x) = \int_{|h|<1} \left| J_1^h(x) + c_n(n-1)\frac{h\cdot x}{|x|^{n+1}}\chi_{|x|>2}(x)\right| dh,$$

c_n is the constant in the Riesz potential (5.10) and R_i^* is the maximal singular integral associated with the Riesz transform R_i.

By Lemma 4.3,

$$\int_{\mathbb{R}^n} |\sum_{i=1}^n R_i^* g(x)|^2 w(x)\,dx \le C(n,[w]_{A_2})\|g\|_{L^2(w)}^2.$$

Therefore, to complete the proof of (5.14) we need to show that

$$\int_{\mathbb{R}^n} |(L_t * |g|)(x)|^2 w(x)\,dx \le C(n,[w]_{A_2})\|g\|_{L^2(w)}^2.$$

But by Lemma 4.1 it will suffice to prove that

$$|L(x)| \le C(n)\min\left(\frac{1}{|x|^{n-1}}, \frac{1}{|x|^{n+1}}\right), \tag{5.15}$$

since the right-hand side is a radial, decreasing function in L^1. To prove this estimate we treat several cases depending on the size of x.

Case 1: $|x| > 2$. In this case, since $h\cdot\nabla|x|^{1-n} = c_n(n-1)h\cdot x/|x|^{n+1}$, if we apply the mean value theorem twice we get that

$$\begin{aligned}|L(x)| &= c_n \int_{|h|<1} ||x+h|^{1-n} - |x|^{1-n} + (n-1)\frac{h\cdot x}{|x|^{n+1}}|\,dh \\ &\le C(n)\int_{|h|<1} \frac{1}{|x|^{n+1}}\,dh \le \frac{C(n)}{|x|^{n+1}}.\end{aligned}$$

Case 2: $|x| \le 2$. In this case we have that

$$|L(x)| \le c_n \sum_{k=0}^\infty \int_{\{2^{-k-1}\le|h|<2^{-k}\}} ||x+h|^{1-n} - |x|^{1-n}|\,dh. \tag{5.16}$$

Fix k. Then we consider the following sub-cases.

Sub-case 2.1: $|x| > 2^{-k+1}$. In this case, $|x|/2 > |h|$. Hence, $|x+h| > |x|/2$, so

$$c_n \int_{\{2^{-k-1} \le |h| < 2^{-k}\}} \big||x+h|^{1-n} - |x|^{1-n}\big|\, dh \le C(n) \frac{2^{-kn}}{|x|^{n-1}}. \tag{5.17}$$

Sub-case 2.2: $|x| < 2^{-k-2}$. In this case, we have that $|x+h| > |h|/2 > 2^{-k-2}$, so we can argue as in the previous case to get (5.17).

Sub-case 2.3: $2^{-k-2} \le |x| \le 2^{-k+1}$. In this case we estimate as follows:

$$\begin{aligned} c_n \int_{\{2^{-k-1} \le |h| < 2^{-k}\}} & \big||x+h|^{1-n} - |x|^{1-n}\big|\, dh \\ &= \frac{c_n}{|x|^{n-1}} \int_{\{2^{-k-1} \le |h| < 2^{-k}\}} \Big|\big|\frac{x}{|x|} + \frac{h}{|x|}\big|^{1-n} - 1\Big|\, dh \end{aligned}$$

and if we make the change of variables $u = h/|x|$, we get

$$\begin{aligned} &= \frac{c_n |x|^n}{|x|^{n-1}} \int_{\{\frac{2^{-k-1}}{|x|} \le |u| < \frac{2^{-k}}{|x|}\}} \Big|\big|\frac{x}{|x|} + u\big|^{1-n} - 1\Big|\, du \\ &\le \frac{C_n 2^{-kn}}{|x|^{n-1}} \int_{\{\frac{1}{4} \le |u| < 4\}} \Big|\big|\frac{x}{|x|} + u\big|^{1-n} - 1\Big|\, du \\ &= \frac{C_n 2^{-kn}}{|x|^{n-1}} \int_{\{\frac{1}{4} \le |u| < 4\}} \big||1+u|^{1-n} - 1\big|\, du, \end{aligned}$$

where the last equality holds by rotational symmetry. Note that this last integral is finite and its value depends on n but is independent of k and x.

If we apply these three sub-cases to (5.16), we get that for all $|x| \le 2$,

$$|L(x)| \le \sum_{k=0}^{\infty} \frac{C(n) 2^{-kn}}{|x|^{n-1}} \le \frac{C(n)}{|x|^{n-1}}.$$

This completes our proof of (5.15), and so of (5.14). Therefore, we have shown that inequality (5.13) holds.

Proof of inequality (5.7). By Lemma 4.10 and inequality (5.6), to prove (5.7) it will suffice to show the corresponding unweighted norm estimate:

$$\|R_t^k Q_s\|_{\mathbf{B}(L^2)} \le C(n, [w]_{A_2}, p, \psi) A^k \min\left(\frac{t}{s}, \frac{s}{t}\right). \tag{5.18}$$

Fix $s, t > 0$, $k \ge 0$, and recall that $Q_s f = \psi_s * f$. By Hölder's inequality and Plancherel's theorem,

$$\begin{aligned} &\int_{\mathbb{R}^n} |R_t^k Q_s g(x)|^2\, dx \\ &\quad \le C(n) \int_{|h|<1} \int_{\mathbb{R}^n} |J_t^h * \psi_s * g(x) + p_{t_k} * (\vec{R} \cdot (h(\psi_s * g))(x)|^2\, dx\, dh \\ &\quad = C(n) \int_{|h|<1} \int_{\mathbb{R}^n} \left| \widehat{J_t^h}(\xi) \widehat{\psi}(s\xi) \widehat{g}(\xi) - i\, \widehat{p}(t_k \xi) \frac{\xi \cdot h}{|\xi|} \widehat{\psi}(s\xi) \widehat{g}(\xi) \right|^2 d\xi\, dh \\ &\quad = C(n) \int_{|h|<1} \int_{\mathbb{R}^n} \left| \frac{e^{2\pi i t h \cdot \xi} - 1 - 2\pi i \widehat{p}(t_k \xi) \xi \cdot h}{2\pi t |\xi|} \right|^2 |\widehat{\psi}(s\xi)|^2 |\widehat{g}(\xi)|^2\, d\xi\, dh, \end{aligned}$$

where we recall that $t_k = 2^{-k}t$. To estimate the last term, we will use the following: since ψ is a Schwartz function such that $\widehat{\psi}(0) = 0$,

$$|\widehat{\psi}(s\xi)| \leq C(\psi)\min(|s\xi|, |s\xi|^{-1}).$$

Since $p \in C_c^\infty$,

$$|\widehat{p}(t_k\xi)| \leq \frac{C(p)2^k}{t|\xi|}.$$

Since $\widehat{p}(0) = 1$,

$$e^{2\pi i t h\cdot\xi} - 1 - 2\pi i\,\widehat{p}(t_k\xi)\xi\cdot h = O(|t\xi|^2).$$

Therefore, we have that

$$\left|\frac{e^{2\pi i t h\cdot\xi} - 1 - 2\pi i\widehat{p}(t_k\xi)\xi\cdot h}{2\pi t|\xi|}\right| \leq C(p)2^k\min(|t\xi|, |t\xi|^{-1}).$$

Combining these estimates we get that

$$\begin{aligned}
&\int_{\mathbb{R}^n} |R_t^k Q_s g(x)|^2\,dx \\
&\quad\leq C(n,p,\psi)4^k\int_{|h|<1}\int_{\mathbb{R}^n}\min(|t\xi|, |t\xi|^{-1})^2\min(|s\xi|, |s\xi|^{-1})^2|\widehat{g}(\xi)|^2\,d\xi\,dh \\
&\quad\leq C(n,p,\psi)4^k\min\left(\frac{s}{t}, \frac{t}{s}\right)^2\int_{|h|<1}\int_{\mathbb{R}^n}|\widehat{g}(\xi)|^2\,d\xi\,dh \\
&\quad\leq C(n,p,\psi)4^k\min\left(\frac{s}{t}, \frac{t}{s}\right)^2\int_{\mathbb{R}^n}|g(x)|^2\,dx.
\end{aligned}$$

Inequality (5.18) follows immediately. This finishes the proof of Lemma 5.1.

Proof of Lemma 5.2. For brevity, let $R_t = P_t - A_t$. Then we have that

$$R_t f = R_t(I - P_t)f + R_t P_t f,$$

so it will suffice to prove that (5.3) holds with R_t replaced by $R_t(I - P_t)$ and $R_t P_t$. By our remarks above, it is clear that both of these operators are uniformly bounded on $L^2(w)$. Therefore, by Lemma 4.10 and Proposition 4.7, it will suffice to show that for all s, $t > 0$, there exist constants C, $\alpha > 0$ such that

$$\|R_t(I - P_t)Q_s\|_{L^2} \leq C\min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha, \tag{5.19}$$

$$\|R_t P_t Q_s\|_{L^2} \leq C\min\left(\frac{t}{s}, \frac{s}{t}\right)^\alpha, \tag{5.20}$$

where $Q_s f = \psi_s * f(x)$ with ψ a radial Schwartz function with $\widehat{\psi}(0) = 0$.

We first prove (5.19). Since R_t is uniformly bounded on L^2,

$$\|R_t(I - P_t)Q_s\|_{L^2} \leq C\|(I - P_t)Q_s\|_{L^2}.$$

We can bound the right-hand side using Plancherel's theorem. By our choice of ψ and since $\hat{p}(0) = 1$,

$$|\hat{\psi}(s\xi)| \leq \frac{C(\psi)}{|s\xi|}, \qquad |1 - \hat{p}(t\xi)| \leq C|t\xi|.$$

Therefore,

$$\begin{aligned}\int_{\mathbb{R}^n} |(I-P_t)Q_s f(x)|^2\,dx &= \int_{\mathbb{R}^n} |(1-\hat{p}(t\xi)|^2|\hat{\psi}(s\xi)^2||\hat{f}(\xi)|^2\,d\xi \\ &\le C\left(\frac{t}{s}\right)^2 \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2\,d\xi = C\left(\frac{t}{s}\right)^2 \|f\|_2^2,\end{aligned}$$

and so

$$\|R_t(I-P_t)Q_s\|_{L^2} \le C\left(\frac{t}{s}\right).$$

On the other hand, since convolution operators commute and $I-P_t$ is uniformly bounded on L^2,

$$\|R_t(I-P_t)Q_s\|_{L^2} = \|R_tQ_s(I-P_t)\|_{L^2} \le C\|R_tQ_s\|_{L^2}.$$

For any α, $0<\alpha<1/2$, there exists a constant C such that

$$\|A_tQ_s\|_{L^2} \le C\left(\frac{s}{t}\right)^{\alpha}.$$

(See Grafakos [20].) Further, since

$$|\hat{p}(t\xi)| \le \frac{C}{|t\xi|}, \qquad |\hat{\psi}(s\xi)| \le C|s\xi|,$$

we can again use Plancherel's theorem to see that

$$\|P_tQ_s\|_{L^2} \le C\left(\frac{s}{t}\right).$$

It follows that for $\alpha<1/2$,

$$\|R_tQ_s\|_{L^2} \le C\left(\frac{s}{t}\right)^{\alpha},$$

and so we get that (5.19) holds.

To prove (5.20) we will apply Lemma 4.9. It will suffice to show that $R_tP_t(1)=0$, and that the kernel K_t of the operator R_tP_t satisfies (4.8) and (4.9). The identity is immediate: both A_t and P_t are bounded on L^∞ and $A_t(1)=P_t(1)=1$.

Let $K_t = J_t + L_t$, where J_t is the kernel of P_t^2 and L_t is the kernel of A_tP_t, that is,

$$J_t(x,y) = \int_{\mathbb{R}^n} p_t(x-z)p_t(y-z)\,dz,$$

$$L_t(x,y) = \int_{\mathbb{R}^n} \sum_j |Q_j^k|\chi_{Q_j^k}(x)\chi_{Q_j^k}(z)p_t(y-z)\,dz.$$

It is immediate from these expressions that there exists a constant $c>0$ such that both $J_t(x,y)$ and $L_t(x,y)$ are non-zero only if $|x-y|<ct$. Further, we have that

$$|J_t(x,y)| \le t^{-n}\|p\|_{L^\infty}\|p\|_{L^1} \le Ct^{-n}.$$

Similarly, since $|Q_j^k| \approx t^{-n}$,

$$|L_t(x,y)| \le Ct^{-n}\|p\|_{L^\infty}.$$

Inequality (4.8) follows at once.

The proof of (4.9) is similar. By the mean value theorem,

$$|p_t(y-z)-p_t(y'-z)| \le t^{-n}\|\nabla p\|_{L^\infty}\left|\frac{y-y'}{t}\right| = \frac{C|y-y'|}{t^{n+1}}.$$

If we use this to estimate $|J_t(x,y)-J_t(x,y')|$ and $|L_t(x,y)-L_t(x,y')|$ and argue as before, we get that both satisfy a similar bound. Inequality (4.9) follows at once. Therefore, we have proved (5.20) and our proof is complete.

6. Reduction to a square function estimate

In this section we begin our proof of Theorem 1.1 by proving that inequality (1.6) holds if we have the square function estimate

$$\left(\int_0^\infty \int_{\mathbb{R}^n} |V_t f(x)|^2\, w(x)\, dx\, \frac{dt}{t}\right)^{1/2} \le C\|\nabla f\|_{L^2(w)}, \tag{6.1}$$

where $V_t = -2t\mathcal{L}_w e^{-t^2\mathcal{L}_w}$. Fix $f \in D(\mathcal{L}_w)$; recall that $D(\mathcal{L}_w)$ is dense in $H_0^1(w)$ (see [14]). If we apply integration by parts to a well-known formula (see, for instance, Kato [26]) we get that

$$\mathcal{L}_w^{1/2} f = \int_0^\infty t^3 e^{-2t^2\mathcal{L}_w}\mathcal{L}_w^2 f\, \frac{dt}{t}. \tag{6.2}$$

Therefore, for all $g \in L^2(w)$,

$$\begin{aligned}
|\langle \mathcal{L}_w^{1/2} f, g\rangle_w|^2 &= \left|\int_0^\infty \left\langle t^3 e^{-2t^2\mathcal{L}_w}\mathcal{L}_w^2 f, g\right\rangle_w \frac{dt}{t}\right|^2 \\
&= \left|\int_0^\infty \left\langle t e^{-t^2\mathcal{L}_w}\mathcal{L}_w f, t^2\mathcal{L}_w^* e^{-t^2\mathcal{L}_w^*} g\right\rangle_w \frac{dt}{t}\right|^2 \\
&\le \left(\int_0^\infty \int_{\mathbb{R}^n} |t e^{-t^2\mathcal{L}_w}\mathcal{L}_w f(x)|^2\, w(x)\, \frac{dt}{t}\right) \\
&\qquad \times \left(\int_0^\infty \int_{\mathbb{R}^n} |t^2\mathcal{L}_w^* e^{-t^2\mathcal{L}_w^*} g(x)|^2\, w(x)\, \frac{dt}{t}\right).
\end{aligned}$$

Hence, by duality we have shown that (1.6) follows from (6.1) provided that we can prove the square function inequality

$$\int_0^\infty \int_{\mathbb{R}^n} |t^2\mathcal{L}_w^* e^{-t^2\mathcal{L}_w^*} g(x)|^2\, w(x)\, \frac{dt}{t} \le C\|g\|_{L^2(w)}^2.$$

Since the semigroups $e^{-t\mathcal{L}_w^*}$ and $e^{-t\mathcal{L}_w}$ satisfy the same estimates, this is equivalent to proving that

$$\int_0^\infty \int_{\mathbb{R}^n} |tV_t g(x)|^2\, w(x)\, \frac{dt}{t} \le C\|g\|_{L^2(w)}^2. \tag{6.3}$$

To prove this square function estimate we use Proposition 4.7. Let $G(x) = \exp(-C_2|x|^2)$. Then by Lemma 3.2,

$$\begin{aligned}
|tV_t f(x)| &\le \int_{\mathbb{R}^n} |tV_t(x,y)||f(y)|\, dy \\
&\le C_1 \int_{\mathbb{R}^n} t^{-n} \exp\left(-C_2\frac{|x-y|^2}{t^2}\right)|f(y)|\, dy = C_1(G_t * |f|)(x).
\end{aligned}$$

Since $G \in L^1$ and is radial, by Lemma 4.1, $\sup_{t>0}(G_t * |f|)(x) \le CMf(x)$ and the operators tV_t are uniformly bounded on $L^2(w)$. Let ψ be a radial Schwartz function such that $\widehat{\psi}(0) = 0$. For $s > 0$, let $Q_s = \psi_s * f$. Again by Lemma 4.1 we have

that the operators Q_s are uniformly bounded on $L^2(w)$. Therefore, there exists a constant C such that for all s, $t > 0$,

$$\|tV_tQ_s\|_{\mathcal{B}(L^2(w))} \leq C.$$

Further, by Lemmas 3.2 and 4.9 there exists $\beta > 0$ such that

$$\|tV_tQ_s\|_{\mathcal{B}(L^2)} \leq C \min\left(\frac{t}{s}, \frac{s}{t}\right)^{\beta}.$$

Hence, by Lemma 4.10 we have that for some $\alpha > 0$,

$$\|tV_tQ_s\|_{\mathcal{B}(L^2(w))} \leq C \min\left(\frac{t}{s}, \frac{s}{t}\right)^{\alpha}.$$

Therefore, the operators $\{tV_t\}$ satisfy the hypotheses of Lemma 4.7, so (6.3) holds.

7. Reduction to a Carleson measure estimate

In this section we prove that (6.1) holds provided that we have a Carleson measure estimate. More precisely, we will show that if $w \in A_2$, then for all Schwartz functions f,

$$\int_0^\infty \int_{\mathbb{R}^n} |V_tf(x)|^2\, w(x)\, dx\, \frac{dt}{t} \leq C(1 + \|\gamma_t\|_{C,w})\|\nabla f\|^2_{L^2(w)}, \tag{7.1}$$

where $\gamma_t = V_t\phi$, with $\phi(x) = x$.

To prove this, we first show the role played by the Carleson measure estimate. Let $p \in C_c^\infty$ be a non-negative, radial, decreasing function such that $\|p\|_1 = 1$ and $\operatorname{supp}(p) \subset B_1(0)$. By Lemma 3.2, $V_t1 = 0$, so

$$\begin{aligned} V_tf(x) &= \int_{\mathbb{R}^n} V_t(x,y)f(y)\, dy \\ &= \left(\int_{\mathbb{R}^n} V_t(x,y)y\, dy\right) \cdot (p_t * \nabla f)(x) \\ &\qquad + \int_{\mathbb{R}^n} V_t(x,y)\left[f(y) - f(x) - (y-x)\cdot(p_t * \nabla f)(x)\right] dy \\ &= \gamma_t(x)\cdot(p_t * \nabla f)(x) + \int_{\mathbb{R}^n} V_t(x,y)G_t(x,y)\, dy, \end{aligned} \tag{7.2}$$

where $G_t(x,y) = f(y) - f(x) - (y-x)\cdot(p_t * \nabla f)(x)$. The first term in (7.2) satisfies a square function estimate. This follows from the Carleson measure estimate: if $\|\gamma_t\|_{C,w} < \infty$, then by Lemma 4.5,

$$\int_0^\infty \int_{\mathbb{R}^n} |(p_t * \nabla f)(x)|^2 |\gamma_t(x)|^2\, w(x)\, dx\, \frac{dt}{t} \leq C\|\gamma_t\|_{C,w}\|\nabla f\|^2_{L^2(w)}.$$

The proof of (7.1) now follows from Lemma 5.1.

8. The weighted Tb theorem for square roots

We have reduced the proof of Theorem 1.1 to proving that $\gamma_t(x) = V_t\phi(x)$ is a Carleson measure with respect to the weight $w(x)$: that is,

$$\|\gamma_t\|_{C,w} = \sup_Q \frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} |\gamma_t(x)|^2 \frac{dt}{t} w(x)\, dx < \infty, \tag{8.1}$$

where $\gamma_t(x) = t\mathcal{L}_w e^{-t^2\mathcal{L}_w}\varphi(x)$ and $\varphi(x) = x$.

In order to prove this fact we will establish a Tb theorem for square roots, a weighted version of a result due to Auscher and Tchamitchian [6]. For technical reasons we actually need a slightly different theorem, which is given in Lemma 8.2 below. However, it seemed clearer to start with this simpler version and then sketch the modifications needed to prove the full result.

Lemma 8.1. *Suppose that for every cube Q there exists a mapping $F_Q : 5Q \to \mathbb{C}^n$ and a constant $C = C(n, \lambda, \Lambda, w) < \infty$ such that*

(i) $\displaystyle\int_{5Q} |\nabla F_Q(x)|^2 w(x)\,dx \le Cw(Q),$

(ii) $\displaystyle\int_{10Q} |\mathcal{L}_w F_Q(x)|^2 w(x)\,dx \le Cw(Q)\,\ell(Q)^{-2},$

(iii)

$$\|\gamma_t\|_{C,w}^2 \le C\left\{\sup_Q \frac{1}{w(Q)}\int_Q\int_0^{\ell(Q)} |\gamma_t(x)\nabla P_t F_Q(x)|^2 \frac{dt}{t} w(x)\,dx + \sup_{t>0}\|\gamma_t\|_\infty^2\right\},$$

where $F_Q = (F_1, F_2, \dots, F_n) \in \mathbb{C}^n$, and ∇F_Q is the matrix $\left(\frac{\partial}{x_i}F_j\right)_{i,j=1}^n$. Then $\|\gamma_t\|_{C,w} \le C(n, \lambda, \Lambda, w) < \infty$, i.e., γ_t is a Carleson measure with respect to w.

Proof. We follow the proof of the unweighted lemma in [21]. Since the kernel of V_t satisfies the Gaussian bounds (3.3), we have that

$$\sup_{t>0}\|\gamma_t\|_\infty \le C(n, \lambda, \Lambda, w) < \infty. \tag{8.2}$$

Indeed, since V_t has zero moment,

$$\gamma_t(x) = V_t\phi(x) = \int_{\mathbb{R}^n} V_t(x,y)y\,dy = -\int_{\mathbb{R}^n} V_t(x,y)(x-y)\,dy;$$

thus, applying the Gaussian estimates,

$$\begin{aligned} |\gamma_t(x)| &\le \frac{C_1}{t^{n+1}}\int_{\mathbb{R}^n} \exp\left(-C_2\frac{|x-y|^2}{t^2}\right)|x-y|\,dy \\ &= C_1\int_0^\infty\int_{S^{n-1}} \exp\left(-C_2\frac{r^2}{t^2}\right)\left(\frac{r}{t}\right)^n d\omega\,\frac{dr}{t} \\ &\le C_1\left|S^{n-1}\right|\int_0^\infty \exp\left(-C_2 s^2\right)s^n\,ds < \infty, \end{aligned} \tag{8.3}$$

and this bound is independent of t. Hence, given (iii) to prove (8.1) we only need to show that

$$\sup_Q \frac{1}{w(Q)}\int_Q\int_0^{\ell(Q)} |\gamma_t(x)\nabla P_t F_Q(x)|^2 \frac{dt}{t} w(x)\,dx \le C(n, \lambda, \Lambda, w) < \infty. \tag{8.4}$$

Recall that $P_t g = p_t * g$ with $p_t(x) = t^{-n}p\left(\frac{x}{t}\right)$, $p \in \mathcal{C}_0^\infty(B_1(0))$ is radial and $\int p(x)\,dx = 1$. Let $\tilde{\chi}(x) \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ satisfy

$$\operatorname{supp}(\tilde{\chi}) \subset 4Q, \quad 0 \le \tilde{\chi} \le 1, \quad \tilde{\chi}|_{3Q} \equiv 1, \quad |\nabla\tilde{\chi}| \le C\ell(Q)^{-1}.$$

Since p is supported in the unit ball and the gradient operator commutes with P_t, for $x \in Q$ and $0 \leq t \leq \ell(Q)$ we have that

$$\begin{aligned} \nabla P_t F_Q(x) = P_t \nabla F_Q(x) &= \int_{B_t(0)} p_t(y) \left(\nabla \tilde{\chi} F_Q\right)(x-y)\ dy \\ &= P_t \nabla \tilde{\chi} F_Q(x) = P_t \nabla \tilde{\chi}\left(F_Q(x) - c_0\right), \end{aligned} \tag{8.5}$$

where c_0 is a constant to be fixed below. Let $\tilde{F}_Q = F_Q - c_0$; then (8.4) is equivalent to

$$\sup_Q \frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} \left|\gamma_t(x) P_t \nabla \tilde{\chi} \tilde{F}_Q(x)\right|^2 \frac{dt}{t} w(x)\, dx \leq C(n, \lambda, \Lambda, w) < \infty. \tag{8.6}$$

Recall that $\frac{1}{2} V_t = -t \mathcal{L}_w e^{-t^2 \mathcal{L}_w} = -t e^{-t^2 \mathcal{L}_w} \mathcal{L}_w = t e^{-t^2 \mathcal{L}_w} w^{-1} \mathrm{div} \mathbf{A} \nabla$. Define $\theta_t = 2t e^{-t^2 \mathcal{L}_w} w^{-1} \mathrm{div} \mathbf{A}$; then θ_t acts on $n \times n$ matrix-valued functions. If $\mathbf{1}$ is the $n \times n$ identity matrix, $\gamma_t = \theta_t \nabla \phi = \theta_t \mathbf{1}$, so for $x \in Q$ and $0 \leq t \leq \ell(Q)$ we can write

$$\begin{aligned} \gamma_t(x) \nabla P_t \tilde{\chi} \tilde{F}_Q(x) &= \theta_t \mathbf{1}(x) \cdot P_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x) \\ &= \left[\theta_t \mathbf{1}(x) \cdot P_t - \theta_t\right] \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x) + \theta_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x) \\ &= R_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x) + \theta_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x), \end{aligned}$$

where $R_t = \theta_t \mathbf{1}(x) P_t - \theta_t$. We claim that

$$\frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} \left|R_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x)\right|^2 \frac{dt}{t} w(x)\, dx \leq \frac{C}{w(Q)} \int_{5Q} \left|\nabla F_Q\right|^2 w(x)\, dx. \tag{8.7}$$

If this is the case, then by assumption (i),

$$\frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} \left|R_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x)\right|^2 \frac{dt}{t} w(x)\, dx \leq C < \infty.$$

Furthermore, since the operator $e^{-t^2 \mathcal{L}_w}$ is a contraction in $L^2(w)$, by (ii), we have

$$\begin{aligned} &\frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} \left|\theta_t \left(\nabla \tilde{\chi} \tilde{F}_Q\right)(x)\right|^2 \frac{dt}{t} w(x)\, dx \\ &\quad = \frac{4}{w(Q)} \int_Q \int_0^{\ell(Q)} \left|e^{-t^2 \mathcal{L}_w} \mathcal{L}_w \tilde{\chi} \tilde{F}_Q(x)\right|^2 t\, dt\, w(x)\, dx \\ &\quad \leq \frac{4}{w(Q)} \int_0^{\ell(Q)} \int_{5Q} \left|\mathcal{L}_w F_Q(x)\right|^2 w(x)\, dx\, t\, dt \\ &\quad \leq C \ell(Q)^{-2} \int_0^{\ell(Q)} t\, dt \\ &\quad \leq C < \infty. \end{aligned}$$

Therefore, to complete the proof we only have to show that (8.7) holds. Since $\int V_t(x,y)\,dy = 0$,

$$\begin{aligned}
&R_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x)\\
&\quad= \theta_t\mathbf{1}(x)\cdot P_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x) - \theta_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x)\\
&\quad= -2te^{-t^2\mathcal{L}_w}\mathcal{L}_w\phi(x)\cdot P_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x) - 2te^{-t^2\mathcal{L}_w}\mathcal{L}_w\tilde{\chi}\tilde{F}_Q(x)\\
&\quad= \left(\int V_t(x,y)\,y\,dy\right)P_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x) + \int V_t(x,y)\left(\tilde{\chi}\tilde{F}_Q\right)(y)\,dy\\
&\quad= \int V_t(x,y)\left[\left(\tilde{\chi}\tilde{F}_Q\right)(y) - \left(\tilde{\chi}\tilde{F}_Q\right)(x) - (y-x)\,P_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x)\right]dy\\
&\quad= \int V_t(x,y)\,\tilde{G}_t(x,y)\,dy.
\end{aligned}$$

By Lemma 5.1,

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_0^\infty \left|R_t\left(\nabla\tilde{\chi}\tilde{F}_Q\right)(x)\right|^2\frac{dt}{t}w(x)\;dx\\
&\qquad= \int_{\mathbb{R}^n}\int_0^\infty\left|\int V_t(x,y)\,\tilde{G}_t\left(\tilde{F}_Q\right)(x,y)\,dy\right|^2\frac{dt}{t}w(x)\;dx\\
&\qquad\le C\left(n,[w]_{A_2}\right)^2\left\|\nabla\tilde{\chi}\tilde{F}_Q\right\|^2_{L^2(w)}.
\end{aligned}\tag{8.8}$$

We estimate the right-hand side using the product rule and the weighted Poincaré inequality (see [17]); if we fix $c_0 = \frac{1}{w(5Q)}\int_{5Q}F_Q(x)\,dx$, then

$$\begin{aligned}
\|\nabla\tilde{\chi}F_Q\|^2_{L^2(w)} &= \int\left|\nabla\tilde{\chi}\tilde{F}_Q(x)\right|^2 w(x)\,dx\\
&\le \frac{C}{\ell^2(Q)}\int_{5Q}|F_Q(x)-c_0|^2\,w(x)\,dx + C\int_{5Q}|\nabla F_Q(x)|^2\,w(x)\,dx\\
&\le \frac{C\ell^2(Q)}{\ell^2(Q)}\int_{5Q}|\nabla F_Q(x)|^2\,w(x)\,dx + C\int_{5Q}|\nabla F_Q(x)|^2\,w(x)\,dx\\
&\le C\int_{5Q}|\nabla F_Q(x)|^2\,w(x)\,dx.
\end{aligned}$$

This proves (8.7) and our proof is complete. □

In our proof we actually need to replace (iii) with a more complicated criterion; for ease of reference we record this as a separate lemma.

Lemma 8.2. *Suppose that there exists a finite index set $\{\nu\}$ with cardinality $N = N(n,\lambda,\Lambda,w)$, such that for each cube Q there are mappings $F_{Q,\nu}:5Q\to\mathbb{C}$ and a constant $C = C(n,\lambda,\Lambda,w) < \infty$ satisfying*

(I) $\displaystyle\int_{5Q}|\nabla F_{Q,\nu}(x)|^2\,w(x)\,dx \le C\;w(Q)\,,$

(II) $\displaystyle\int_{10Q} |\mathcal{L}_w F_{Q,\nu}(x)|^2 w(x)\,dx \le C w(Q)\,\ell(Q)^{-2}$,

(III) $\displaystyle\|\gamma_t\|_{C,w}^2 \le C \sup_{t>0} \|\gamma_t\|_\infty^2 + C \sum_\nu \frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \nabla P_t F_{Q,\nu}(x)|^2 \frac{dt}{t}\, w(x)\,dx.$

Then $\|\gamma_t\|_{C,w} \le C(n,\lambda,\Lambda,w) < \infty$*: i.e.,* γ_t *is a Carleson measure with respect to* w*.*

The proof is essentially identical to the proof of Lemma 8.1. By (8.3), it is enough to show that

$$\sup_Q \frac{1}{w(Q)} \int_Q \int_0^{\ell(Q)} |\gamma_t(x)\,\nabla P_t F_{Q,\nu}(x)|^2 \frac{dt}{t}\, w(x)\,dx \le C(n,\lambda,\Lambda,w) < \infty,$$

and this is done exactly as in the proof of (8.4).

9. Proof of the weighted Kato theorem

To complete the proof of Theorem 1.1 we will construct a finite index set $\{\nu\}$ and for each cube $Q \subset \mathbb{R}^n$ a family of functions $F_{Q,\nu} : 5Q \to \mathbb{C}$ that satisfy the hypotheses of Lemma 8.2. To do so we adapt the proof of the non-weighted case in [21].

Recall that $\varphi(x) = x$. Given a cube $Q \subset \mathbb{R}^n$ define the function

$$F_Q(x) = e^{-\varepsilon^2 \ell(Q)^2 \mathcal{L}_w} \varphi(x), \tag{9.1}$$

where $\varepsilon > 0$ will be chosen below. The set $\{\nu\}$ will be a finite collection of vectors in $\mathbb{C}^n$, $|\nu| = 1$, also to be chosen below. Define

$$F_{Q,\nu} = F_Q \cdot \bar{\nu}.$$

Since $|F_{Q,\nu}| \le |F_Q|$, and similarly for the gradient, to prove (I) and (II) in Lemma 8.2 it will suffice to prove that F_Q satisfies (i) and (ii) in Lemma 8.1. To prove (ii): from (8.3) we have

$$\left| t\mathcal{L}_w e^{-t^2 \mathcal{L}_w} \varphi(x) \right| = |V_t \varphi(x)| = \left| \int V_t(x,y)\, y\, dy \right| \le C < \infty$$

for some constant independent of t. Then

$$|\mathcal{L}_w F_Q(x)| = \frac{1}{\varepsilon \ell(Q)} \left| V_{\varepsilon \ell(Q)} \varphi(x) \right| = \frac{1}{\varepsilon \ell(Q)} \left| \int V_{\varepsilon \ell(Q)}(x,y)\, y\, dy \right| \le \frac{C}{\varepsilon \ell(Q)}.$$

Since $w \in A_2$, it is a doubling measure, and so

$$\int_{10Q} |\mathcal{L}_w F_Q(x)|^2 w(x)\,dx \le \frac{C}{\varepsilon^2} w(Q)\,\ell(Q)^{-2}. \tag{9.2}$$

This proves (ii).

To prove (i), fix $\eta \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ such that $\eta \equiv 1$ in $5Q$, $\eta \equiv 0$ in $\mathbb{R}^n \setminus 10Q$, $\|\eta\|_\infty \le 1$ and $\|\nabla\eta\|_\infty \le C\ell(Q)^{-1}$. Then, by the ellipticity condition (1.1),

$$\begin{aligned}(9.3)\qquad &\int_{5Q} |\nabla F_Q(x)|^2 w(x)\,dx \\ &\le \int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 \eta(x)^2 w(x)\,dx \le \lambda^{-1}\int_{\mathbb{R}^n} \mathbf{A}(x)\nabla F_Q(x)\cdot\nabla F_Q(x)\eta(x)^2\,dx.\end{aligned}$$

By the conservation property, Lemma 3.1, $e^{-t^2\mathcal{L}_w}\mathbf{1} = \mathbf{1}$; hence, $\nabla_x e^{-t^2\mathcal{L}_w}\mathbf{1} = 0$ and so we can write

$$\begin{aligned}\nabla_x e^{-t^2\mathcal{L}_w}\varphi(x) &= \nabla_x \int W_{t^2}(x,y)\,y\,dy \\ &= \nabla_x \int W_{t^2}(x,y)\,(y-x)\,dy + \mathbf{1} = \nabla_x G_t(x) + \mathbf{1},\end{aligned}$$

where $G_t(x) = \big(e^{-t^2\mathcal{L}_w} - I\big)\varphi(x)$. Thus $\nabla F_Q = \mathbf{1} + \nabla G_{\varepsilon\ell(Q)}$. By the Gaussian decay (1.3) of the heat kernel,

$$(9.4)\qquad |G_t(x)| \le \frac{C_1}{t^n}\int \exp\left(-C_2\frac{|x-y|^2}{t^2}\right)|y-x|\,dy \le Ct.$$

Integrating by parts, we have that

$$\begin{aligned}&\int_{\mathbb{R}^n} \mathbf{A}(x)\nabla F_Q(x)\cdot\nabla G_{\varepsilon\ell(Q)}(x)\eta(x)^2\,dx \\ &\quad= \int_{\mathbb{R}^n} \mathbf{A}(x)\nabla F_Q(x)\cdot\nabla(G_{\varepsilon\ell(Q)}\eta^2)(x)\,dx \\ &\qquad\qquad - 2\int_{\mathbb{R}^n} \mathbf{A}(x)\nabla F_Q(x)\cdot G_{\varepsilon\ell(Q)}(x)\nabla\eta(x)\eta(x)\,dx \\ &\quad= \int_{\mathbb{R}^n} \mathcal{L}_w F_Q(x) G_{\varepsilon\ell(Q)}(x)\eta(x)^2\,dx \\ &\qquad\qquad - 2\int_{\mathbb{R}^n} \mathbf{A}(x)\nabla F_Q(x)\cdot G_{\varepsilon\ell(Q)}(x)\nabla\eta(x)\eta(x)\,dx.\end{aligned}$$

If we combine this with (9.3) we get that

$$\begin{aligned}&\int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 w(x)\eta(x)^2\,dx \\ &\quad\le \lambda^{-1}\int_{\mathbb{R}^n} \mathbf{A}(x)\nabla F_Q(x)\cdot\big(\mathbf{1} + \nabla G_{\varepsilon\ell(Q)}(x)\big)\,\eta(x)^2 dx \\ &\quad\le \lambda^{-1}\int_{\mathbb{R}^n} |\mathbf{A}(x)\nabla F_Q(x)\cdot\mathbf{1}|\eta(x)^2 dx \\ &\qquad + \lambda^{-1}\int_{\mathbb{R}^n} |\mathcal{L}_w F_Q(x)||G_{\varepsilon\ell(Q)}(x)|\eta(x)^2 w(x)\,dx \\ &\qquad + 2\lambda^{-1}\int_{\mathbb{R}^n} \eta(x)|\mathbf{A}(x)\nabla F_Q(x)\cdot\nabla\eta(x)||G_{\varepsilon\ell(Q)}(x)|\,dx.\end{aligned}$$

By ellipticity, the inequality $|2ab| \leq \delta a^2 + \delta^{-1} b^2$, $\delta > 0$, and (9.4) we get

$$\begin{aligned}
&\int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 w(x)\eta(x)^2\, dx \\
&\quad \leq \frac{n\delta\Lambda}{2\lambda}\int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 w(x)\eta(x)^2 dx + \frac{n\Lambda}{2\lambda\delta}\int_{\mathbb{R}^n} \eta^2 w(x)\, dx \\
&\qquad + C\frac{\varepsilon^2 \ell(Q)^2}{\lambda}\int_{10Q} |\mathcal{L}_w F_Q(x)|^2 w(x)\, dx + \frac{C}{\lambda}\int_{\mathbb{R}^n} \eta(x)^2 w(x)\, dx \\
&\qquad + \frac{\delta\Lambda}{\lambda}\int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 \eta(x)^2 w(x)\, dx + \frac{\Lambda}{\lambda\delta}\varepsilon^2 \ell(Q)^2 \int_{10Q} |\nabla\eta(x)|^2 w(x)\, dx.
\end{aligned}$$

If we apply (ii), use the fact that $\varepsilon \leq 1$, $|\nabla\eta| \approx \ell(Q)^{-1}$ and w is doubling, and if we take $\delta > 0$ sufficiently small, then

$$\int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 w(x)\eta(x)^2\, dx \leq \frac{1}{2}\int_{\mathbb{R}^n} |\nabla F_Q(x)|^2 w(x)\eta(x)^2\, dx + C\frac{\Lambda}{\lambda} w(Q).$$

Rearranging terms and using the fact that $\eta \equiv 1$ on $5Q$, we get (i).

To prove that (III) holds we make two reductions. First, recall that the averaging operator A_t is defined by

$$A_t f(x) = \fint_{Q_t(x)} f(y)\, dy,$$

where $Q_t(x)$ is the unique dyadic cube containing x such that $t \leq \ell(Q_t(x)) < 2t$. As before, let $\tilde{\chi}(x) \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ be such that $\operatorname{supp}(\tilde{\chi}) \subset 4Q$, $\tilde{\chi}|_{3Q} \equiv 1$ and $|\nabla\tilde{\chi}| \leq C\ell(Q)^{-1}$. Then $P_t f(x) = P_t\tilde{\chi} f(x)$, and $A_t f(x) = A_t\tilde{\chi} f(x)$ for all $x \in Q$ and t, $0 \leq t \leq \ell(Q)$. Since P_t commutes with the gradient, by Lemma 5.2 we have that

$$\begin{aligned}
&\int_Q \int_0^{\ell(Q)} |\gamma_t(x)\nabla P_t F_{Q,\nu}(x)|^2 \frac{dt}{t} w(x)\, dx \\
&\quad \leq \int_Q \int_0^{\ell(Q)} |\gamma_t(x)(P_t - A_t)\tilde{\chi}\nabla F_{Q,\nu}(x)|^2 \frac{dt}{t} w(x)\, dx \\
&\qquad + \int_Q \int_0^{\ell(Q)} |\gamma_t(x) A_t \nabla F_{Q,\nu}(x)|^2 \frac{dt}{t} w(x)\, dx \\
&\quad \leq \sup_{t>0}\|\gamma_t\|_\infty^2 \int_Q \int_0^{\ell(Q)} |(P_t - A_t)\tilde{\chi}\nabla F_{Q,\nu}(x)|^2 \frac{dt}{t} w(x)\, dx \\
&\qquad + \int_Q \int_0^{\ell(Q)} |\gamma_t(x) A_t \nabla F_{Q,\nu}(x)|^2 \frac{dt}{t} w(x)\, dx \\
&\quad \leq C\sup_{t>0}\|\gamma_t\|_\infty^2 \int_{5Q} |\nabla F_{Q,\nu}(x)|^2 w(x)\, dx \\
&\qquad + \int_Q \int_0^{\ell(Q)} |\gamma_t(x) A_t \nabla F_{Q,\nu}(x)|^2 \frac{dt}{t} w(x)\, dx.
\end{aligned}$$

Therefore, by (I) we have

$$\frac{1}{w(Q)}\int_Q\int_0^{\ell(Q)}\left|\gamma_t(x)\nabla P_tF_{Q,\nu}(x)\right|^2\frac{dt}{t}w(x)\,dx$$
$$\leq C\sup_{t>0}\|\gamma_t\|_\infty^2+\frac{1}{w(Q)}\int_Q\int_0^{\ell(Q)}\left|\gamma_t(x)A_t\nabla F_{Q,\nu}(x)\right|^2\frac{dt}{t}w(x)\,dx,$$

and so to prove (III) it is enough to establish this estimate with P_t replaced by A_t.

For our second reduction, we need a lemma that was proved in [21] in the unweighted case; essentially the same argument works in the weighted case.

Lemma 9.1. *Suppose that there exist $0<\eta\leq 1$ such that for all cubes Q, there exists a subset $E_Q\subset Q$ with $w(E_Q)>\eta\, w(Q)$, and $Q\backslash E_Q=\bigcup Q_j$, where $\{Q_j\}$ is a set of non-overlapping dyadic sub-cubes of Q. Let $E_Q^*=R_Q\backslash\bigcup R_{Q_j}$, where R_Q denotes the Carleson rectangle $Q\times(0,\ell(Q))$ above Q. If for every cube Q,*

$$\mu\left(E_Q^*\right)\leq C_1\ w(Q),$$

then μ is a Carleson measure with respect to w, and

$$\mu(Q^*)\leq\frac{C_1}{\eta}w(Q).$$

Therefore, to complete our proof that (III) holds, we need to find our finite index set $\{\nu\}$ and construct sets E_Q as in Lemma 9.1 such that

$$\frac{1}{w(Q)}\iint_{E_Q^*}|\gamma_t(x)|^2\frac{dt}{t}\,w(x)\,dx \tag{9.5}$$
$$\leq C\sup_{t>0}\|\gamma_t\|_\infty^2+C\sum_\nu\frac{1}{w(Q)}\int_Q\int_0^{\ell(Q)}|\gamma_t(x)\nabla A_tF_{Q,\nu}(x)|^2\frac{dt}{t}\,w(x)\,dx.$$

We have now come to "the heart of the matter", as was said in [21]. For every $\nu\in\mathbb{C}^n$ with $|\nu|=1$, define the cone

$$\Gamma_\nu=\{z\in\mathbb{C}^n:|z-\nu(z\cdot\bar{\nu})|<\varepsilon|z\cdot\bar{\nu}|\}.$$

Clearly, for each $\varepsilon>0$ there exists a positive integer $N=N(\varepsilon,n)$ and unit vectors $\nu_1,\dots,\nu_N\in\mathbb{C}^n$ such that $\mathbb{C}^n\subset\bigcup_{j=1}^N\Gamma_{\nu_j}$. Below we will fix our ε and will then let $\{\nu\}=\{\nu_j\}_{j=1}^N$ be our index set.

We will first construct the sets E_Q. To do so, we will construct sets $E_{Q,\nu}$ that depend on ν; the desired set E_Q will be the union of these sets over our index set $\{\nu\}$. We need two more lemmas. The first is from [21]; since its proof depends only on the Gaussian bounds for the kernel of $e^{-t\mathcal{L}_w}$, it holds in the weighted case without change.

Lemma 9.2. *There exists a constant $C>0$ depending only on the constants C_1, C_2 in the Gaussian bounds, such that for any cube $Q\subset\mathbb{R}^n$,*

$$\left|\fint_Q(\nabla F_Q(x)-\mathbf{1})\,dx\right|\leq C\varepsilon.$$

The second lemma gives two properties of A_2 weights. The first is the well-known A_∞ condition, and the second is closely related. For a proof, see Theorems 2.9 and 2.11 in [19], or Theorem 9.3.3 in [20].

Lemma 9.3. *If $w \in A_2$, there exist constants α, $\delta > 0$ and constants β, $\epsilon > 0$ such that given any cube Q and measurable set $E \subset Q$,*

$$\frac{w(E)}{w(Q)} \leq \alpha \left(\frac{|E|}{|Q|}\right)^{\delta} \qquad \text{and} \qquad \frac{|E|}{|Q|} \leq \beta \left(\frac{w(E)}{w(Q)}\right)^{\epsilon}.$$

To prove (III) we first construct the sets $\{E_{Q,\nu}\}$ via the stopping time argument used in [21]. We include the details to show how the proof adapts to the weighted case. Let $\mathcal{S}_1$ be the collection of all maximal dyadic cubes $Q' \subset Q$ such that

$$\operatorname{Re} \fint_{Q'} \nu \cdot \nabla F_{Q,\nu}(y)\, dy \leq \frac{3}{4}, \tag{9.6}$$

and let $\mathcal{B}^1 = \mathcal{B}^1_{Q,\nu} = \bigcup_{\mathcal{S}_1} Q'$. Similarly, let $\mathcal{S}_2$ be the collection of maximal subcubes of Q such that

$$\fint_{Q'} |\nabla F_{Q,\nu}(y)|\, dy > \frac{1}{8\varepsilon}, \tag{9.7}$$

and let $\mathcal{B}^2 = \mathcal{B}^2_{Q,\nu} = \bigcup_{\mathcal{S}_2} Q'$. Set $\mathcal{B} = \mathcal{B}_{Q,\nu} = \mathcal{B}^1 \cup \mathcal{B}^2$ and $E_{Q,\nu} = Q \backslash \mathcal{B}$. By Lemma 9.2,

$$\left| \fint_Q (\nu \cdot \nabla F_{Q,\nu}(y)\, dy - 1) \right| = \left| \nu \left(\fint_Q (\nabla F_Q(y) - \mathbf{1})\, dy \right) \cdot \bar{\nu} \right| \leq C\varepsilon,$$

and thus

$$\begin{aligned} (1 - C\varepsilon)\, |Q| &\leq \operatorname{Re} \int_Q \nu \cdot \nabla F_{Q,\nu}(y)\, dy \\ &= \operatorname{Re} \left(\int_{E_{Q,\nu}} \nu \cdot \nabla F_{Q,\nu}(y)\, dy + \int_{\mathcal{B}^1} \nu \cdot \nabla F_{Q,\nu}(y)\, dy + \int_{\mathcal{B} \backslash \mathcal{B}^1} \nu \cdot \nabla F_{Q,\nu}(y)\, dy \right). \end{aligned} \tag{9.8}$$

By (9.6) and the definition of $\mathcal{B}^1$,

$$\operatorname{Re} \int_{\mathcal{B}^1} \nu \cdot \nabla F_{Q,\nu}(y)\, dy = \operatorname{Re} \sum_{Q' \in \mathcal{S}_1} \int_{Q'} \nu \cdot \nabla F_{Q,\nu}(y)\, dy \leq \frac{3}{4} \sum_{Q' \in \mathcal{S}_1} |Q'| \leq \frac{3}{4} |Q|. \tag{9.9}$$

By Hölder's inequality, (I) and the definition of A_2 weights,

$$\begin{aligned} &\operatorname{Re} \int_{\mathcal{B} \backslash \mathcal{B}^1} \nu \cdot \nabla F_{Q,\nu}(y)\, dy \\ &\leq \left(\int_{\mathcal{B}^2} |\nabla F_{Q,\nu}(y)|^2\, w(y)\, dy \right)^{1/2} w^{-1}\left(\mathcal{B}^2\right)^{1/2} \leq C w\left(Q\right)^{1/2} w^{-1}\left(\mathcal{B}^2\right)^{1/2} \\ &\leq C \left(w^{-1}\left(Q\right) w\left(Q\right)\right)^{1/2} \left(\frac{w^{-1}\left(\mathcal{B}^2\right)}{w^{-1}\left(Q\right)} \right)^{1/2} \leq C\, [w]_{A_2}\, |Q| \left(\frac{w^{-1}\left(\mathcal{B}^2\right)}{w^{-1}\left(Q\right)} \right)^{1/2}. \end{aligned} \tag{9.10}$$

Similarly, by the definition of $\mathcal{B}^2$ and (I),

$$\begin{aligned} \left|\mathcal{B}^2\right| &\leq \sum_{\mathcal{S}_2} |Q'| < 8\varepsilon \sum_{\mathcal{S}_2} \int_{Q'} |\nabla F_{Q,\nu}(y)|\, dy = 8\varepsilon \int_{\mathcal{B}^2} |\nabla F_{Q,\nu}(y)|\, dy \\ &\leq 8\varepsilon \left(\int_{\mathcal{B}^2} |\nabla F_{Q,\nu}(y)|^2\, w(y)\, dy \right)^{1/2} w^{-1}\left(\mathcal{B}^2\right)^{1/2} \\ &\leq 8C\varepsilon w\left(Q\right)^{1/2} w^{-1}\left(Q\right)^{1/2} \leq 8C\varepsilon\, [w]_{A_2}\, |Q| = C\varepsilon\, |Q|. \end{aligned} \tag{9.11}$$

By Lemma 9.3 we may assume that ε is so small that (9.11) implies

$$\frac{w^{-1}\left(\mathcal{B}^2\right)}{w^{-1}\left(Q\right)} \leq \left(\frac{1/8}{1+C\left[w\right]_{A_2}}\right)^2.$$

Combining this estimate with (9.10) we obtain

$$\operatorname{Re}\int_{\mathcal{B}\setminus\mathcal{B}^1} \nu\cdot\nabla F_{Q,\nu}(y)\,dy \leq \frac{1}{8}\left|Q\right|,$$

and putting this together with inequalities (9.8) and (9.9), for ε small enough we get

$$\frac{1}{16}\left|Q\right| \leq \operatorname{Re}\left(\int_{E_{Q,\nu}} \nu\cdot\nabla F_{Q,\nu}(y)\,dy\right).$$

But then, since $w\in A_2$,

$$\begin{aligned}\frac{1}{16}\left|Q\right| &\leq \operatorname{Re}\left(\int_{E_{Q,\nu}} \nu\cdot\nabla F_{Q,\nu}(y)\,dy\right)\\ &\leq C\left(w\left(Q\right)w^{-1}\left(E_{Q,\nu}\right)\right)^{1/2} \leq C\left|Q\right|\left(\frac{w^{-1}\left(E_{Q,\nu}\right)}{w^{-1}\left(Q\right)}\right)^{1/2}.\end{aligned}$$

Since w^{-1} is also in A_2, by Lemma 9.3 (applied twice) there exists $\eta>0$ such that

$$w\left(E_{Q,\nu}\right) \geq \eta\, w\left(Q\right). \tag{9.12}$$

Now write $\mathcal{B}_{Q,\nu} = Q\setminus E_{Q,\nu} = \bigcup Q_{\nu,j}$, where the $Q_{\nu,j}$ are disjoint maximal dyadic cubes. Let $E^*_{Q,\nu} = R_Q\setminus\bigcup R_{Q_{\nu,j}}$ be the sawtooth region above $E_{Q,\nu}$. If $(x,t)\in E^*_{Q,\nu}$, and $Q_t\left(x\right)$ is the biggest dyadic sub-cube of Q containing x with $t\leq \ell\left(Q_t\left(x\right)\right)<2t$, then by the maximality of the cubes in $\mathcal{S}_1$ and $\mathcal{S}_2$ we have that $Q_t\left(x\right)\notin\mathcal{B}$; hence

$$\frac{3}{4} < \operatorname{Re}\fint_{Q_t(x)} \nu\cdot\nabla F_{Q,\nu}(y)\,dy \leq \fint_{Q_t(x)} \left|\nabla F_{Q,\nu}(y)\right|\,dy \leq \frac{1}{8\varepsilon}.$$

By the definition of A_t, this implies that

$$\frac{3}{4} < \operatorname{Re}\nu\cdot A_t\nabla F_{Q,\nu}\left(x\right) \leq \left|A_t\nabla F_{Q,\nu}\left(x\right)\right| \leq \frac{1}{8\varepsilon}.$$

By the definition of Γ_ν, if $z\in\Gamma_\nu$, then $|z|<(1+\varepsilon)|z\cdot\bar\nu|$. Hence, for $\varepsilon>0$ sufficiently small, we have that

$$\begin{aligned}|z\cdot A_tF_{Q,\nu}(x)| &\geq |(z\cdot\bar\nu)(\nu\cdot A_tF_{Q,\nu}(x))| - |(z-\nu(z\cdot\bar\nu)\cdot A_tF_{Q,\nu}(x)|\\ &\geq \frac{3}{4}|z\cdot\bar\nu| - \varepsilon|z\cdot\bar\nu||A_tF_{Q,\nu}(x)| \geq \frac{5}{8}|z\cdot\bar\nu| \geq \frac{1}{2}|z|.\end{aligned}$$

Now fix ε small; form our index set $\{\nu\}=\{\nu_j\}_{j=1}^N$ as described above, and let $E_Q=\bigcup_j E_{Q,\nu_j}$. Therefore, if we let $z=\gamma_t(x)$, then $\gamma_t\left(x\right)\in\Gamma_{\nu_j}$ for some j, $1\leq j\leq N$, and we have

$$\frac{1}{4}\left|\gamma_t(x)\right|^2 \leq \sum_{j=1}^N \left|\gamma_t(x)\cdot A_t\nabla F_{Q,\nu_j}\left(x\right)\right|^2$$

for all $(x,t) \in E_Q^*$. It follows that

$$\begin{aligned}
&\frac{1}{w(Q)} \iint_{E_Q^*} |\gamma_t(x)|^2 \, \frac{dt}{t} \, w(x) \, dx \\
&\quad \le \frac{4}{w(Q)} \sum_{j=1}^{N} \iint_{E_Q^*} |\gamma_t(x) \cdot A_t \nabla F_{Q,\nu}(x)|^2 \, \frac{dt}{t} \, w(x) \, dx \\
&\quad \le \frac{4}{w(Q)} \sum_{j=1}^{N} \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot A_t \nabla F_{Q,\nu}(x)|^2 \, \frac{dt}{t} \, w(x) \, dx,
\end{aligned}$$

where, by (9.12), $w(E_Q) \ge \eta \, w(Q)$. This proves (9.5); thus we have shown that (III) in Lemma 8.2 holds, and so we have completed the proof of Theorem 1.1.

References

1. J. Alvarez and C. Pérez, *Estimates with A_∞ weights for various singular integral operators*, Boll. Un. Mat. Ital. A (7) **8** (1994), no. 1, 123–133. MR1273194 (95f:42027)
2. P. Auscher, A. Axelsson, and A. McIntosh, *On quadratic estimates related to the Kato conjecture and boundary value problems*, Preprint (2008).
3. P. Auscher, T. Coulhon, and P. Tchamitchian, *Absence de principe du maximum pour certaines équations paraboliques complexes*, Colloq. Math. **71** (1996), no. 1, 87–95. MR1397370 (97e:35019)
4. P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, and P. Tchamitchian, *The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$*, Ann. of Math. (2) **156** (2002), no. 2, 633–654. MR1933726 (2004c:47096c)
5. P. Auscher, S. Hofmann, J. Lewis, and P. Tchamitchian, *Extrapolation of Carleson measures and the analyticity of Kato's square-root operators*, Acta Math. **187** (2001), no. 2, 161–190. MR1879847 (2004c:47096a)
6. P. Auscher and P. Tchamitchian, *Square root problem for divergence operators and related topics*, Astérisque (1998), no. 249, viii+172. MR1651262 (2000c:47092)
7. J. Bergh and J. Löfström, *Interpolation spaces. An introduction*, Grundlehren der Mathematischen Wissenschaften, No. 223, Springer-Verlag, Berlin, 1976. MR0482275 (58:2349)
8. Y. Chen, *Regularity of the solution to the Dirichlet problem in Morrey spaces*, J. Partial Differential Equations **15** (2002), no. 2, 37–46. MR1909284 (2003e:35059)
9. ———, *Morrey regularity of solutions to degenerate elliptic equations in $\mathbf{R}^n$*, J. Partial Differential Equations **16** (2003), no. 2, 127–134. MR1979262 (2004c:35148)
10. ———, *Regularity of solutions to the Dirichlet problem for degenerate elliptic equation*, Chinese Ann. Math. Ser. B **24** (2003), no. 4, 529–540. MR2024992 (2004k:35150)
11. F. Chiarenza and M. Franciosi, *Quasiconformal mappings and degenerate elliptic and parabolic equations*, Matematiche (Catania) **42** (1987), no. 1-2, 163–170 (1989). MR1030914 (91b:35019)
12. M. Christ and J. Journé, *Polynomial growth estimates for multilinear singular integral operators*, Acta Math. **159** (1987), no. 1-2, 51–80. MR906525 (89a:42024)
13. D. Cruz-Uribe and C. Pérez, *On the two-weight problem for singular integral operators*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **1** (2002), no. 4, 821–849. MR1991004 (2004d:42021)
14. D. Cruz-Uribe and C. Rios, *Gaussian bounds for degenerate parabolic equations*, J. Funct. Anal. **255** (2008), no. 2, 283–312. MR2419963
15. J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics, vol. 29, American Mathematical Society, Providence, RI, 2001. MR1800316 (2001k:42001)
16. J. Duoandikoetxea and J. Rubio de Francia, *Maximal and singular integral operators via Fourier transform estimates*, Invent. Math. **84** (1986), no. 3, 541–561. MR837527 (87f:42046)
17. E. Fabes, C. Kenig, and R. Serapioni, *The local regularity of solutions of degenerate elliptic equations*, Comm. Partial Differential Equations **7** (1982), no. 1, 77–116. MR643158 (84i:35070)

18. B. Franchi, *Weighted Sobolev-Poincaré inequalities and pointwise estimates for a class of degenerate elliptic equations*, Trans. Amer. Math. Soc. **327** (1991), no. 1, 125–158. MR1040042 (91m:35095)
19. J. García-Cuerva and J. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies, vol. 116, North-Holland Publishing Co., Amsterdam, 1985. MR807149 (87d:42023)
20. L. Grafakos, *Classical and modern Fourier analysis*, Pearson/Prentice Hall, Upper Saddle River, NJ, 2004. MR2449250
21. S. Hofmann, *A short course on the Kato problem*, Second Summer School in Analysis and Mathematical Physics (Cuernavaca, 2000), Contemp. Math., vol. 289, Amer. Math. Soc., Providence, RI, 2001, pp. 61–77. MR1864539 (2002k:47096)
22. S. Hofmann, M. Lacey, and A. McIntosh, *The solution of the Kato problem for divergence form elliptic operators with Gaussian heat kernel bounds*, Ann. of Math. (2) **156** (2002), no. 2, 623–631. MR1933725 (2004c:47096b)
23. J. Journé, *Calderón-Zygmund operators, pseudodifferential operators and the Cauchy integral of Calderón*, Lecture Notes in Mathematics, vol. 994, Springer-Verlag, Berlin, 1983. MR706075 (85i:42021)
24. ———, *Remarks on Kato's square-root problem*, Publ. Mat. **35** (1991), no. 1, 299–321, Conference on Mathematical Analysis (El Escorial, 1989). MR1103623 (92d:47050)
25. T. Kato, *Fractional powers of dissipative operators*, J. Math. Soc. Japan **13** (1961), 246–274. MR0138005 (25:1453)
26. ———, *Perturbation theory for linear operators*, Die Grundlehren der mathematischen Wissenschaften, Band 132, Springer-Verlag New York, Inc., New York, 1966. MR0203473 (34:3324)
27. C. Kenig, *Featured review: The solution of the Kato square root problem for second order elliptic operators on* $\mathbb{R}^n$, Ann. of Math. (2) **156** (2002), no. 2, 633–654. MR1933726 (2004c:47096c)
28. K. Kurata and S. Sugano, *Fundamental solution, eigenvalue asymptotics and eigenfunctions of degenerate elliptic operators with positive potentials*, Studia Math. **138** (2000), no. 2, 101–119. MR1749075 (2001c:35172)
29. J. Lewis and K. Nyström, *Boundary behaviour for p harmonic functions in Lipschitz and starlike Lipschitz ring domains*, Ann. Sci. École Norm. Sup. (4) **40** (2007), no. 5, 765–813. MR2382861
30. S. Semmes, *Square function estimates and the* $T(b)$ *theorem*, Proc. Amer. Math. Soc. **110** (1990), no. 3, 721–726. MR1028049 (91h:42018)
31. E. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, NJ, 1970. MR0290095 (44:7280)
32. E. Stein and G. Weiss, *Interpolation of operators with change of measures*, Trans. Amer. Math. Soc. **87** (1958), 159–172. MR0092943 (19,1184d)
33. M. Wilson, *The convergence of the Calderón reproducing formula*, preprint (2007).

Department of Mathematics, Trinity College, Hartford, Connecticut 06106
E-mail address: david.cruzuribe@trincoll.edu

Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada
E-mail address: crios@math.ucalgary.ca