Acta Math., 187 (2001), 161–190

Extrapolation of Carleson measures and the analyticity of Kato's square-root operators

by

PASCAL AUSCHER

Université de Picardie–Jules Verne
Amiens, France

STEVE HOFMANN

University of Missouri
Columbia, MO, U.S.A.

JOHN L. LEWIS

University of Kentucky
Lexington, KY, U.S.A.

and

PHILIPPE TCHAMITCHIAN

Université d'Aix–Marseille III
Marseille, France

1. Introduction, history and statement of the main theorem

Let A be an $(n\times n)$-matrix of complex L^∞-coefficients, defined on $\mathbf{R}^n$, with $\|A\|_\infty\leqslant\Lambda$, and satisfying the ellipticity (or "accretivity") condition

$$\lambda|\xi^2|\leqslant \operatorname{Re}\langle A\xi,\xi\rangle\leqslant\Lambda|\xi|^2, \tag{1.1}$$

for $\xi\in\mathbf{C}^n$ and for some λ,Λ such that $0<\lambda\leqslant\Lambda<\infty$. Here $\langle\,\cdot\,,\cdot\,\rangle$ denotes the usual inner product in $\mathbf{C}^n$, so that

$$\langle A\xi,\xi\rangle\equiv\sum_{i,j}A_{ij}(x)\xi_j\cdot\bar{\xi}_i.$$

We define a divergence-form operator

$$Lu\equiv-\operatorname{div}(A(x)\nabla u), \tag{1.2}$$

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

The accretivity condition (1.1) enables one to define an accretive square root $\sqrt{L}\equiv L^{1/2}$ (see [14]), and a fundamental question is to determine when one can solve the "square-root problem", i.e. to establish the estimate

$$\|\sqrt{L}f\|_{L^2(\mathbf{R}^n)}\leqslant C\|\nabla f\|_{L^2(\mathbf{R}^n)}, \tag{1.3}$$

Hofmann and Lewis were each partially supported by NSF research grants. Lewis also was partially supported by the Mittag-Leffler Institute, who provided both financial support and a pleasant working atmosphere during the spring of 2000.

with C depending only on n, λ and Λ. The latter estimate is connected with the question of the analyticity of the mapping $A \to L^{1/2}$, which in turn has applications to the perturbation theory for certain classes of hyperbolic equations (see [15], [19]). We note that it is well known, and easy to see, that (1.3) holds when L is self-adjoint.

A long-standing open problem, essentially posed by Kato [14] (but refined by McIntosh [19], [21]—we shall explain this point more fully below), is the following:

QUESTION 1. *Let A_z, $z \in \mathbf{C}$, denote a family of accretive matrices as above, which in addition are holomorphic in z, and self-adjoint for real z. Let*

$$L_z \equiv -\operatorname{div} A_z(x)\nabla.$$

Is $L_z^{1/2}$ holomorphic in z, in a neighborhood of $z=0$?

In fact, Kato actually formulated this question for a more general class of abstract accretive operators. A counterexample to the abstract problem was found by McIntosh [21]. However, it has been pointed out in [19] that, in posing the problem, Kato had been motivated by the special case of elliptic differential operators, and by the applicability of a positive result, in that special case, to the perturbation theory for hyperbolic evolution equations. A positive answer to the question posed above can be restated as

CONJECTURE 1.4. *The estimate* (1.3) *holds in a complex neighborhood in L^∞ of any self-adjoint matrix A satisfying* (1.1); *i.e.* (1.3) *holds for the operator $\tilde{L}$ (as in* (1.2)) *associated to any complex-valued matrix $\tilde{A}$, whenever*

$$\|\tilde{A} - A\|_\infty \leqslant \varepsilon_0,$$

with ε_0 depending only on n, λ and Λ.

Indeed, given Conjecture 1.4, then by the operator-valued version of Cauchy's theorem, one obtains analyticity at $z=0$ of the mapping

$$z \to L_z^{1/2},$$

where $L_z \equiv -\operatorname{div}(A_z)\nabla$, $z \to A_z$ is analytic, and $A_0 \equiv A$ is self-adjoint. It was this analyticity result that Kato had sought, in particular for real, symmetric matrices, in connection with the theory of hyperbolic equations.

In [14], Kato also framed a more general conjecture for square roots of abstract accretive operators belonging to some broad class (see [22] for the details). Again the abstract question was shown to have a negative answer: a counterexample was obtained by McIntosh [20], who then reformulated the conjecture for the special case of elliptic differential operators:

CONJECTURE 1.5. *The estimate* (1.3) *holds for any operator* L *defined as in* (1.2), *associated to an* L^∞-$(n\times n)$-*matrix* A *with complex entries, for which* (1.1) *holds.*

To establish the validity of Conjecture 1.5 has become known as the *Kato problem*, or *square-root problem.* Until recently, both Conjecture 1.4 and Conjecture 1.5 had been proved completely only when $n=1$.

In the 1-dimensional case, the square-root problem is essentially equivalent to the problem of establishing the L^2-boundedness of the Cauchy integral operator along a Lipschitz curve. Thus Conjecture 1.5, and hence also Conjecture 1.4, were proved in one dimension in the celebrated paper of Coifman, McIntosh and Meyer [5]. The precise nature of the relationship between the Cauchy integral operator along a Lipschitz curve, and the 1-dimensional Kato problem, was obtained in [16].

In higher dimensions, both Conjecture 1.4 and Conjecture 1.5 had been proved only in the case that A is close, in some sense, to a constant matrix (or in the case that one imposes some additional structure on the matrix—see [2] for some examples).

The first result involving perturbations of constant matrices was due independently to Coifman, Deng and Meyer [4], and Fabes, Jerison and Kenig [9], who established the square-root estimate (1.3) whenever $\|A-I\|_\infty\leqslant\varepsilon(n)$. Clearly, their methods allowed one also to replace the identity matrix I by any constant accretive matrix, and this was certainly understood at that time (see [10]). Sharper bounds for the constant $\varepsilon(n)$ on the order of $n^{-1/2}$ were obtained by Journé [13]. Another result in the same spirit was due to Fabes, Jerison and Kenig [unpublished], who proved that an appropriate analogue of (1.3) holds when A is continuous (and hence, at least locally, close to a constant matrix). Extensions of these "small constant" results, with L^∞ replaced by BMO, and C replaced by VMO, were obtained by Escauriaza (VMO, unpublished), and by Auscher and Tchamitchian [2] (BMO with small norm; ABMO, a space somewhat beyond VMO; and, more generally, small perturbations of ABMO in BMO). In the latter results, one still supposes that $A\in L^\infty$; the point is that the smallness of the perturbation is measured in a more general sense.

In the present paper, we present the solution to Conjecture 1.4, in all dimensions, at least in the case that A is real, symmetric. Our main result is

THEOREM 1.6. *Let* $n\geqslant 1$. *Suppose that* A *is a real, symmetric* $(n\times n)$-*matrix of* L^∞-*coefficients satisfying* (1.1). *Then there exists* $\varepsilon_0\equiv\varepsilon_0(n,\lambda,\Lambda)$ *such that for any complex-valued* $(n\times n)$-*matrix* $\tilde{A}$, *with* $\|A-\tilde{A}\|_\infty\leqslant\varepsilon_0$, *the operator*

$$\tilde{L}\equiv-\operatorname{div}(\tilde{A}(x)\nabla)$$

satisfies (1.3), *with a constant* C *which depends only on* n,λ,Λ. *Moreover,*

$$\big\|\sqrt{\tilde{L}}\,f-\sqrt{L}\,f\big\|_{L^2(\mathbf{R}^n)}\leqslant C(n,\lambda,\Lambda)\|A-\tilde{A}\|_\infty\|\nabla f\|_{L^2(\mathbf{R}^n)}. \tag{1.7}$$

It is worthwhile to make several comments at this point. The first is that it is enough to establish (1.3) for $\sqrt{L}$, for then (1.7) follows immediately by our previous remarks concerning Conjecture 1.4 and analyticity. Second, we observe that, more generally, our proof actually yields that Conjecture 1.4 holds if A is merely self-adjoint (not necessarily real, symmetric), if we assume also that the heat kernel $W_{t^2}(x,y)$, which is the kernel of the operator e^{-t^2L}, satisfies the "Gaussian" property

$$|W_{t^2}(x,y)| \leqslant C(n,\lambda,\Lambda)\, t^{-n} \exp\left\{\frac{-|x-y|^2}{Ct^2}\right\}, \tag{1.8i}$$

$$\begin{aligned}|W_{t^2}(x+h,y)-W_{t^2}(x,y)|+|W_{t^2}(x,y+h)-W_{t^2}(x,y)|& \\ \leqslant C(n,\lambda,\Lambda)\frac{|h|^\alpha}{t^{n+\alpha}}\exp\left\{\frac{-|x-y|^2}{Ct^2}\right\}&,\end{aligned} \tag{1.8ii}$$

where the latter inequality holds for some positive exponent α depending only on n, λ and Λ, whenever either $|h|\leqslant t$ or $|h|\leqslant\frac{1}{2}|x-y|$. Of course, (1.8) always holds for A real, symmetric, by the classical parabolic regularity theory of Nash–Moser–Aronson. To simplify matters as much as possible, we shall assume in the sequel that A is real, symmetric. We leave it to the interested reader to check that the same arguments yield a proof in the slightly more general case that L is merely self-adjoint and Gaussian. We shall not insist on this point here, as we plan, in a future paper, to prove a more general result. Indeed, three of us (Auscher, Hofmann and Tchamitchian), along with M. Lacey and A. McIntosh, will present the proof of Conjecture 1.5, in general. In dimension 2, the solution to Conjecture 1.5 has recently been obtained by one of the present authors (Hofmann), jointly with McIntosh [12]. It was then observed by M. Lacey [17] that the use of an appropriate sectorial decomposition of $\mathbf{C}^n$ allows one to extend the argument in [12] to higher dimensions, assuming that the above-mentioned "Gaussian" property holds. The removal of the Gaussian hypothesis was then permitted by means of an argument due to Auscher and Tchamitchian. A summary of these combined efforts, giving the complete solution to the Kato problem, will appear in a forthcoming paper.

The paper is organized as follows. In the next section, we discuss the strategy of our proof and make some preliminary reductions. In particular, we shall state an "extrapolation lemma" for Carleson measures, which lies at the heart of our approach here. In §3, we prove the extrapolation lemma. In §4, we state another key lemma, and use it, along with the extrapolation lemma, to prove Theorem 1.6. The proof of this key lemma is given in §5. §6 is an appendix, in which we give the proof of one technical lemma.

Acknowledgements. The work described in this paper was initiated during the visit in 1996 of the second author to Macquarie University, where the question was posed to him by A. McIntosh. We thank him for his encouragement and support, and for numerous illuminating conversations during the subsequent three plus years of effort to solve the problem. We are also very grateful to the referee, for many insightful suggestions which have improved the exposition.

2. The strategy of the proof, notation and preliminary arguments

In the sequel, we shall use the convention that the generic constant C may depend upon n, λ and Λ, but that when a constant depends upon other parameters, we shall note that dependence explicitly, while leaving any dependence upon n, λ and Λ implicit. We shall also suppose in the sequel that A is real, symmetric, and that $\|\tilde{A}-A\|_\infty \leqslant \varepsilon_0$. Moreover, by [2, Chapter 0.5, Proposition 7], we may assume, and do, that $\tilde{A}, A \in C^\infty$. Our estimates, of course, will depend only on n, λ and Λ.

Let us now state some notation that we shall use in the sequel. Given a cube $Q \in \mathbf{R}^n$, let $l(Q)$ denote the side length of Q, and let R_Q^b and T_Q^b denote respectively the Carleson box above Q of height $bl(Q)$, and the Carleson tent above Q with slope b. That is, we set $R_Q^b \equiv Q \times (0, bl(Q))$ and $T_Q^b \equiv \{(x,t): x \in Q, 0<t<b \operatorname{dist}(x, Q^c)\}$. In the case that $b=1$ we shall write merely R_Q and T_Q. Given a positive Borel measure ν on the upper half-space, we denote its "Carleson norm" by

$$\|\nu\|_C \equiv \sup |Q|^{-1} \nu(R_Q),$$

where the supremum runs over all cubes Q with sides parallel to the coordinate axes. Given a Lipschitz function ψ defined on $\mathbf{R}^n$, we denote by Ω_ψ the domain above the graph of ψ; i.e.,

$$\Omega_\psi \equiv \{(x,t) \in \mathbf{R}^{n+1} : t > \psi(x)\}.$$

It is known [2] (although for the reader's convenience we shall sketch a proof below) that one may reduce the proof of (1.3) for $\sqrt{\tilde{L}}$ to proving a certain Carleson measure estimate, which we shall now describe. We define a measure on the upper half-space by

$$d\tilde{\mu}(x,t) \equiv |\tilde{\gamma}_t(x)|^2 \, dx \, \frac{dt}{t}, \tag{2.1}$$

where

$$\tilde{\gamma}_t(x) \equiv e^{-t^2 \tilde{L}} t \tilde{L} \varphi(x). \tag{2.2}$$

Here,

$$\varphi(x) \equiv x.$$

(In the sequel, φ will always be this $\mathbf{R}^n$-valued function.) Our goal is to establish the Carleson measure estimate

$$\|\tilde{\mu}\|_C \equiv \sup_Q \frac{1}{|Q|}\int_Q\int_0^{l(Q)} |e^{-t^2\tilde{L}}t\tilde{L}\varphi(x)|^2\,\frac{dt}{t}\,dx \leqslant C(1+\varepsilon_0^2\|\tilde{\mu}\|_C). \tag{2.3}$$

By smoothly truncating in the time variable, we may suppose *a priori* that $\|\tilde{\mu}\|_C$ is finite. Thus, assuming that (2.3) holds, and taking ε_0 small enough, depending only on ellipticity and dimension, we may hide the small term on the left side of the inequality to obtain that $\|\tilde{\mu}\|_C \leqslant C$. The bounds that we obtain are of course independent of the truncation, which may then be removed by a limiting argument. We shall not tire the reader with such routine details, and shall therefore suppress the truncation in the sequel.

Let us now sketch a proof that this last estimate implies that (1.3) holds for $\sqrt{\tilde{L}}$. A complete proof, by another method, may be found in [2]. We begin by noting that, since (1.8) holds for L (with A real, symmetric), a perturbation result of [1] (which is given also as [2, Chapter 1.2, Theorem 6 (ii)]) implies the following: for $\|\tilde{A}-A\|_\infty \leqslant \varepsilon_0$, $\varepsilon_0(n,\lambda,\Lambda)$ small enough, the heat kernel $\widetilde{W}_{t^2}(x,y)$, which is the kernel of the operator $e^{-t^2\tilde{L}}$, satisfies

$$|\widetilde{W}_{t^2}(x,y)| \leqslant Ct^{-n}\exp\biggl\{\frac{-|x-y|^2}{Ct^2}\biggr\}, \tag{2.4i}$$

$$|\widetilde{W}_{t^2}(x+h,y)-\widetilde{W}_{t^2}(x,y)|+|\widetilde{W}_{t^2}(x,y+h)-\widetilde{W}_{t^2}(x,y)| \leqslant C\frac{|h|^\alpha}{t^{n+\alpha}}\exp\biggl\{\frac{-|x-y|^2}{Ct^2}\biggr\}, \tag{2.4ii}$$

where the latter inequality holds for some positive exponent α depending only on n, λ and Λ, whenever either $|h|\leqslant t$ or $|h|\leqslant \frac{1}{2}|x-y|$. Letting $\tilde{L}^*$ denote the adjoint of $\tilde{L}$, we see that the same bounds hold also for the kernel of $t^2\tilde{L}^*e^{-t^2\tilde{L}^*}$. Moreover, $t^2\tilde{L}^*e^{-t^2\tilde{L}^*}1=0$. Thus, standard real-variable orthogonality techniques imply the square-function estimate

$$\int_0^\infty\int_{\mathbf{R}^n} |t^2\tilde{L}^*e^{-t^2\tilde{L}^*}g(x)|^2\,dx\,\frac{dt}{t} \leqslant C\|g\|^2_{L^2(\mathbf{R}^n)}.$$

Consequently, if we resolve the square root as

$$\tilde{L}^{1/2} = \int_0^\infty e^{-2t^2\tilde{L}}\tilde{L}^2t^2\,dt,$$

and then dualize and apply Schwarz's inequality, we see that to prove (1.3) for $\tilde{L}^{1/2}$, it is enough to establish the inequality

$$\int_0^\infty\int_{\mathbf{R}^n} |t\tilde{L}e^{-t^2\tilde{L}}f(x)|^2\,dx\,\frac{dt}{t} \leqslant C\|\nabla f\|^2_{L^2(\mathbf{R}^n)}.$$

We now claim that, in the spirit of the $T1$-theorem, this last square-function estimate follows from (2.3). Let us sketch a simple proof of the claim.

For a function $G(x,t)$, define the triple bar norm by

$$|||G|||^2=\int_0^\infty\int_{\mathbf{R}^n}|G(x,t)|^2\,dx\,\frac{dt}{t},$$

and set $R_tf=t\tilde{L}e^{-t^2\tilde{L}}f$. We want to prove that the Carleson measure estimate (2.3) implies $|||R_tf|||\leqslant C\|\nabla f\|_2$. Let P_t denote a nice convolution-type approximate identity. By a slight abuse of notation, let y denote the variable of integration in the definition of the integral operator R_t applied to f, i.e., $R_tf(x)=R_t(f(y))(x)$, and $R_t(y)=R_t\varphi$, since $\varphi(y)=y$. Since $R_t1=0$, we have, following [3], that

$$R_tf(x)=R_t(f(y)-f(x)-\langle y-x,\nabla P_tf(x)\rangle)+\langle R_t(y)(x),\nabla P_tf(x)\rangle\equiv \mathrm{I}+\mathrm{II}.$$

Now, the non-tangential maximum of ∇P_tf is L^2-bounded, so the triple bar norm of II satisfies the desired bound, given that we have an appropriate Carleson measure estimate for $R_t(y)$, namely (2.3). Moreover, it is essentially known that the triple bar norm of I is bounded. Indeed, just take the absolute value of the integrand, and use, in effect, the results of Dorronsoro [8], along with property (G), to estimate the tail of the kernel of the operator R_t. We leave the routine details to the reader. This completes our sketch of the proof of the fact that (2.3) implies (1.3) for $\sqrt{\tilde{L}}$.

Thus, our goal is to prove (2.3). Our method of proof is one which has been used in [18] and [11], to establish parabolic measure estimates for certain classes of parabolic equations. This technique is an inductive procedure which, roughly speaking, utilizes a stopping time argument, reminiscent of Carleson's "corona" construction, to "extrapolate" the constant which bounds a certain Carleson measure estimate. In the present setting, this extrapolation method may be formalized as follows. Let μ, $\tilde{\mu}$ be two positive measures defined on the upper half-space $\mathbf{R}^{n+1}_+$, with

$$d\mu\equiv K_t(x)\,dx\,\frac{dt}{t},\quad d\tilde{\mu}\equiv \widetilde{K}_t(x)\,dx\,\frac{dt}{t},$$

where $0\leqslant K_t(x),\widetilde{K}_t(x)\leqslant\beta_0$. In the next section we shall prove the following "extrapolation lemma for Carleson measures".

LEMMA 2.5. *Suppose that $\mu,\tilde{\mu}$ are given as above, and that μ is a Carleson measure with $\|\mu\|_C\leqslant C_0$. Suppose also that there are positive constants δ and C_1 such that*

$$\tilde{\mu}(R_Q\cap\Omega_\psi)\leqslant C_1|Q|,$$

for every cube Q and every positive Lipschitz function ψ with $\|\nabla\psi\|_\infty\leqslant 1$ which satisfy

$$\sup|Q'|^{-1}\mu(T_{Q'}\cap\Omega_\psi)\leqslant\delta,$$

where the supremum runs over all dyadic subcubes $Q'\subseteq Q$. Then $\tilde{\mu}$ is a Carleson measure, with

$$\|\tilde{\mu}\|_C \leqslant C(n,\delta,C_0,\beta_0)(1+C_1).$$

For our purposes, we shall apply the extrapolation lemma with $\tilde{\mu}$ defined by (2.1), and with μ defined by

$$d\mu(x,t) \equiv |\gamma_t(x)|^2\,dx\,\frac{dt}{t}, \tag{2.6}$$

where

$$\gamma_t(x) \equiv \gamma_t^\varepsilon(x) \equiv e^{-\varepsilon^2 t^2 L}\varepsilon t L\varphi(x), \tag{2.7}$$

and ε is a small, fixed number, to be chosen later, and which will ultimately depend only on n, λ and Λ. Since A is real, symmetric, it is not hard to see that

$$\|\mu\|_C \leqslant C_0 \tag{2.8}$$

(independently of ε), where C_0 depends only on ellipticity and dimension. Indeed, this follows easily from the fact that the heat kernel $W_{t^2}(x,y)$, the kernel of e^{-t^2L}, satisfies (1.8), plus the fact that $\sqrt{L}$ satisfies (1.3). We omit the details, which are standard. Moreover, it follows readily from (1.8) and (2.4) that $|\gamma_t(x)|^2, |\tilde{\gamma}_t(x)|^2 \leqslant \beta_0$, with β_0 depending only on ellipticity and dimension, and again we omit the routine details. We are therefore left with two main tasks. One of these is to prove the extrapolation lemma; the other is to verify that μ and $\tilde{\mu}$ satisfy the remaining hypothesis of the extrapolation lemma, for some δ depending only on ellipticity and dimension, and with

$$C_1 = C(\varepsilon^{-1})(1+\varepsilon_0^2\|\tilde{\mu}\|_C),$$

where ε is the same as in (2.7). In carrying out the latter task, we shall exploit the circle of ideas surrounding the proof of a sort of "Tb"-theorem for square roots, given in [2]. Let P_t denote a nice approximate identity, given by convolution with a function $t^{-n}p(x/t)$, $p\in C_0^\infty$, with support in the unit ball, and $\int p=1$. Suppose that there are constants C' and C'' such that for each cube Q, there exists a mapping $F\equiv F_Q: 5Q\to\mathbf{C}^n$ (here $5Q$ denotes the concentric dilate of Q having side length $5l(Q)$), satisfying

$$\int_{5Q} |\nabla F_Q|^2 \leqslant C'|Q|, \tag{2.9i}$$

$$\int_{5Q} |LF_Q|^2 \leqslant C''\frac{|Q|}{(l(Q))^2}. \tag{2.9ii}$$

(Here ∇F_Q denotes the transpose of the Jacobian matrix.) We shall utilize the ideas of [2] in the form of the following lemma, whose proof may be deduced from the proofs of Theorems 3 and 4 of [2, §3.2], although for the sake of self-containment, we shall give the proof here momentarily.

LEMMA 2.10 [2]. *Suppose that* $\|\tilde{A}-A\|_{L^\infty}\leqslant\varepsilon_0$. *Fix* Q *and suppose that there exists* F_Q *satisfying* (2.9i) *and* (2.9ii) *with respect to* Q. *Then,*

$$\frac{1}{|Q|}\int_Q\int_0^{l(Q)}|\tilde{\gamma}_t(x)P_t(\nabla F_Q)(x)|^2\,\frac{dt}{t}\,dx\leqslant C(1+C'+C''+\varepsilon_0^2\|\tilde{\mu}\|_C).$$

Before proving the lemma, we note that in our case we shall define the mapping F_Q as follows. Given Q, with side length $l(Q)$, we define $F_Q:\mathbf{R}^n\to\mathbf{R}^n$ by

$$F_Q\equiv e^{-(\varepsilon/2)^2(l(Q))^2L}\varphi, \tag{2.11}$$

where ε is the same small number that first appeared in (2.7). We remind the reader, also, that $\varphi(x)\equiv x$, throughout this paper. We observe that this is the same F_Q that was introduced previously in the solution of the 2-dimensional Kato conjecture in [12]. It is a routine matter to prove that this particular choice of F_Q satisfies

$$\int_{5Q}|\nabla F_Q|^2\leqslant C|Q|, \tag{2.12i}$$

$$\int_{5Q}|LF_Q|^2\leqslant C\frac{|Q|}{\varepsilon^2(l(Q))^2}, \tag{2.12ii}$$

and we omit the details. These estimates are, of course, restatements of (2.9i) and (2.9ii).

For the reader's convenience, let us now sketch the proof of Lemma 2.10, following [2, §3.2]. As usual, let $H^1(5Q)$ denote the homogeneous Sobolev space of complex-valued functions having a gradient in $L^2(5Q)$, and let $H_0^1(5Q)$ denote the closure of $C_0^1(5Q)$ in $H^1(5Q)$. By ellipticity, the sesquilinear form

$$B_{\tilde{A}}(\phi,\psi)\equiv\int_{5Q}\tilde{A}\nabla\phi\cdot\nabla\bar{\psi}$$

is coercive and bounded on $H_0^1(5Q)$. Also, the mapping

$$\psi\to-\int_{5Q}(\tilde{A}-A)\nabla F_Q\cdot\nabla\bar{\psi}$$

defines a bounded anti-linear functional on $H_0^1(5Q)$. Thus, by the Lax–Milgram lemma, there exists a unique $H_Q\in H_0^1(5Q)$ such that

$$\int_{5Q}|\nabla H_Q|^2\leqslant C\varepsilon_0^2|Q|,$$

and $\tilde{L}H_Q=\operatorname{div}(\tilde{A}-A)\nabla F_Q$ in the weak sense. Setting $G_Q\equiv F_Q+H_Q$, we then have that

$$\tilde{L}G_Q=LF_Q$$

in the weak sense. Choosing ε_0 small enough, we therefore obtain from (2.9) that

$$\int_{5Q} |\nabla G_Q|^2 \leqslant 2C'|Q| \tag{2.13i}$$

and

$$\int_{5Q} |\tilde{L} G_Q|^2 \leqslant C'' \frac{|Q|}{(l(Q))^2}. \tag{2.13ii}$$

We now define an operator, mapping matrix-valued L^2-functions into $\mathbf{C}^n$-valued functions, by

$$\theta_t \mathbf{f}(x) \equiv -te^{-t^2\tilde{L}} \operatorname{div} \tilde{A}\mathbf{f},$$

so that, by the definition (2.2), we have

$$\widetilde{\gamma}_t(x) = \theta_t \mathbf{1}(x),$$

where $\mathbf{1}=\nabla\varphi$ denotes the identity $(n\times n)$-matrix. To prove the lemma, we therefore need to show that

$$\frac{1}{|Q|}\int_Q \int_0^{l(Q)} |\theta_t \mathbf{1}(x) P_t(\nabla F_Q)(x)|^2 \frac{dt}{t}\,dx \leqslant C(1+C'+C''+\varepsilon_0^2\|\tilde{\mu}\|_C).$$

We may replace F_Q by G_Q, as the resulting error is no larger than $C\varepsilon_0^2\|\tilde{\mu}\|_C$. We may also multiply ∇G_Q by a smooth, non-negative cut-off function χ_Q, supported in $4Q$ and identically 1 in $3Q$, since the convolution kernel of P_t has support in a ball of radius $t\leqslant l(Q)$. Furthermore, we may replace G_Q by $\widetilde{G}_Q \equiv \widetilde{\chi}_Q(G_Q - c_Q)$, where c_Q denotes the mean value of G_Q on $5Q$, and where $\widetilde{\chi}_Q$ is another smooth, non-negative cut-off function, supported in $5Q$ and identically 1 in $4Q$. We note that by Poincaré's inequality, $\widetilde{G}_Q$ satisfies (2.13i) with constant CC', and that $\nabla\widetilde{G}_Q = \nabla G_Q$ on $4Q$. Following a trick of Coifman and Meyer [6], we write

$$\theta_t = \theta_t - \theta_t \mathbf{1} P_t + \theta_t \mathbf{1} P_t \equiv S_t + \theta_t \mathbf{1} P_t.$$

Since $S_t\mathbf{1}=0$, it follows from a slight variation of standard orthogonality arguments that

$$\frac{1}{|Q|}\int_Q \int_0^{l(Q)} |S_t(\nabla\widetilde{G}_Q)(x)|^2 \frac{dt}{t}\,dx \leqslant CC',$$

where we have used that S_t is being applied to a gradient field. Moreover,

$$\frac{1}{|Q|}\int_Q \int_0^{l(Q)} |S_t((1-\chi_Q)\nabla\widetilde{G}_Q)(x)|^2 \frac{dt}{t}\,dx \leqslant CC',$$

since the kernel of S_t decays rapidly, at least in the sense of L^2-averages, by a standard argument using ellipticity, integration by parts, and the Gaussian bounds for the heat kernel of $\tilde{L}$. Also,

$$-\operatorname{div}\tilde{A}\chi_Q\nabla\widetilde{G}_Q=-\operatorname{div}\tilde{A}\chi_Q\nabla G_Q=\chi_Q\tilde{L}G_Q-\tilde{A}\nabla\chi_Q\cdot\nabla G_Q,$$

so that

$$\theta_t(\chi_Q\nabla\widetilde{G}_Q)=te^{-t^2\tilde{L}}\chi_Q\tilde{L}G_Q-te^{-t^2\tilde{L}}\tilde{A}\nabla\chi_Q\cdot\nabla G_Q.$$

Hence

$$\int_Q\int_0^{l(Q)}|\theta_t(\chi_Q\nabla\widetilde{G}_Q)(x)|^2\,\frac{dt}{t}\,dx\leqslant C(l(Q))^2\int_{5Q}(|\tilde{L}G_Q|^2+|\nabla\chi_Q|^2|\nabla G_Q|^2),$$

and the conclusion of Lemma 2.10 now follows from (2.13).

We finish this section by stating a lemma which we shall find useful in the sequel. It is a sort of "John–Nirenberg lemma for Carleson measures".

LEMMA 2.14. *Fix Q. Suppose that $0\leqslant H_t(x)\leqslant\beta_0$ in Q, and that*

$$|H_t(x)-H_t(x')|\leqslant\beta_0\frac{|x-x'|^\alpha}{t^\alpha},$$

for some $\alpha>0$, whenever $x,x'\in Q$. Suppose also that there is a number $\eta\in(0,1]$, and a number β, such that for every dyadic subcube $Q'\subseteq Q$ there is a subset $E'\subseteq Q'$, with

$$|E'|\geqslant\eta|Q'|$$

and

$$\int_{E'}\int_0^{l(Q')}H_t(x)\,\frac{dt}{t}\,dx\leqslant\beta|Q'|.$$

Then the following estimate holds in Q:

$$\frac{1}{|Q|}\int_Q\int_0^{l(Q)}H_t(x)\,\frac{dt}{t}\,dx\leqslant C(\alpha,\eta)(\beta_0+\beta).$$

We defer the proof of Lemma 2.14 to an appendix (§6). We note that $\widetilde{K}_t(x)\equiv|\widetilde{\gamma}_t(x)|^2$ satisfies the size and Hölder continuity hypotheses of the function $H_t(x)$ of the lemma, with β_0 and α depending only on n, λ and Λ, as the reader may readily verify using (2.4). We omit the routine details.

3. Proof of the extrapolation lemma

We begin with a few preliminary observations. Recall (in the statement of Lemma 2.5) that

$$d\mu \equiv K_t(x)\,\frac{dx\,dt}{t}, \quad d\tilde{\mu} \equiv \widetilde{K}_t(x)\,\frac{dx\,dt}{t},$$

where $0 \leqslant K_t, \widetilde{K}_t \leqslant \beta_0$. Given a cube Q, let Q^* denote its immediate dyadic ancestor. We note that, for every $b \in [0,1]$, we have that

$$\sup_Q \frac{\mu((R^b_{Q^*} \setminus T_Q) \cap (Q \times (0,\infty)))}{|Q|} \leqslant C\beta_0 b, \tag{3.1}$$

as the reader may verify by an elementary computation. A similar computation shows that

$$\sup_Q \frac{\mu(R_Q \setminus T_Q)}{|Q|} \leqslant C\beta_0, \tag{3.2}$$

and moreover the same holds for $\tilde{\mu}$.

We now prove the following variant of Lemma 2.14:

LEMMA 3.3. *Suppose that*

$$d\tilde{\mu} \equiv \widetilde{K}_t(x)\,\frac{dx\,dt}{t}$$

with $0 \leqslant \widetilde{K}_t \leqslant \beta_0$. Suppose that there exists a number $\eta \in (0,1]$, and a number $\beta_1 \in (0,\infty)$, such that on every cube Q we have a decomposition $Q = E_Q \cup B_Q$, $E_Q \cap B_Q = \varnothing$, satisfying

(i) $|E_Q| > \eta |Q|$,

(ii) $B_Q = \bigcup Q_j$, *where the dyadic subcubes Q_j are non-overlapping, and*

(iii) $\tilde{\mu}(R_Q \setminus (\bigcup R_{Q_j})) \leqslant \beta_1 |Q|$.

Then $\tilde{\mu}$ is a Carleson measure, with

$$\|\tilde{\mu}\|_C \leqslant \frac{\beta_1}{\eta}.$$

Proof. By truncating in t, we may make the qualitative a priori assumption that $\tilde{\mu}$ is a Carleson measure. This assumption may be removed by a limiting argument.

Fix Q. We have that

$$\tilde{\mu}(R_Q) = \tilde{\mu}(R_Q \setminus (\bigcup R_{Q_j})) + \sum \tilde{\mu}(R_{Q_j}) \leqslant \beta_1 |Q| + \|\tilde{\mu}\|_C \sum_j |Q_j| \leqslant \beta_1 |Q| + \|\tilde{\mu}\|_C (1-\eta)|Q|.$$

Dividing by $|Q|$, and taking the supremum over all Q, we obtain the conclusion of Lemma 3.3. □

We remark that since $T_{Q_j} \subseteq R_{Q_j}$, the hypotheses of Lemma 3.3 will be verified if, in particular, $\tilde{\mu}(R_Q \setminus (\bigcup T_{Q_j})) \leqslant \beta_1 |Q|$.

Next, we prove the following "Calderón–Zygmund decomposition" for Carleson measures.

LEMMA 3.4. *Let μ be given as above, and let $a\geqslant 0$, $b\equiv 2^{-N}$ for some positive integer N. Suppose that Q is a cube such that*

$$\mu(R_Q)\leqslant (a+b)|Q|.$$

Then there exists a family $S=\{Q_k\}$ of non-overlapping dyadic subcubes of Q such that

$$\sup \frac{\mu(T_{Q'}\setminus(\bigcup T_{Q_k}))}{|Q'|}\leqslant C(1+\beta_0)b, \tag{3.5}$$

where the supremum runs over all dyadic subcubes $Q'\subseteq Q$, and

$$|B|\leqslant \frac{a+b}{a+2b}|Q|, \tag{3.6}$$

where B denotes the union of those Q_k such that $\mu(R^b_{Q_k})>a|Q_k|$.

Proof. If $\mu(R^b_Q)\leqslant a|Q|$, then let $S=\{Q\}$ and $B\equiv\varnothing$, and we are done. Otherwise, $\mu(R_Q\setminus R^b_Q)\leqslant b|Q|$. In this case, we perform a stopping time argument, subdividing Q dyadically and stopping the first time that

$$\mu((R_Q\setminus R^b_{Q'})\cap(Q'\times(0,\infty))>2b|Q'|. \tag{3.7}$$

Let S be the collection of selected cubes which are maximal with respect to (3.7). In particular, if $Q_k\in S$, and Q_k^* denotes its immediate dyadic ancestor (or "parent"), then

$$\mu((R_Q\setminus R^b_{Q_k^*})\cap(Q_k^*\times(0,\infty)))\leqslant 2b|Q_k^*|=2^{n+1}b|Q_k|. \tag{3.8}$$

We shall show that this collection S satisfies (3.5) and (3.6). To verify the latter, we note that by (3.7) and the definition of B, we have that

$$(a+2b)|B|\leqslant \sum\mu(R^b_{Q_k})+\sum\mu((R_Q\setminus R^b_{Q_k})\cap(Q_k\times(0,\infty)))\leqslant\mu(R_Q)\leqslant(a+b)|Q|,$$

and (3.6) follows.

We now proceed to show that (3.5) holds. Fix a dyadic cube Q', and observe that if $Q'\subseteq Q_k$, for some Q_k in S, then $\mu(T_{Q'}\setminus(\bigcup T_{Q_k}))=0$. Thus, we may suppose that Q' is not contained in any $Q_k\in S$. Then

$$T_{Q'}\setminus(\bigcup T_{Q_k})=(T_{Q'}\cap(E_0\times(0,\infty)))\cup\Big(\bigcup_{Q_k\subseteq Q'}((T_{Q'}\setminus T_{Q_k})\cap(Q_k\times(0,\infty)))\Big), \tag{3.9}$$

where $E_0\equiv Q\setminus(\bigcup Q_k)$. Note that by the stopping time construction, we have that

$$\frac{1}{|Q'|}\int_{Q'}\int_{bl(Q')}^{l(Q)}K_t(x)\,\frac{dt}{t}\,dx\leqslant 2b,$$

for all Q' which meet E_0. Hence, for a.e. $x\in E_0$, we have that

$$\int_0^{l(Q)} K_t(x)\,\frac{dt}{t} \leqslant 2b.$$

Integrating this last estimate over $Q'\cap E_0$, we obtain in particular that

$$\mu(T_{Q'}\cap(E_0\times(0,\infty))) \leqslant 2b|Q'|. \tag{3.10}$$

Moreover, if Q_k^* is the dyadic parent of Q_k, then

$$T_{Q'}\setminus T_{Q_k}\subset (R_{Q'}\setminus R^b_{Q_k^*})\cup(R^b_{Q_k^*}\setminus T_{Q_k}).$$

Hence by (3.10), (3.8) and (3.1) applied to Q_k, we have that the μ-measure of the set in (3.9) is no larger than

$$2b|Q'|+(2^{n+1}+C\beta_0)b\sum_{Q_k\subseteq Q'}|Q_k|,$$

and (3.5) follows. This concludes the proof of Lemma 3.4. $\square$

We now proceed to give the proof of the extrapolation theorem. The proof is based on an inductive, boot-strapping procedure. Our induction hypothesis is the following statement, which is defined for $a\geqslant 0$.

H(a): Let Q be a cube such that $\mu(R_Q)\leqslant a|Q|$. Then there exists numbers $\eta=\eta(a)\in(0,1]$, $\beta_2\equiv\beta_2(a)\equiv\beta_2(a,n,\beta_0,b)\in(0,\infty)$, and a decomposition $Q\equiv E_Q\cup B_Q$, $E_Q\cap B_Q=\varnothing$, with $|E_Q|\geqslant\eta|Q|$, and $B_Q\equiv\bigcup Q_j$, where $\{Q_j\}$ is a (possibly empty) collection of non-overlapping dyadic subcubes of Q, such that

$$\tilde{\mu}(R_Q\setminus(\bigcup T_{Q_j}))\leqslant\beta_2(1+C_1)|Q|.$$

The proof of the theorem proceeds now in two steps.

Step 1. Observe that $H(0)$ is true. Indeed, in this case $\mu(R_{Q'})=0$, for all dyadic subcubes $Q'\subseteq Q$, so by the hypotheses of our theorem, applied with $\psi\equiv 0$, we have $\tilde{\mu}(R_Q)\leqslant C_1|Q|$.

Step 2. Show that there exists $b>0$, depending only on n, β_0 and δ, such that for all $a\geqslant 0$, $H(a)\Rightarrow H(a+b)$.

Once Step 2 is completed, we are done. Indeed since $\|\mu\|_C\leqslant C_0$, we have that $H(C_0)$ can be achieved in finitely many steps, with the number of steps depending only on η,δ,β_0,C_0, in which case Lemma 3.3 may be invoked, with $\eta\equiv\eta(n,\delta,\beta_0,C_0)$ and $\beta_1=\beta_2(\eta,\delta,\beta_0,C_0)(1+C_1)$.

Let us now carry out Step 2. In order to do so, we first prove

LEMMA 3.11. *Suppose that $H(a)$ holds, that $b \equiv 2^{-N}$, and that Q is a cube for which*

$$\mu(R_Q^b) \leqslant a|Q|.$$

Then there exists a decomposition $Q = E_Q \cup B_Q$, $E_Q \cap B_Q = \varnothing$, with

(i) $|E_Q| \geqslant b^n \eta(a)|Q|$,

(ii) $B_Q = \bigcup Q_j$, *where Q_j are non-overlapping dyadic subcubes of Q, and*

(iii) $\tilde{\mu}(R_Q \setminus (\bigcup T_{Q_j})) \leqslant C(\beta_0, b, n, \beta_2(a))(1+C_1)|Q|$.

Proof. Write $Q \equiv \bigcup_{j=1}^{2^{Nn}} \widetilde{Q}_k$, where $\widetilde{Q}_k$ are non-overlapping dyadic subcubes of Q with side length $2^{-N} l(Q)$, and observe that there is at least one $\widetilde{Q}_k$, which we designate as $\widetilde{Q}_{k_0}$, such that

$$\mu(R_{\widetilde{Q}_{k_0}}) \leqslant a|\widetilde{Q}_{k_0}|,$$

since $R_Q^b \equiv \bigcup R_{\widetilde{Q}_k}$. As we are assuming that $H(a)$ holds, there exist non-overlapping dyadic subcubes $Q_j^{k_0} \subseteq \widetilde{Q}_{k_0}$ such that

$$|\widetilde{Q}_{k_0} \setminus (\bigcup Q_j^{k_0})| \geqslant \eta(a)|\widetilde{Q}_{k_0}| = \eta(a) 2^{-Nn}|Q|$$

and

$$\tilde{\mu}(R_{\widetilde{Q}_{k_0}} \setminus (\bigcup T_{Q_j^{k_0}})) \leqslant \beta_2(a)(1+C_1)|\widetilde{Q}_{k_0}| = \beta_2(a) 2^{-Nn}(1+C_1)|Q|. \tag{3.12}$$

Moreover,

$$\tilde{\mu}(R_Q \setminus R_Q^b) \leqslant \beta_0 \log(1/b)|Q|. \tag{3.13}$$

Thus, setting

$$B_Q \equiv \Big(\bigcup_{k \neq k_0} \widetilde{Q}_k\Big) \cup \big(\bigcup Q_j^{k_0}\big) \equiv \bigcup Q_j,$$

and invoking (3.2) (with $\tilde{\mu}$ in place of μ) in each $\widetilde{Q}_k$, $k \neq k_0$, (3.12) and (3.13), we obtain the conclusion of Lemma 3.11. $\square$

We now proceed to Step 2. Suppose that $a \geqslant 0$, that $H(a)$ holds, and that Q is a cube for which

$$\mu(R_Q) \leqslant (a+b)|Q|,$$

where we choose $b \equiv 2^{-N}$ so small that $\delta \geqslant C(1+\beta_0)b$ (this is the constant on the right side of (3.5)). By Lemma 3.4, there exists a family $S \equiv \{Q_k\}$ of non-overlapping dyadic subcubes satisfying (3.5) and (3.6). Let us denote by S' the subcollection of $Q_k \in S$ such that $\mu(R_{Q_k}^b) \leqslant a|Q_k|$. Let $(a+b)/(a+2b) \equiv 1-\theta$, and observe that either

$$|E_0| \equiv \Big|Q \setminus \Big(\bigcup_{Q_k \in S} Q_k\Big)\Big| \geqslant \tfrac{1}{2}\theta|Q| \tag{3.14}$$

or

$$\Big|\bigcup_{Q_k\in S'} Q_k\Big| \geqslant \tfrac{1}{2}\theta|Q|. \tag{3.15}$$

For each $Q_k\in S'$, we invoke Lemma 3.11 to construct a family $\{Q_j^k\}$ of non-overlapping dyadic subcubes of Q_k such that Lemma 3.11 (i)–(iii) hold with Q replaced by Q_k, B_Q replaced by $B_{Q_k}\equiv\bigcup Q_j^k$, and $E_{Q_k}\equiv Q_k\setminus B_{Q_k}$. We now define

$$B_Q\equiv\Big(\bigcup_{k:Q_k\in S'}\bigcup_j Q_j^k\Big)\cup\Big(\bigcup_{S\setminus S'} Q_k\Big), \tag{3.16}$$

and observe that its complement $E_Q\equiv Q\setminus B_Q\equiv E_0\cup\big(\bigcup_{k:Q_k\in S'} E_{Q_k}\big)$ satisfies

$$|E_Q|\geqslant b^n\eta(a)\tfrac{1}{2}\theta|Q|, \tag{3.17}$$

by virtue of (3.14), (3.15) and Lemma 3.11 (i) applied to every $Q_k\in S'$. Moreover, if we let S'' denote the collection of all the cubes whose union is the set B_Q in (3.16), then we have that

$$R_Q\setminus\Big(\bigcup_{Q''\in S''} T_{Q''}\Big)\subseteq\Big(R_Q\setminus\Big(\bigcup_{Q_k\in S} T_{Q_k}\Big)\Big)\cup\Big(\bigcup_{Q_k\in S'} T_{Q_k}\setminus\Big(\bigcup T_{Q_j^k}\Big)\Big)\equiv R_1\cup R_2.$$

Now, since (3.5) holds, with $C(1+\beta_0)b\leqslant\delta$, the hypotheses of the extrapolation theorem imply that $\tilde\mu(R_1)\leqslant C_1|Q|$. Also, since $T_{Q_k}\subseteq R_{Q_k}$, and since Lemma 3.11 applies to every $Q_k\in S'$, we obtain that

$$\tilde\mu(R_2)\leqslant C(\beta_0,b,n,\beta_2(a))(1+C_1)\sum_{Q_k\in S'}|Q_k|\leqslant C(\beta_0,b,n,\beta_2(a))(1+C_1)|Q|.$$

Thus, in view of (3.17), and the fact that $E_Q=Q\setminus\big(\bigcup_{Q''\in S''}Q''\big)$, we have that $H(a+b)$ holds. This concludes the proof of the extrapolation theorem. □

4. Deducing (2.3) from the extrapolation lemma

As mentioned in the previous section, we shall apply the extrapolation lemma to the measures $\tilde\mu$ and μ defined by (2.1) and (2.6), respectively. To establish (2.3), it is enough, given Lemma 2.5, to prove that there is a constant $\delta>0$, depending only on ellipticity and dimension, and a constant

$$C_1=C(\varepsilon^{-1})(1+\varepsilon_0^2)\|\tilde\mu\|_C, \tag{4.1}$$

such that

$$\tilde{\mu}(R_Q \cap \Omega_\psi) \leqslant C_1 |Q|, \tag{4.2}$$

whenever Q is a cube and ψ is a non-negative Lipschitz function with Lip norm at most 1, for which

$$\sup_{Q' \subseteq Q} |Q|^{-1} \mu(T_{Q'} \cap \Omega_\psi) \leqslant \delta. \tag{4.3}$$

Here, the supremum runs over dyadic subcubes of Q. Indeed, we have already observed above that the other hypotheses of the extrapolation lemma hold, with constants that depend only on ellipticity and dimension. Let us now proceed to prove that (4.2) holds, given (4.3).

To this end, we recall that $W_{t^2} \equiv e^{-t^2 L}$, and we denote the kernel of this operator by $W_{t^2}(x,y)$. Thus, $W_{(\varepsilon/2)^2 (l(Q))^2} \varphi \equiv F_Q$ (recall that F_Q was defined in (2.11), and satisfies (2.12)). We define also

$$F_\psi(x) \equiv W_{\varepsilon^2 (\psi(x))^2} \varphi(x).$$

Let $\varkappa Q$ denote the concentric cube with side length $\varkappa l(Q)$. Our fundamental estimate for F_Q is

LEMMA 4.4. *Fix a cube Q. Let $\psi(x)$ be a Lipschitz function, defined on $\mathbf{R}^n$, with $0 \leqslant \psi(x) \leqslant \varepsilon l(Q)$, for all $x \in Q$, and with $\|\nabla \psi\|_{L^\infty(\mathbf{R}^n)} \leqslant 1$. Then*

$$\int_{Q/8} |\nabla (F_Q - F_\psi)|^2 \, dx \leqslant C(\mu(T_Q \cap \Omega_\psi) + \varepsilon |Q|). \tag{4.5}$$

We shall defer the proof of this lemma until the next section. Let us now show that Lemma 4.4, together with (4.3), imply (4.2). Let $\widetilde{\mathcal{X}}(r)$ be a smooth cut-off function, with $\widetilde{\mathcal{X}}(r) \equiv 1$ if $r > 2$, $\widetilde{\mathcal{X}}(r) \equiv 0$ if $r < 1$, $0 \leqslant \widetilde{\mathcal{X}} \leqslant 1$, and we note that, by (2.4),

$$\widetilde{H}_t(x) \equiv |\widetilde{\gamma}_t(x)|^2 \widetilde{\mathcal{X}}\left(\frac{\psi(x)}{t}\right)$$

satisfies the size and Hölder continuity hypotheses of Lemma 2.14, with constants that depend only on ellipticity and dimension. Consequently, it is enough to prove that there exists $\eta > 0$, depending only on dimension and ε, such that for every dyadic $Q' \subseteq Q$, there is a set $E_{Q'} \subseteq Q'$ with $|E_{Q'}| > \eta |Q'|$, on which the following estimate holds:

$$\frac{1}{|Q'|} \int_{E_{Q'}} \int_{\psi(x)}^{l(Q')} |\widetilde{\gamma}_t(x)|^2 \frac{dt}{t} \, dx \leqslant C(\varepsilon^{-1})(1 + \varepsilon_0^2 \|\tilde{\mu}\|_C). \tag{4.6}$$

Indeed, given (4.6), we may apply Lemma 2.14 to the function $\widetilde{H}_t(x)$ defined above, to deduce that

$$\frac{1}{|Q|}\int_Q\int_{2\psi(x)}^{l(Q)}|\widetilde{\gamma}_t(x)|^2\,\frac{dt}{t}\,dx\leqslant C(\varepsilon^{-1})(1+\varepsilon_0^2\|\tilde{\mu}\|_C).$$

Since $\int_{\psi(x)}^{2\psi(x)}|\widetilde{\gamma}_t(x)|^2\,dt/t\leqslant C$, it follows that (4.2) will hold, once we have established (4.6).

We now show that (4.3) and Lemma 4.4 imply (4.6), as long as we choose δ and ε small enough depending only on ellipticity and dimension. We consider two cases: either $\psi(x)\leqslant\varepsilon l(Q')$, for all $x\in Q'$, or else there exists $x_0\in Q'$, with $\psi(x_0)>\varepsilon l(Q')$. In the latter case, since $\|\nabla\psi\|_\infty\leqslant 1$, we have that

$$|\psi(x)-\psi(x_0)|\leqslant\tfrac{1}{2}\varepsilon l(Q'),$$

as long as $x\in B(x_0)\equiv\{|x-x_0|\leqslant\frac{1}{2}\varepsilon l(Q')\}$. Thus, $\psi(x)\geqslant\frac{1}{2}\varepsilon l(Q')$ for $x\in B(x_0)\cap Q'\equiv E_{Q'}$. Since $x_0\in Q'$, it follows that $|E_{Q'}|\geqslant C^{-1}\varepsilon^n|Q'|$. We have also that

$$\int_{E_{Q'}}\int_{\psi(x)}^{l(Q')}|\widetilde{\gamma}_t(x)|^2\,\frac{dt}{t}\,dx\leqslant\|\widetilde{\gamma}_t\|_\infty^2\int_{E_{Q'}}\int_{\psi(x)}^{2\varepsilon^{-1/8}\psi(x)}\frac{dt}{t}\,dx\leqslant C\|\widetilde{\gamma}_t\|_\infty^2|Q'|\log\frac{1}{\varepsilon},$$

which yields (4.6) in the present case.

Otherwise, if $\psi(x)\leqslant\varepsilon l(Q')$ for all $x\in Q'$, then we may apply Lemma 4.4 to Q', and use (4.3) to obtain that

$$\int_{Q'/8}|\nabla(F_{Q'}-F_\psi)|^2\,dx\leqslant C(\delta+\varepsilon|Q'|)\leqslant C\varepsilon|Q'|,\tag{4.7}$$

if we set $\delta=\varepsilon$. Now, let M_r denote the Hardy–Littlewood maximal operator, taken with respect to balls of radius at most r. Then from (4.7) we deduce that

$$\int_{Q'/16}|M_{l(Q')/100}(\nabla(F_{Q'}-F_\psi))|^2\,dx\leqslant C\varepsilon|Q'|.\tag{4.8}$$

Hence,

$$\left|\left\{x\in\tfrac{1}{16}Q':M_{l(Q')/100}(\nabla(F_{Q'}-F_\psi))(x)>\varepsilon^{1/4}\right\}\right|\leqslant C\varepsilon^{1/2}|Q'|.$$

Thus, for ε small enough, there exists η depending only on n, and a set $E_{Q'}\subseteq\frac{1}{16}Q'$, with $|E_{Q'}|\geqslant\eta|Q'|$, and such that, for all $x\in E_{Q'}$, we have

$$M_{l(Q')/100}(\nabla(F_{Q'}-F_\psi))(x)\leqslant\varepsilon^{1/4}.\tag{4.9}$$

Now, $\int_{l(Q')/100}^{l(Q')}|\widetilde{\gamma}_t(x)|^2\,dt/t\leqslant C\|\widetilde{\gamma}_t\|_\infty^2\leqslant C$. Hence, in (4.6), it is enough to integrate over the t-interval $\psi(x)\leqslant t\leqslant\frac{1}{100}l(Q')$. By the triangle inequality,

$$\begin{aligned}|\widetilde{\gamma}_t(x)|&\leqslant|\widetilde{\gamma}_t(x)P_t(\nabla(I-W_{\varepsilon^2(\psi(x))^2})\varphi)(x)|\\&\quad+|\widetilde{\gamma}_t(x)P_t(\nabla(W_{\varepsilon^2(\psi(x))^2}-W_{\varepsilon^2(\psi(\cdot))^2})\varphi)(x)|\\&\quad+|\widetilde{\gamma}_t(x)P_t(\nabla(F_\psi-F_{Q'}))(x)|+|\widetilde{\gamma}_t(x)P_t(\nabla F_{Q'})(x)|\\&\equiv\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV},\end{aligned}\tag{4.10}$$

where $F_{Q'}\equiv e^{-(\varepsilon/2)^2(l(Q'))^2L}\varphi\equiv W_{(\varepsilon/2)^2(l(Q'))^2}\varphi$ and $F_\psi\equiv W_{\varepsilon^2\psi^2}\varphi$. Since $F_{Q'}$ satisfies (2.12), with Q' in place of Q, we may invoke Lemma 2.10 in Q', to deduce that

$$\int_{Q'}\int_0^{l(Q')}\mathrm{IV}^2\,\frac{dt}{t}\,dx\leqslant C(\varepsilon^{-1})|Q'|(1+\varepsilon_0^2\|\tilde{\mu}\|_C).$$

By (4.9), we have that, for $x\in E_{Q'}$ and $t<\frac{1}{100}l(Q')$,

$$\mathrm{III}\leqslant C\varepsilon^{1/4}|\tilde{\gamma}_t(x)|,$$

which we may hide on the left side of (4.10), if ε is small. Since φ is Lipschitz, and $\nabla P_t\equiv\vec{Q}_t/t$, where $\vec{Q}_t$ is an operator given by convolution with a smooth kernel which is supported in the the ball of radius t, we have that

$$\mathrm{I}\leqslant C\varepsilon\psi(x)t^{-1}|\tilde{\gamma}_t(x)|.$$

In proving (4.6), we only integrate where $t\geqslant\psi(x)$. It follows that $\mathrm{I}\leqslant C\varepsilon|\tilde{\gamma}_t(x)|$, which may also be hidden on the left side of (4.10). Finally,

$$|(W_{(\varepsilon\psi(y))^2}-W_{(\varepsilon\psi(x))^2})\varphi(y)|\leqslant\varepsilon|\psi(y)-\psi(x)|\sup_{t>0}\left|\frac{\partial}{\partial t}W_{t^2}\varphi(y)\right|\leqslant C\varepsilon|y-x|,\tag{4.11}$$

since $\|\nabla\psi\|_\infty\leqslant 1$, and

$$\left\|\sup_{t>0}\frac{\partial}{\partial t}W_{t^2}\varphi\right\|_\infty\leqslant C.$$

(In the last estimate, we have used that $\varphi\in\mathrm{Lip}_1$, and that the kernel $W_{t^2}\equiv e^{-t^2L}$ has Gaussian bounds.) But (4.11) implies that

$$|P_t(\nabla(W_{(\varepsilon\psi(\cdot))^2}-W_{(\varepsilon\psi(x))^2})\varphi)(x)|\leqslant Ct^{-n-1}\int_{\{|x-y|<t\}}|x-y|\,dy\cdot\varepsilon\leqslant C\varepsilon.$$

Thus, $\mathrm{II}\leqslant C\varepsilon|\tilde{\gamma}_t(x)|$, which may also be hidden on the left side of (4.10), if ε is chosen small enough depending only on n, λ and Λ. This proves (4.6), given Lemma 4.4. The proof of Theorem 1.6 is now complete, modulo Lemma 4.4 and Lemma 2.14. We give the proof of the former in the next section, and of the latter in §6.

5. Proof of Lemma 4.4

Throughout this section, Q is a fixed cube, with $\varrho\equiv l(Q)$. We recall that we are using the notation $W_{t^2}\equiv e^{-t^2L}$, and we denote the kernel of this operator by $W_{t^2}(x,y)$. Recall also that $W_{(\varepsilon/2)^2\varrho^2}\varphi\equiv F_Q$, and that $W_{\varepsilon^2(\psi(x))^2}\varphi\equiv F_\psi$.

Our goal is to prove

$$\int_{Q/8} |\nabla(F_Q - F_\psi)|^2\,dx \leqslant C(\mu(T_Q \cap \Omega_\psi) + \varepsilon|Q|), \tag{5.1}$$

whenever ψ is a Lipschitz function with $\|\nabla\psi\|_\infty \leqslant 1$ and satisfying $0 \leqslant \psi \leqslant \varepsilon l(Q)$ on Q. We begin by defining some cut-off functions. Fix Q, and let $\theta, \tilde\theta, \tilde{\tilde\theta} \in C_0^\infty$, $0 \leqslant \theta, \tilde\theta, \tilde{\tilde\theta} \leqslant 1$, $\|\nabla\theta\|_\infty + \|\nabla\tilde\theta\|_\infty + \|\nabla\tilde{\tilde\theta}\|_\infty \leqslant C/\varrho$, and suppose that $\theta \equiv 1$ on $\frac{1}{6}Q$, $\operatorname{supp}\theta \subseteq \frac{1}{4}Q$, $\tilde\theta \equiv 1$ on $\frac{1}{8}Q$, $\operatorname{supp}\tilde\theta \subseteq \frac{1}{7}Q$, $\tilde{\tilde\theta} \equiv 1$ on $\frac{3}{4}Q$, $\operatorname{supp}\tilde{\tilde\theta} \subseteq Q$. Let $x_Q \equiv$ center of Q. Now, since $W_\tau 1 = 1$, we have that $\nabla(W_\tau 1) = 0$ (even if τ depends on x). Hence, on the left side of (5.1), we may replace φ by $\varphi(\cdot) - \varphi(x_Q) \equiv \varphi_1 + \varphi_2$, where

$$\begin{aligned}\varphi_1(y) &\equiv (y - x_Q)\theta(y),\\ \varphi_2(y) &\equiv (y - x_Q)[1 - \theta(y)].\end{aligned}$$

Also, we may replace $\psi(x)$ by $\psi_Q(x) \equiv \psi(x)\tilde{\tilde\theta}(x)$. Our first step is to prove that

$$\int_{Q/8} |\nabla(W_{(\varepsilon/2)^2\varrho^2} - W_{(\varepsilon\psi_Q)^2})\varphi_2|^2\,dx \leqslant C\varepsilon|Q|. \tag{5.2}$$

By ellipticity and the definition of $\tilde\theta$, we have that the left side of (5.2) is dominated by a constant times

$$\begin{aligned}\int (\tilde\theta)^2 A\nabla(W_{(\varepsilon/2)^2\varrho^2} &- W_{(\varepsilon\psi_Q)^2})\varphi_2 \cdot \nabla(W_{(\varepsilon/2)^2\varrho^2} - W_{(\varepsilon\psi_Q)^2})\varphi_2\\ &= \int (\tilde\theta)^2 A\nabla W_{(\varepsilon/2)^2\varrho^2}\varphi_2 \cdot \nabla(W_{(\varepsilon/2)^2\varrho^2} - W_{(\varepsilon\psi_Q)^2})\varphi_2\\ &\qquad - \int (\tilde\theta)^2 A\nabla W_{\varepsilon^2\psi_Q^2} \cdot \nabla(W_{(\varepsilon/2)^2\varrho^2} - W_{\varepsilon^2\psi_Q^2})\varphi_2\\ &\equiv \mathrm{I} + \mathrm{II}.\end{aligned} \tag{5.3}$$

Since $\nabla W_{t^2}\varphi$ is a matrix, we should explain our notation: if $F = (F_1, ..., F_n)$ is a vector, then $|\nabla F|^2 \equiv \sum_j \nabla F_j \cdot \nabla F_j$ and $A\nabla F \cdot \nabla F \equiv \sum_j A\nabla F_j \cdot \nabla F_j$. Now

$$\begin{aligned}\mathrm{I} &= -\int \nabla(\tilde\theta)^2 \cdot A\nabla W_{(\varepsilon/2)^2\varrho^2}\varphi_2 (W_{(\varepsilon/2)^2\varrho^2} - W_{\varepsilon^2\psi_Q^2})\varphi_2\\ &\qquad + \int (\tilde\theta)^2 L W_{(\varepsilon/2)^2\varrho^2}\varphi_2 (W_{(\varepsilon/2)^2\varrho^2} - W_{\varepsilon^2\psi_Q^2})\varphi_2\\ &\equiv \mathrm{I}_1 + \mathrm{I}_2.\end{aligned}$$

We now claim that, for $t \leqslant \varepsilon \varrho$,

$$\int (\tilde{\theta})^2 |\nabla W_{t^2} \varphi_2|^2 \leqslant C\varepsilon^2 |Q|, \tag{5.4}$$

$$\int (\tilde{\theta})^2 |\nabla W_{\varepsilon^2 \psi_Q^2} \varphi_2|^2 \leqslant C\varepsilon^2 |Q| \tag{5.5}$$

(the latter will be used, and proved, later) and

$$\frac{1}{\varrho^2} \int_{Q/7} |(W_{2\varepsilon^2 \varrho^2} - W_{\varepsilon^2 \psi_Q^2}) \varphi_2|^2 \leqslant C\varepsilon^4 |Q|. \tag{5.6}$$

Assuming momentarily that the claim is valid, we immediately obtain that $|\mathrm{I}_1| \leqslant C\varepsilon^3 |Q| \leqslant C\varepsilon |Q|$ as desired. To prove the claim, and also to handle I_2, we note that for $t \leqslant \varepsilon \varrho$, $x \in \frac{1}{7} Q$, we have

$$|t^2 L W_{t^2} \varphi_2(x)| + |W_{t^2} \varphi_2(x)| \leqslant C t^{-n} \int_{|x-y|>C\varrho} e^{-|x-y|^2/Ct^2} |x-y| \, dy \leqslant C \frac{t^2}{\varrho} \leqslant C\varepsilon^2 \varrho. \tag{5.7}$$

(In the first inequality, we have used that $|x_Q - y| \approx |x-y|$, under the present circumstances.) This last bound yields (5.6) immediately, and also (5.4), by an argument similar to the proof of Caccioppoli's inequality. The proof of (5.5) is a bit more delicate, owing to the x-dependence of ψ_Q, and we defer it until the end of this section. Moreover, an application of (5.6) and (5.7) also yield the bound $|\mathrm{I}_2| \leqslant C\varepsilon^2 |Q| \leqslant C\varepsilon |Q|$ as desired.

Next, we turn to the bounds for II. We have

$$\begin{aligned} \mathrm{II} &= \int (\tilde{\theta})^2 A \nabla W_{\varepsilon^2 \psi_Q^2} \varphi_2 \cdot \nabla W_{\varepsilon^2 \psi_Q^2} \varphi_2 - \int (\tilde{\theta})^2 W_{\varepsilon^2 \psi_Q^2} \varphi_2 L W_{(\varepsilon/2)^2 \varrho^2} \varphi_2 \\ &\quad + \int \nabla (\tilde{\theta})^2 W_{\varepsilon^2 \psi_Q^2} \varphi_2 \cdot A \nabla W_{(\varepsilon/2)^2 \varrho^2} \varphi_2 \\ &\equiv \mathrm{II}_1 + \mathrm{II}_2 + \mathrm{II}_3. \end{aligned}$$

Now for $x \in \frac{1}{7} Q$, we have

$$|W_{\varepsilon^2 \psi_Q^2} \varphi_2(x)| \equiv |(W_{\varepsilon^2 \psi_Q^2} - I) \varphi_2(x)| \leqslant C\varepsilon \|\psi_Q\|_{L^\infty(Q)} \leqslant C\varepsilon^2 \varrho, \tag{5.8}$$

where we have used that φ_2 is Lipschitz with $\|\nabla \varphi_2\|_\infty \leqslant C$. The bound $|\mathrm{II}_3| \leqslant C\varepsilon |Q|$ now follows easily from (5.4) and (5.8), and we have $|\mathrm{II}_2| \leqslant C\varepsilon |Q|$, by (5.7) and (5.8). Also $|\mathrm{II}_1| \leqslant C\varepsilon^2 |Q|$, by (5.5). This concludes the proof of estimate (5.2), modulo the proof of (5.5), which we continue to defer for the moment.

We now return to the proof of Lemma 4.4 (that is, the proof of estimate (5.1)), and observe that by (5.2) we may replace φ by φ_1 on the left side of (5.1), which we then dominate by a constant times

$$\begin{aligned}\int_{\mathbf{R}^n} & A\nabla(W_{(\varepsilon/2)^2\varrho^2}-W_{\varepsilon^2\psi_Q^2})\varphi_1\cdot\nabla(W_{(\varepsilon/2)^2\varrho^2}-W_{\varepsilon^2\psi_Q^2})\varphi_1\\ &=\int_{\mathbf{R}^n}A\nabla W_{(\varepsilon/2)^2\varrho^2}\varphi_1\cdot\nabla W_{(\varepsilon/2)^2\varrho^2}\varphi_1-2\int_{\mathbf{R}^n}A\nabla W_{(\varepsilon/2)^2\varrho^2}\varphi_1\cdot\nabla W_{\varepsilon^2\psi_Q^2}\varphi_1\\ &\quad+\int_{\mathbf{R}^n}A\nabla W_{\varepsilon^2\psi_Q^2}\varphi_1\cdot\nabla W_{\varepsilon^2\psi_Q^2}\varphi_1\\ &\equiv V_1-2V_2+V_3.\end{aligned}\tag{5.9}$$

We note that, since $\|\nabla\varphi_1\|_2\leqslant C|Q|^{1/2}$, we have

$$\|(W_{\varepsilon^2\psi_Q^2}-\mathrm{I})\varphi_1\|_2\leqslant C\varepsilon\|\psi_Q\|_\infty|Q|^{1/2}\leqslant C\varepsilon^2\varrho|Q|^{1/2}.$$

Also,

$$\|LW_{(\varepsilon/2)^2\varrho^2}\varphi_1\|_2\leqslant\frac{C}{\varepsilon\varrho}|Q|^{1/2}.$$

Integrating by parts, and then combining the last two estimates, we see that we may replace $W_{\varepsilon^2\psi_Q^2}\varphi_1$ by φ_1, in V_2. Let us call the resulting term $\widetilde{V}_2$. The error $|V_2-\widetilde{V}_2|$ is on the order of $C\varepsilon|Q|$, which we allow. We write

$$\begin{aligned}\widetilde{V}_2&\equiv\int A\nabla W_{(\varepsilon/2)^2\varrho^2}\varphi_1\cdot\nabla\varphi_1=\int LW_{(\varepsilon/2)^2\varrho^2}\varphi_1\varphi_1\\ &=\int W_{\varepsilon^2\varrho^2/8}LW_{\varepsilon^2\varrho^2/8}\varphi_1\varphi_1=\int A\nabla W_{\varepsilon^2\varrho^2/8}\varphi_1\cdot\nabla W_{\varepsilon^2\varrho^2/8}\varphi_1.\end{aligned}\tag{5.10}$$

Next, we claim that we may replace V_3 by

$$\widetilde{V}_3\equiv\int_{\mathbf{R}^n}A(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\cdot(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q},\tag{5.11}$$

at the expense of introducing another allowable error. Indeed, by the chain rule,

$$\nabla W_{\varepsilon^2\psi_Q^2}\varphi_1-(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\equiv\left(\frac{\partial}{\partial t}W_{t^2}\varphi_1\Big|_{t=\varepsilon\psi_Q}\right)\varepsilon\nabla\psi_Q,$$

which is bounded in absolute value by

$$\varepsilon\sup_{t>0}\left|\frac{\partial}{\partial t}W_{t^2}\varphi_1\right|.$$

Moreover, for $t>0$, since φ_1 is Lipschitz, and since $(\partial/\partial t)W_{t^2}1=0$, we have that for all $x\in\mathbf{R}^n$ (and in particular for $x\in Q$),

$$\left|\frac{\partial}{\partial t}W_{t^2}\varphi_1(x)\right|\leqslant\frac{C}{t^{n+1}}\int_{\mathbf{R}^n}e^{-|x-y|^2/Ct^2}|x-y|\,dy\leqslant C.$$

Also, since φ_1 is supported in $\frac{1}{4}Q$, we have that for $x\in Q^c$,

$$\left|\frac{\partial}{\partial t}W_{t^2}\varphi_1(x)\right|\leqslant\frac{C}{t^{n+1}}\int_{|x-y|>\varrho}e^{-|x-y|^2/Ct^2}|\varphi_1|\,dy\leqslant C\varrho^{-1}M\varphi_1(x).$$

Combining the last two inequalities, we see that

$$\varepsilon\left\|\sup_{t>0}\left|\frac{\partial}{\partial t}W_{t^2}\varphi_1\right|\right\|_{L^2(\mathbf{R}^n)}\leqslant C\varepsilon|Q|^{1/2},$$

since $\|\varphi_1\|_2\leqslant C\varrho$. Thus, the claim that the difference $|V_3-\widetilde{V}_3|$ is small follows immediately from the inequality

$$\|(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\|_{L^2(\mathbf{R}^n)}\leqslant C|Q|^{1/2}. \tag{5.12}$$

We shall prove the latter estimate momentarily. Assuming for now that (5.12) holds, we therefore have that (5.9) equals $O(\varepsilon|Q|)$ plus

$$\begin{aligned}V_1-2\widetilde{V}_2+\widetilde{V}_3&\equiv\int_{\mathbf{R}^n}\int_{\varrho/(2\sqrt{2})}^{\varrho/2}\frac{\partial}{\partial t}(A\nabla W_{\varepsilon^2t^2}\varphi_1\cdot\nabla W_{\varepsilon^2t^2}\varphi_1)\,dt\,dx\\&\quad-\int_{\mathbf{R}^n}\int_{\psi_Q(x)}^{\varrho/(2\sqrt{2})}\frac{\partial}{\partial t}(A\nabla W_{\varepsilon^2t^2}\varphi_1\cdot\nabla W_{\varepsilon^2t^2}\varphi_1)\,dt\,dx\\&=-4\varepsilon^2\left\{\int_{\mathbf{R}^n}\int_{\varrho/(2\sqrt{2})}^{\varrho/2}tA\nabla W_{\varepsilon^2t^2}\varphi_1\cdot\nabla W_{\varepsilon^2t^2}L\varphi_1\,dt\,dx\right.\\&\quad\left.-\int_{\mathbf{R}^n}\int_{\psi_Q(x)}^{\varrho/(2\sqrt{2})}tA\nabla W_{\varepsilon^2t^2}\varphi_1\cdot\nabla W_{\varepsilon^2t^2}L\varphi_1\,dt\,dx\right\}\\&\equiv-4\varepsilon^2\{\Gamma_1-\Gamma_2\}.\end{aligned}$$

We note that

$$-4\varepsilon^2\Gamma_1\equiv-4\varepsilon^2\int_{\varrho/(2\sqrt{2})}^{\varrho/2}\int_{\mathbf{R}^n}|W_{\varepsilon^2t^2}L\varphi_1|^2\,dx\,t\,dt\leqslant 0.$$

We now claim also that

$$4\varepsilon^2\Gamma_2\leqslant C(\mu(T_Q\cap\Omega_\psi)+\varepsilon|Q|+\varepsilon|Q|^{1/2}\|(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\|_2). \tag{5.13}$$

Let us show that the last claim establishes the conclusion of Lemma 4.4. Indeed,

$$\begin{aligned}4\varepsilon^2\Gamma_2 &\equiv \widetilde{V}_3-\widetilde{V}_2\\ &\equiv \int_{\mathbf{R}^n} A(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\cdot(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}-\int_{\mathbf{R}^n} A\nabla W_{\varepsilon^2\varrho^2/8}\varphi_1\cdot\nabla W_{\varepsilon^2\varrho^2/8}\varphi_1.\end{aligned}$$

Combining this last identity with (5.13), ellipticity, the fact that μ is a Carleson measure, and the fact that

$$\|\nabla W_{\varepsilon^2\varrho^2/8}\varphi_1\|_{L^2(\mathbf{R}^n)}\leqslant C|Q|^{1/2}$$

(the proof of which is routine, and omitted), and then hiding a small term on the left, we obtain (5.12). But given (5.12) and (5.13), the conclusion of Lemma 4.4 follows, since we have shown that (5.9) is dominated by $4\varepsilon^2\Gamma_2$ plus small errors. It is therefore enough to prove (5.13).

To this end, let $\Omega\equiv\{(x,t)\in\mathbf{R}^{n+1}_+:\psi_Q(x)<t<\varrho/(2\sqrt{2})\}$. Then,

$$\begin{aligned}4\varepsilon^2\Gamma_2 &= 4\varepsilon^2\iint_\Omega tA\nabla W_{\varepsilon^2t^2}\varphi_1\cdot\nabla W_{\varepsilon^2t^2}L\varphi_1\,dt\,dx\\ &=4\varepsilon^2\iint_\Omega \operatorname{div}_x(tA\nabla W_{\varepsilon^2t^2}\varphi_1 W_{\varepsilon^2t^2}L\varphi_1)\,dt\,dx+4\varepsilon^2\iint_\Omega t|W_{\varepsilon^2t^2}L\varphi_1|^2\,dx\,dt\\ &\equiv 4\varepsilon^2\Gamma'+4\varepsilon^2\Gamma''.\end{aligned}$$

We observe that, by the divergence theorem,

$$\begin{aligned}4\varepsilon^2\Gamma' &= 4\varepsilon^2\iint_\Omega \operatorname{div}_{x,t}[(tA\nabla W_{\varepsilon^2t^2}\varphi_1,0)W_{\varepsilon^2t^2}L\varphi_1]\\ &=4\varepsilon^2\int_{\partial\Omega}\langle N,(A\nabla W_{\varepsilon^2t^2}\varphi_1,0)\rangle\, tW_{\varepsilon^2t^2}L\varphi_1\,d\sigma(x,t),\end{aligned}$$

where N denotes the outer unit normal to $\partial\Omega$. But along the "top" part of $\partial\Omega$, when $t\equiv\varrho/(2\sqrt{2})$, we have that $N\equiv(0,...,0,1)$. Hence

$$|4\varepsilon^2\Gamma'|\leqslant C\varepsilon^2\int_{\mathbf{R}^n}|\nabla W_{\varepsilon^2t^2}\varphi_1|_{t=\psi_Q}|\cdot|\psi_Q(x)W_{\varepsilon^2\psi_Q^2}L\varphi_1|\sqrt{1+(\nabla\psi_Q)^2}\,dx.$$

But $\psi_Q(x)\equiv\psi(x)\tilde{\tilde{\theta}}(x)$, so that

$$\|\nabla\psi_Q\|_\infty\leqslant\|\nabla\psi\|_{L^\infty(Q)}+\frac{C}{\varrho}\|\psi\|_{L^\infty(Q)}\leqslant C.$$

In particular, $\sqrt{1+|\nabla\psi_Q|^2}\leqslant C$. Moreover,

$$|\varepsilon\psi_Q(x)W_{\varepsilon^2\psi_Q^2}L\varphi_1(x)|\leqslant\sup_{t>0}\left|tW_{t^2}\sqrt{L}\,\sqrt{L}\,\varphi_1\right|\leqslant CM(\sqrt{L}\,\varphi_1),$$

where M denotes the Hardy–Littlewood maximal operator. Consequently,

$$|4\varepsilon^2\Gamma'| \leqslant C\varepsilon \big\|\sqrt{L}\,\varphi_1\big\|_2 \|(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\|_2 \leqslant C\varepsilon|Q|^{1/2}\|(\nabla W_{\varepsilon^2t^2}\varphi_1)|_{t=\psi_Q}\|_2,$$

as desired.

It remains to treat the term $4\varepsilon^2\Gamma''$, which equals

$$4\int_{\mathbf{R}^n}\int_{\psi_Q(x)}^{\varrho/(2\sqrt{2})} |e^{-\varepsilon^2t^2L}\varepsilon tL\varphi_1(x)|^2\,\frac{dt}{t}\,dx.$$

Let $k_{\varepsilon t}(x,y)$ denote the kernel of $e^{-\varepsilon^2t^2L}\varepsilon tL$. Clearly, $\int k_{\varepsilon t}(x,y)\,dy=0$. Hence,

$$\begin{aligned}\int k_{\varepsilon t}(x,y)\varphi_1(y)\,dy &= \int k_{\varepsilon t}(x,y)[\varphi_1(y)-\varphi_1(x)-((y-x)\cdot\nabla)\varphi_1(x)]\,dy\\ &\quad+\Big(\int k_{\varepsilon t}(x,y)y\,dy\cdot\nabla\Big)\varphi_1(x)\\ &\equiv f_{\varepsilon t}(x)+g_{\varepsilon t}(x).\end{aligned}$$

Now, $\|\nabla\varphi_1\|_{L^\infty}\leqslant C$, and $\nabla\varphi_1$ is supported in $\frac14Q$. Moreover, for all $x\in\frac14Q$, we have that $\operatorname{dist}(x,Q^c)\geqslant\frac38>1/(2\sqrt{2})$, so that $\frac14Q\times(0,\varrho/(2\sqrt{2}))\subseteq T_Q$. Also,

$$\int k_{\varepsilon t}(x,y)y\,dy = e^{-\varepsilon^2t^2L}\varepsilon tL\varphi(x).$$

Thus

$$\int_{\mathbf{R}^n}\int_{\psi_Q(x)}^{l(Q)/(2\sqrt{2})} |g_{\varepsilon t}(x)|^2\,\frac{dt}{t}\,dx \leqslant C\mu(T_Q\cap\Omega_\psi)$$

as desired, since $\psi_Q=\psi$ on $\frac14Q$.

Next, we note that $\|\nabla^2\varphi_1\|_\infty\leqslant C\varrho^{-1}$, so that the expression in square brackets in the definition of $f_{\varepsilon t}(x)$ is dominated in absolute value by

$$C\frac{|y-x|^2}{\varrho}.$$

Therefore,

$$|f_{\varepsilon t}(x)|\leqslant C\frac{1}{\varepsilon t\varrho}\int(\varepsilon t)^{-n}e^{-|x-y|^2/(C\varepsilon t)^2}|x-y|^2\,dy\leqslant C\varepsilon t\varrho, \tag{5.14}$$

and consequently,

$$\int_{2Q}\int_0^\varrho|f_{\varepsilon t}(x)|^2\,\frac{dt}{t}\,dx\leqslant C\varepsilon^2|Q|.$$

On the other hand, for $x\in(2Q)^c$, $\varphi_1(x)$ and $\nabla\varphi_1(x)=0$, so that, in the definition of $f_{\varepsilon t}(x)$, we may multiply the integrand by $\mathcal{X}_{3Q/2}(y)$. Hence, for $x\in(2Q)^c$, we have by the same computation as in (5.14), with $\mathcal{X}_{3Q/2}(y)$ inserted in the integral, that

$$|f_{\varepsilon t}(x)|\leqslant C\frac{\varepsilon t}{\varrho}M(\mathcal{X}_{3Q/2})(x),$$

which in turn implies that

$$\int_{(2Q)^c}\int_0^\varrho |f_{\varepsilon t}(x)|^2\,\frac{dt\,dx}{t}\leqslant C\varepsilon^2|Q|.$$

We have thus established (5.13). Modulo the proof of estimate (5.5), which we had deferred, the proof of Lemma 4.4 is now complete.

It remains only to prove (5.5). We note first that, by the chain rule,

$$\nabla W_{\varepsilon^2\psi_Q^2}\varphi_2-(\nabla W_{\varepsilon^2t^2}\varphi_2)|_{t=\psi_Q}\equiv\left(\frac{\partial}{\partial t}W_{t^2}\varphi_2\right)\bigg|_{t=\varepsilon\psi_Q}\varepsilon\nabla\psi_Q. \tag{5.15}$$

Consequently, by (5.7), and the fact that $\partial W_{t^2}/\partial t=-2tLW_{t^2}$,

$$\int(\tilde\theta)^2|\nabla W_{\varepsilon^2\psi_Q^2}\varphi_2-(\nabla W_{\varepsilon^2t^2}\varphi_2)|_{t=\psi_Q}|^2\leqslant C\varepsilon^4|Q|. \tag{5.16}$$

Thus, by (5.4), (5.16) and ellipticity, the left side of (5.5) is dominated by a constant times

$$\int(\tilde\theta)^2A(\nabla W_{\varepsilon^2t^2}\varphi_2)|_{t=\psi}\cdot(\nabla W_{\varepsilon^2t^2}\varphi_2)|_{t=\psi}-\int(\tilde\theta)^2A\nabla W_{\varepsilon^2\varrho^2}\varphi_2\cdot\nabla W_{\varepsilon^2\varrho^2}\varphi_2+O(\varepsilon^2|Q|)$$

(here we have used that $\psi_Q\equiv\psi$ on $\operatorname{supp}\tilde\theta$)

$$\begin{aligned}
&=-\int(\tilde\theta)^2\int_{\psi(x)}^{\varrho}\frac{\partial}{\partial t}(A\nabla W_{\varepsilon^2t^2}\varphi_2\cdot\nabla W_{\varepsilon^2t^2}\varphi_2)+O(\varepsilon^2|Q|)\\
&=4\int(\tilde\theta)^2\int_{\psi(x)}^{\varrho}\varepsilon^2t\nabla e^{-\varepsilon^2t^2L}L\varphi_2\cdot A\nabla W_{\varepsilon^2t^2}\varphi_2+O(\varepsilon^2|Q|)\\
&=4\int(\tilde\theta)^2\int_{\psi(x)}^{\varrho}|e^{-\varepsilon^2t^2L}\varepsilon tL\varphi_2(x)|^2\,\frac{dt}{t}\,dx\\
&\quad-4\int\nabla(\tilde\theta)^2\int_{\psi(x)}^{\varrho}\varepsilon^2te^{-\varepsilon^2t^2L}L\varphi_2\cdot A\nabla W_{\varepsilon^2t^2}\varphi_2\,dt\,dx\\
&\quad+\int_{\partial\{(x,t)\in\mathbf{R}^{n+1}_+:\psi(x)<t<\varrho\}}(\tilde\theta)^2\varepsilon^2te^{-\varepsilon^2t^2L}L\varphi_2\langle N,(A\nabla W_{\varepsilon^2t^2}\varphi_2,0)\rangle\,d\sigma(x,t)+O(\varepsilon^2|Q|)\\
&\equiv Z_1+Z_2+Z_3+O(\varepsilon^2|Q|),
\end{aligned}$$

where we have obtained the last equality by the same argument, involving the divergence theorem, that we had used to treat $4\varepsilon^2\Gamma_2$ above. By (5.7), and the definition of $\tilde{\theta}$, we have that

$$Z_1 \leqslant C \int_{Q/7} \int_0^{\varrho} \left(\frac{\varepsilon t}{\varrho}\right)^2 \frac{dt}{t}\, dx = C\varepsilon^2 |Q|.$$

Also, by (5.7), and then Schwarz's inequality and (5.4),

$$\begin{aligned}|Z_2| &\leqslant C\varepsilon^2 \int \tilde{\theta} \int_0^{\varrho} \frac{t}{\varrho^2} |\nabla W_{\varepsilon^2 t^2} \varphi_2|\, dt\, dx \\ &\leqslant C\varepsilon^2 \int_0^{\varrho} \frac{t}{\varrho^2} |Q|^{1/2} \left(\int (\tilde{\theta})^2 |\nabla W_{\varepsilon^2 t^2} \varphi_2(x)|^2\, dx\right)^{1/2} dt \leqslant C\varepsilon^3 |Q|.\end{aligned}$$

Finally, since $N=(0, ..., 0, 1)$ along the "upper" boundary $t=\varrho$, we have that, by (5.7),

$$\begin{aligned}|Z_3| &\leqslant C\varepsilon \int (\tilde{\theta})^2 |\varepsilon\psi(x) e^{-\varepsilon^2(\psi(x))^2 L} L\varphi_2(x)| \cdot |(\nabla W_{\varepsilon^2 t^2}\varphi_2)|_{t=\psi}|\, dx \\ &\leqslant C\varepsilon^2 \frac{\|\psi\|_\infty}{\varrho} \int (\tilde{\theta})^2 |(\nabla W_{\varepsilon^2 t^2}\varphi_2)|_{t=\psi}|\, dx.\end{aligned}$$

Using the elementary inequality $a \leqslant \frac{1}{2}(1+a^2)$, (5.16) and the fact that $\|\psi\|_\infty \leqslant \varepsilon\varrho$, we see that the last expression is dominated by

$$C\varepsilon^3 \left[|Q| + \int |\nabla W_{\varepsilon^2 \psi^2} \varphi_2|^2 (\tilde{\theta})^2\, dx\right].$$

For ε small enough, we may hide the second summand on the left side of (5.5), and the proof of Lemma 4.4 is now complete.

6. Appendix: Proof of Lemma 2.14

Fix Q. Our goal is to establish the estimate

$$\frac{1}{|Q|} \int_Q \int_0^{l(Q)} H_t(x) \frac{dt}{t}\, dx \leqslant C(\eta, \alpha)(\beta_0 + \beta), \tag{6.1}$$

given the hypotheses of Lemma 2.14.

Let $0<\varepsilon<l(Q)$ (we remark that the present ε has no connection with the number used in §§2–5; it is now merely a small, arbitrary number). Let

$$M(\varepsilon) \equiv \sup_{Q' \subseteq Q} \frac{1}{|Q'|} \int_\varepsilon^{l(Q')} \int_{Q'} H_t(x)\, dx \frac{dt}{t},$$

where the sup runs over all dyadic subcubes $Q'\subseteq Q$, and where the integral is taken to be zero if $l(Q')\leqslant\varepsilon$. Clearly then, $M(\varepsilon)<\infty$, and $x\to\int_\varepsilon^{l(Q)} H_t(x)\,dt/t$ is continuous.

Let $N\equiv(2/\eta)\beta$. Then the set $\Omega\equiv\{x\in Q:\int_\varepsilon^{l(Q)} H_t(x)\,dt/t>N\}$ is open, and moreover,

$$|\Omega|\leqslant|Q\backslash E|+\frac{\beta}{N}|Q|\leqslant\left(1-\tfrac{1}{2}\eta\right)|Q|, \tag{6.2}$$

by Chebyshev's inequality and the hypotheses of Lemma 2.14.

Next let $\Omega\equiv\bigcup Q_j$ denote the usual Whitney decomposition of Ω. Then

$$\begin{aligned}\int_\varepsilon^{l(Q)}\int_Q H_t(x)\,dx\,\frac{dt}{t}&\leqslant N|Q\backslash\Omega|+\sum_j\int_\varepsilon^{l(Q)}\int_{Q_j} H_t(x)\,dx\,\frac{dt}{t}\\&\leqslant N|Q|+\sum_j\int_\varepsilon^{l(Q_j)}\int_{Q_j} H_t(x)\,dx\,\frac{dt}{t}\\&\quad+\sum_j\int_{\max(l(Q_j),\varepsilon)}^{l(Q)}\int_{Q_j} H_t(x)\,dx\,\frac{dt}{t},\end{aligned}\tag{6.3}$$

where again we use the convention that the integrals in the middle term are zero if $l(Q_j)\leqslant\varepsilon$.

By the Whitney construction, there exists $x_j\in Q\backslash\Omega$, with $\operatorname{dist}(x_j,Q_j)\leqslant Cl(Q_j)$. We therefore have

$$\int_{Q_j} H_t(x)\,dx=\int_{Q_j}(H_t(x)-H_t(x_j))\,dx+\int_{Q_j} H_t(x_j)\,dx\leqslant\left[C\beta_0\left(\frac{l(Q_j)}{t}\right)^\alpha+H_t(x_j)\right]|Q_j|.$$

Hence,

$$\int_{\max(l(Q_j),\varepsilon)}^{l(Q)}\int_{Q_j} H_t(x)\,dx\,\frac{dt}{t}\leqslant\left(\frac{C}{\alpha}\beta_0+N\right)|Q_j|,$$

since $x_j\in Q\backslash\Omega$. Thus, returning to (6.3), we obtain that

$$\begin{aligned}\int_\varepsilon^{l(Q)}\int_Q H_t(x)\,dx\,\frac{dt}{t}&\leqslant\left(\frac{C}{\alpha}\beta_0+2N\right)|Q|+M(\varepsilon)|\Omega|\\&\leqslant C(\alpha,\eta)(\beta_0+\beta)|Q|+M(\varepsilon)\left(1-\tfrac{1}{2}\eta\right)|Q|,\end{aligned}\tag{6.4}$$

where in the last inequality we have used (6.2). But we can repeat the previous argument to show that (6.4) holds also with Q replaced by any dyadic subcube $Q'\subseteq Q$. Thus

$$M(\varepsilon)\leqslant\frac{2}{\eta}C(\alpha,\eta)(\beta_0+\beta),$$

and the conclusion of the lemma follows by letting $\varepsilon\to0$.

References

[1] Auscher, P., Regularity theorems and heat kernels for elliptic operators. *J. London Math. Soc.* (*2*), 54 (1996), 284–296.

[2] Auscher, P. & Tchamitchian, P., *Square Root Problem for Divergence Operators and Related Topics.* Astérisque, 249. Soc. Math. France, Paris, 1998.

[3] Christ, M., *Lectures on Singular Integral Operators.* CBMS Regional Conf. Ser. in Math., 77. Conf. Board Math. Sci., Washington, DC, 1990.

[4] Coifman, R. R., Deng, D. G. & Meyer, Y., Domains de la racine carrée de certains opérateurs différentiels accrétifs. *Ann. Inst. Fourier* (*Grenoble*), 33 (1983), 123–134.

[5] Coifman, R. R., McIntosh, A. & Meyer, Y., L'intégrale de Cauchy définit un opérateur borné sur L^2 pour les courbes lipschitziennes. *Ann. of Math.* (*2*), 116 (1982), 361–387.

[6] Coifman, R. R. & Meyer, Y., Nonlinear harmonic analysis, operator theory and P.D.E., in *Beijing Lectures in Harmonic Analysis* (Beijing, 1984), pp. 3–45. Ann. of Math. Stud., 112. Princeton Univ. Press, Princeton, NJ, 1986.

[7] Dahlberg, B. E. J. & Kenig, C. E., *Harmonic Analysis and Partial Differential Equations.* Department of Mathematics, Chalmers University of Technology and the University of Göteborg, Göteborg, 1985.

[8] Dorronsoro, J. R., A characterization of potential spaces. *Proc. Amer. Math. Soc.*, 95 (1985), 21–31.

[9] Fabes, E. B., Jerison, D. S. & Kenig, C. E., Multilinear square functions and partial differential equations. *Amer. J. Math.*, 107 (1985), 1325–1368.

[10] — Necessary and sufficient conditions for absolute continuity of elliptic-harmonic measure. *Ann. of Math.* (*2*), 119 (1984), 121–141.

[11] Hofmann, S. & Lewis, J. L., *The Dirichlet Problem for Parabolic Operators with Singular Drift Terms.* Mem. Amer. Math. Soc., 151 (719). Amer. Math. Soc., Providence, RI, 2001.

[12] Hofmann, S. & McIntosh, A., The solution of the Kato problem in two dimensions. Unpublished manuscript.

[13] Journé, J.-L., Remarks on Kato's square-root problem, in *Mathematical Analysis* (El Escorial, 1989). *Publ. Math.*, 35 (1991), 299–321.

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

[15] — Integration of the equation of evolution in a Banach space. *J. Math. Soc. Japan*, 5 (1953), 208–234.

[16] Kenig, C. & Meyer, Y., Kato's square roots of accretive operators and Cauchy kernels on Lipschitz curves are the same, in *Recent Progress in Fourier Analysis* (El Escorial, 1983), pp. 123–143. North-Holland Math. Stud., 111. North-Holland, Amsterdam, 1985.

[17] Lacey, M., Personal communication.

[18] Lewis, J. L. & Murray, M., *The Method of Layer Potentials for the Heat Equation in Time-Varying Domains.* Mem. Amer. Math. Soc., 114 (545). Amer. Math. Soc., Providence, RI, 1995.

[19] McIntosh, A., Square roots of operators and applications to hyperbolic PDEs, in *Miniconference on Operator Theory and Partial Differential Equations* (Canberra, 1983), pp. 124–136. Proc. Centre Math. Anal. Austral. Nat. Univ., 5. Austral. Nat. Univ., Canberra, 1984.

[20] — On the comparability of $A^{1/2}$ and $A^{*1/2}$. *Proc. Amer. Math. Soc.*, 32 (1972), 430–434.

[21] — On representing closed accretive sesquilinear forms as $(A^{1/2}u, A^{*1/2}v)$, in *Nonlinear Partial Differential Equations and Their Applications, Collège de France Seminar, Vol.* III (Paris, 1980/1981), pp. 252–267. Res. Notes in Math., 70. Pitman, Boston, MA–London, 1982.

[22] — The square root problem for elliptic operators: a survey, in *Functional-Analytic Methods for Partial Differential Equations* (Tokyo, 1989), pp. 122–140. Lecture Notes in Math., 1450. Springer-Verlag, Berlin, 1990.

PASCAL AUSCHER
Faculté de Mathématiques et d'Informatique
Université de Picardie–Jules Verne
LAMFA, CNRS-FRE 2270
33, rue Saint Leu
FR-80039 Amiens Cedex 1
France
auscher@mathinfo.u-picardie.fr

STEVE HOFMANN
Department of Mathematics
University of Missouri at Columbia
Columbia, MO 65211
U.S.A.
hofmann@math.missouri.edu

JOHN L. LEWIS
Department of Mathematics
University of Kentucky
Lexington, KY 40506-0027
U.S.A.
john@ms.uky.edu

PHILIPPE TCHAMITCHIAN
Faculté des Sciences et Techniques de Saint-Jérôme
Université d'Aix–Marseille III
LATP, CNRS-UMR 6632
Avenue Escadrille Normandie–Niemen
FR-13397 Marseille Cedex 20
France
tchamphi@math.u-3mrs.fr

Received May 22, 2000