Michigan Math. J. 51 (2003)

Sharp Estimate of the Ahlfors–Beurling Operator via Averaging Martingale Transforms

Oliver Dragičević & Alexander Volberg

1. Introduction

Our most important object will be the so-called two-dimensional martingale transform. In order to define it properly, we should start with the notion of a Haar basis.

We call the family $\mathcal{L} := \{[m2^n, (m+1)2^n] \mid m, n \in \mathbb{Z}\}$ the standard *dyadic lattice.* Observe that 0 is the only real number that is not contained in the interior of any *dyadic interval,* that is, any member of $\mathcal{L}$. Each interval $I \subset \mathbb{R}$ gives rise to its *Haar function* h_I, defined by

$$h_I := |I|^{-1/2}(\chi_{I_+} - \chi_{I_-}),$$

where I_- and I_+ denote (respectively) the left and the right half of the interval I and χ_E stands for the characteristic function of the set E, as usual. It is a well-known fact that the set $\{h_I \mid I \in \mathcal{L}\}$ forms an orthonormal basis of the space $L^2(\mathbb{R})$.

At this point we should emphasize that our attention will be concentrated on the planar case. Toward that end we shall introduce a similar basis for the space $L^2(\mathbb{R}^2)$. This will be described in detail in the continuation of this preface.

We may now define the operator T_σ on $L^2(\mathbb{R})$ by

$$T_\sigma f := \sum_{I \in \mathcal{L}} \sigma(I)\langle f, h_I\rangle h_I,$$

where $\sigma : \mathcal{L} \to \{-1, 1\}$ is arbitrary. Such operators are called *martingale transforms.* Observe that T_σ is an isometry satisfying $T_\sigma^2 = I$.

The symbol $\langle f\rangle_I$ shall stand for $|I|^{-1}\int_I f\,dm$, the average of the function f over the interval I. We say that a measurable function $w : \mathbb{R} \to \mathbb{R}_+$ satisfies the *dyadic A_2 condition* if

$$Q_{w,2} := \sup_{I \in \mathcal{L}} \langle w\rangle_I \langle w^{-1}\rangle_I < \infty.$$

It is well known that, for any such weight w, the martingale transforms are uniformly bounded on $L^2(w)$, that is, on the Hilbert space of measurable complex functions on $\mathbb{R}$ endowed with the scalar product

$$\langle f, g\rangle_w := \int_{\mathbb{R}} f(x)\overline{g(x)}w(x)\,dx.$$

Received April 5, 2002. Revision received June 12, 2002.
The second author is partially supported by NSF Grant nos. DMS-9970395 and DMS-0200713.

More precisely, in [9] it was shown that there is a constant $C > 0$ such that, for arbitrary choice of functions $w \in A_2$, $\sigma \colon \mathcal{L} \to \{-1, 1\}$, and $f \in L^2(w)$, we have the condition

$$\|T_\sigma f\|_{L^2(w)} \le C Q_{w,2} \|f\|_{L^2(w)}. \tag{1}$$

The estimate (1) is sharp in the sense that one cannot replace $CQ_{w,2}$ in it by $\phi(Q_{w,2})$, where ϕ grows slower than a linear function.

Our interest will focus on studying the case of $\mathbb{R}^2$ instead of $\mathbb{R}$, and a certain important singular integral operator on $\mathbb{R}^2$ will play the role of T_σ. In the planar case, all the definitions simply proceed from the one-dimensional case in a natural way. Thus the term *dyadic lattice* will now stand for the collection of all squares of the form $I \times J \subset \mathbb{R}^2$, where I and J are dyadic intervals of the same length. To each such square $Q = I \times J$ we will assign three Haar functions:

$$\begin{aligned} h_Q^I(s,t) &:= h_I(s)\chi_J(t)|J|^{-1/2}, \\ h_Q^J(s,t) &:= |I|^{-1/2}\chi_I(s)h_J(t), \\ h_Q(s,t) &:= h_I(s)h_J(t). \end{aligned}$$

As previously, one can verify that the set $\{h_Q^I, h_Q^J, h_Q \mid Q \in \mathcal{L}\}$ builds an orthonormal basis in $L^2(\mathbb{R}^2)$. Now the *two-dimensional martingale transform* becomes the operator

$$T_\sigma f := \sum_{Q\in\mathcal{L}} \sigma_I(Q)\langle f, h_Q^I\rangle h_Q^I + \sum_{Q\in\mathcal{L}} \sigma_J(Q)\langle f, h_Q^J\rangle h_Q^J + \sum_{Q\in\mathcal{L}} \sigma(Q)\langle f, h_Q\rangle h_Q,$$

where, as before, $\sigma_I, \sigma_J, \sigma \colon \mathcal{L} \to \{-1, 1\}$ and $f \in L^2(\mathbb{R}^2)$ are arbitrary functions.

The term (two-dimensional) A_2 *weight* now stands for a positive measurable function w on $\mathbb{R}^2$ such that

$$Q_{w,2} := \sup_{Q\subset\mathbb{R}^2} \langle w\rangle_Q \langle w^{-1}\rangle_Q < \infty.$$

Here, unlike previously, the supremum runs over all squares in $\mathbb{R}^2$, not merely the dyadic ones; in the latter case we would be referring to the *dyadic* A_2 *weight*. Certainly, $\langle\cdot\rangle_Q$ denotes the average over the square Q with respect to the planar Lebesgue measure. One can immediately obtain a two-dimensional version of Wittwer's result (1).

Let us also introduce

$$Q_{w,p} := \sup_{Q\subset\mathbb{R}^2} \langle w\rangle_Q \langle w^{-1/(p-1)}\rangle_Q^{p-1}, \quad 1 < p < \infty.$$

We shall be studying the operator $T \colon L^2(w) \to L^2(w)$ of convolution with the kernel z^{-2}, that is,

$$Tf(x,y) = \iint_{\mathbb{R}\times\mathbb{R}} \frac{f(x-u, y-v)}{(u+iv)^2}\, du\, dv.$$

Here w is a planar A_2 weight, of course. The operator T, sometimes multiplied by $1/\pi$, is called the *Ahlfors–Beurling operator*.

Our main result is the following theorem.

THEOREM. *T is in the weakly closed linear span of operators of the type T_σ.*

This yields an immediate corollary as follows.

COROLLARY. $\|T\|_{L^p(w)\to L^p(w)} \le C(p)Q_{w,p}$.

It is enough to prove the Corollary for $p = 2$ (see [7]). This corollary was the main goal of [7], where it was proved by different methods. It seems to us that our proof is much more streamlined and perhaps more conceptual.

The theorem looks slightly unexpected because the same result would not be true for T replaced by the first-order Riesz transforms on the plane. In fact, all our operators T_σ have symmetric kernels k_σ, meaning that $k_\sigma(x, y) = k_\sigma(y, x)$, but the first-order Riesz transforms have antisymmetric kernels. This is why a completely different set of dyadic singular operators was used in [6] to represent Riesz transforms. The operators from [6] are slightly complicated. We already explained that, in representing first-order Riesz transforms, one cannot average our simpler operators T_σ. But we do not know whether one can average something that is as simple as T_σ and also antisymmetric in order to obtain the first-order Riesz transforms.

ACKNOWLEDGMENT. We are grateful to the referee for several useful suggestions.

2. Motivation

Consider the standard differential operators

$$\partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right),$$

$$\bar\partial = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

The regularity of solutions of the Beltrami equation

$$\bar\partial f = \mu \cdot \partial f \tag{2}$$

has received a lot of attention from mathematicians since the 1940s. Function μ, called the *Beltrami coefficient*, belongs to the space $L^\infty(\mathbb{C})$. Its norm k is strictly smaller than 1. The result of Bojarski, Ahlfors, Bers, and Lavrentiev states that there is a W_1^2 solution to (2) that is a global homeomorphism of the extended complex plane $\hat{\mathbb{C}}$. For $z \in \mathbb{C}$, it maps the infinitely small ellipse, centered at z and with ratio of the axes $\frac{1+|\mu(z)|}{1-|\mu(z)|}$, into some infinitely small circle centered at $f(z)$. For this reason it is important to consider the constant

$$K = \frac{1+k}{1-k}.$$

Every homeomorphism $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ belonging to the Sobolev class W_1^2 for which (2) is fulfilled is called a *K-quasiconformal* mapping. Any local solution of (2) from $W_{1,\mathrm{loc}}^2$ is called a *quasiregular* mapping.

Denoting $g = \bar{\partial} f$, we are able to write f (such that $f(z) \sim z + C$, $|z| \to \infty$), as

$$f(z) = \frac{1}{\pi} \int \frac{g(\zeta)}{\zeta - z}\, dA(\zeta) + z + C.$$

This is where our operator T begins to play its role, for $\partial f = Tg + 1$ and so $g = \mu T g + \mu$, or

$$(I - \mu T) g = \mu.$$

Since

$$\|\mu T\|_{L^2 \to L^2} \le \|\mu\|_\infty \|T\|_{L^2 \to L^2} \le k < 1,$$

the operator $(I - \mu T)^{-1}$ exists and

$$g = (I - \mu T)^{-1} \mu.$$

It has been shown that the norm of μT as an operator on L^p is still less than 1 if p is slightly greater than 2. On the other hand, the word "slightly" is rather important. More precisely, it is known [4] that p should not exceed $1 + k^{-1}$. This fact, combined with our awareness that $\|T\|_{L^2 \to L^2} = 1$, gives rise to the assumption that

$$\|T\|_{L^p \to L^p} = p - 1.$$

For in that case,

$$\|\mu T\|_{L^p \to L^p} \le \|\mu\|_\infty \|T\|_{L^p \to L^p} < k(1 + k^{-1} - 1) = 1.$$

This is still an open question. The best known estimate has been obtained by Nazarov and Volberg in [5]. They proved that $\|T\|_{L^p \to L^p} \le 2(p - 1)$. This improves the previous estimate in [3] (namely, with $4(p-1)$). Recently, the estimate $2(p - 1)$ was obtained in [2] by using methods that differ from those in [5].

It is relevant that we also study weighted L^p spaces. For it was shown in [1] that

$$\|(I - \mu T)^{-1}\|_{L^p \to L^p} \le C(k) \|T\|_{L^p(w) \to L^p(w)},$$

where $w = |f_z \circ f^{-1}|^{p-2}$ for $p \in (1 + k, 1 + k^{-1})$ and f is a quasiconformal homeomorphism satisfying (2) together with the normalization $f(z) = z + o(1)$ as $|z| \to \infty$.

This estimate, together with the Corollary stated in Section 1, gives the linear growth of $\|(I - \mu T)^{-1}\|_{L^p \to L^p}$ where $p \to 1 + k$. We know (see [1]) this implies that weakly quasiregular maps on the plane are quasiregular. Hence this geometric fact becomes the corollary of our main theorem about representation of the Ahlfors–Beurling transform as a closure of the linear span of martingale transforms T_σ and the correct weighted estimates of T_σ obtained by Wittwer [9].

3. The Main Idea

As we have announced, we will represent our T as the result of averaging of operators similar to T_σ. After that, the desired estimate of $\|T\|_{L^2(w) \to L^2(w)}$ will follow from the two-dimensional version of (1) for $\|T_\sigma\|_{L^2(w) \to L^2(w)}$.

4. The Averaging

Instead of a dyadic lattice, let us consider for a moment a *grid* $\mathcal{G}$ of squares. This is a family of squares of the form $I \times J$, where I and J are dyadic intervals of unit length. Furthermore, for $t \in \mathbb{R}^2$ define $\mathcal{G}_t := \mathcal{G} + t$, that is, the grid of unit squares such that one (in fact, four) of them contain point t as one of its vertices.

Now we are ready to introduce our "core" operators $\mathbb{P}_t : L^2(w) \to L^2(w)$ by

$$\mathbb{P}_t f := \sum_{Q \in \mathcal{G}_t} [\langle f, h_Q^I \rangle h_Q^I - \langle f, h_Q^J \rangle h_Q^J].$$

Since $\langle f, h_Q^I \rangle h_Q^I$ can be written as

$$\tfrac{1}{2}[\langle f, h_Q^I \rangle h_Q^I + \langle f, h_Q^J \rangle h_Q^J + \langle f, h_Q \rangle h_Q] \\ + \tfrac{1}{2}[\langle f, h_Q^I \rangle h_Q^I - \langle f, h_Q^J \rangle h_Q^J - \langle f, h_Q \rangle h_Q]$$

and similarly for $\langle f, h_Q^J \rangle h_Q^J$, we see that $\mathbb{P}_t f$ is a linear combination (with coefficients 1 and -1) of two convex combinations of two martingale transforms (one of which is the identity). This means that the analogue of condition (1) also holds (with some other constant) for all operators $\mathbb{P}_t$. That we have translated our standard grid does not cause any problems, as we shall see later.

Notice that the family $\Omega := \{\mathcal{G}_t \mid t \in \mathbb{R}^2\}$ of all unit grids naturally corresponds to the torus $\mathbb{R}^2/\mathbb{Z}^2$, which is of course in one-to-one correspondence with the square $[0,1)^2$. Thus we are able to regard Ω as a probability space where the probability measure equals to the Lebesgue measure on $[0,1)^2$.

This now leads to the "mathematical expectation" of the "random variable" $\mathbb{P}$. This will again be an operator on $L^2(w)$, defined pointwise (for $f \in L^2(w)$) as

$$(\mathbb{E}\mathbb{P} f)(x) := \int_\Omega \mathbb{P}_t f(x)\, dm(t).$$

Since $\mathbb{E}\mathbb{P}$ is a result of integrating over a certain probability space (more generally, a set of finite measure), it makes sense to call this process the *averaging.* The significance of this operator is revealed in the following proposition.

Proposition 1. *With notation as before, the operator $\mathbb{E}\mathbb{P}$ is a convolution operator with kernel*

$$F(x_1, x_2) = \alpha(x_1)\beta(x_2) - \beta(x_1)\alpha(x_2) = \begin{vmatrix} \alpha(x_1) & \alpha(x_2) \\ \beta(x_1) & \beta(x_2) \end{vmatrix},$$

where

$$\alpha = h_0 * h_0 \quad \textit{and} \quad \beta = \chi_0 * \chi_0.$$

Here χ_0 and h_0 stand (respectively) for the characteristic and Haar function of the interval $[-1/2, 1/2]$.

Proof. Choose $t = (t_1, t_2) \in \mathbb{R}^2$ and $Q = I \times J \in \mathcal{G}_t$. Then

$$\begin{aligned} \langle f, h_Q^I \rangle &= \int_\mathbb{R} \int_\mathbb{R} f(s_1, s_2) h_Q^I(s_1, s_2)\, ds_1\, ds_2 \\ &= \int_J \int_I f(s_1, s_2) h_I(s_1) \chi_J(s_2)\, ds_1\, ds_2 \end{aligned}$$

and similarly for $\langle f, h_Q^J\rangle$. Thus for (fixed) $f \in L^2(\mathbb{R}^2)$ and $x = (x_1, x_2) \in \mathbb{R}^2$ we have

$$
\begin{aligned}
&(\mathbb{P}_t f)(x)\\
&\quad = \sum_{Q\in\mathcal{G}_t} [\langle f, h_Q^I\rangle h_Q^I - \langle f, h_Q^J\rangle h_Q^J](x)\\
&\quad = \sum_{Q\in\mathcal{G}_t} \bigg[\int_{\mathbb{R}}\int_{\mathbb{R}} f(s_1,s_2) h_I(s_1)\chi_J(s_2)\, ds_1\, ds_2 \cdot h_Q^I(x)\\
&\qquad\qquad - \int_{\mathbb{R}}\int_{\mathbb{R}} f(s_1,s_2)\chi_I(s_1) h_J(s_2)\, ds_1\, ds_2 \cdot h_Q^J(x)\bigg]\\
&\quad = \int_{\mathbb{R}}\int_{\mathbb{R}} f(s_1,s_2)\bigg(\sum_{Q\in\mathcal{G}_t}[h_I(s_1)\chi_J(s_2)h_Q^I(x) - \chi_I(s_1)h_J(s_2)h_Q^J(x)]\bigg)\, ds_1\, ds_2.
\end{aligned}
$$

The expression under the summation in the last row is nonzero for exactly one $Q \in \mathcal{G}_t$; namely, one such that $h_Q^I(x) \neq 0 \neq h_Q^J(x)$. This means that $x = (x_1, x_2) \in Q$ and hence $x_1 \in I$ and $x_2 \in J$. Thus $h_Q^I(x) = h_I(x_1)$ and $h_Q^J(x) = h_J(x_2)$. We thus obtain

$$
\begin{aligned}
&(\mathbb{P}_t f)(x)\\
&\quad = \int_{\mathbb{R}}\int_{\mathbb{R}} f(s_1,s_2)[h_I(s_1)\chi_J(s_2)h_I(x_1) - i\chi_I(s_1)h_J(s_2)h_J(x_2)]\, ds_1\, ds_2. \qquad (3)
\end{aligned}
$$

Because $\mathcal{G}_t$ does not change if we increase or decrease any component of t by 1, we may assume that $I = (t_1 - 1, t_1)$ and $J = (t_2 - 1, t_2)$. Denoting $I_0 = [-1/2, 1/2]$, this assumption implies

$$I = t_1 - \frac{1}{2} + I_0 \quad \text{and} \quad J = t_2 - \frac{1}{2} + I_0.$$

Now let χ_0 and h_0 be as in the formulation of the proposition and let $k_0 := -h_0$. The fact that I_0 is symmetric with respect to 0 yields the equalities

$$\chi_I(z) = \chi_0(t_1 - 1/2 - z)$$

and

$$h_I(z) = k_0(t_1 - 1/2 - z)$$

for all $z \in \mathbb{R}$. The analogue pair is valid also for J, of course.

The point here is that our goal was to modify the expressions on the left to look more like a part of a convolution integral with t_1 and t_2 as integration variables and z as a center of convolving.

Together with (3), the last two equalities imply

$$
\begin{aligned}
&(\mathbb{P}_t f)(x)\\
&\quad = \int_{\mathbb{R}}\int_{\mathbb{R}} f(s_1,s_2)\big[k_0\big(t_1 - \tfrac{1}{2} - s_1\big)\chi_0\big(t_2 - \tfrac{1}{2} - s_2\big)k_0\big(t_1 - \tfrac{1}{2} - x_1\big)\\
&\qquad\qquad - \chi_0\big(t_1 - \tfrac{1}{2} - s_1\big)k_0\big(t_2 - \tfrac{1}{2} - s_2\big)k_0\big(t_2 - \tfrac{1}{2} - x_2\big)\big]\, ds_1\, ds_2.
\end{aligned}
$$

Recall that $x_1 \in I = (t_1 - 1, t_1)$ and $x_2 \in J = (t_2 - 1, t_2)$. Hence $x_i < t_i < x_i + 1$ for $i = 1, 2$. Averaging in our case means integrating over all admissible t_i. Therefore,

$$(\mathbb{E}\mathbb{P}f)(x) = \int_{x_2}^{x_2+1} \int_{x_1}^{x_1+1} (\mathbb{P}_{(t_1,t_2)} f)(x)\, dt_1\, dt_2.$$

By using the most recent expression for $(\mathbb{P}_t f)(x)$ and changing variables (to $t_i - 1/2$) we obtain

$$(\mathbb{E}\mathbb{P}f)(x) = \int_{x_2-1/2}^{x_2+1/2} \int_{x_1-1/2}^{x_1+1/2} \int_{\mathbb{R}} \int_{\mathbb{R}} f(s_1, s_2)[A]\, ds_1\, ds_2\, dt_1\, dt_2,$$

where

$$A = k_0(t_1 - x_1)k_0(t_1 - s_1)\chi_0(t_2 - s_2) - k_0(t_2 - x_2)k_0(t_2 - s_2)\chi_0(t_1 - s_1).$$

Applying Fubini's theorem yields

$$(\mathbb{E}\mathbb{P}f)(x_1, x_2) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(s_1, s_2) \int_{x_2-1/2}^{x_2+1/2} \int_{x_1-1/2}^{x_1+1/2} [A]\, dt_1\, dt_2\, ds_1\, ds_2. \tag{4}$$

This is how we obtained the candidate for the convolution kernel F of the operator $\mathbb{E}\mathbb{P}$. Namely, equation (4) gives us the relation

$$F(x_1 - s_1, x_2 - s_2) = \int_{x_2-1/2}^{x_2+1/2} \int_{x_1-1/2}^{x_1+1/2} [A]\, dt_1\, dt_2.$$

Taking $s_1 = s_2 = 0$ gives

$$\begin{aligned} &F(x_1, x_2) \\ &= \int_{x_2-1/2}^{x_2+1/2} \int_{x_1-1/2}^{x_1+1/2} [k_0(t_1 - x_1)k_0(t_1)\chi_0(t_2) - k_0(t_2 - x_2)k_0(t_2)\chi_0(t_1)]\, dt_1\, dt_2. \end{aligned}$$

We are able to separate variables t_1 and t_2, so

$$\begin{aligned} F(x_1, x_2) = &\int_{x_1-1/2}^{x_1+1/2} k_0(t_1 - x_1)k_0(t_1)\, dt_1 \cdot \int_{x_2-1/2}^{x_2+1/2} \chi_0(t_2)\, dt_2 \\ &- \int_{x_1-1/2}^{x_1+1/2} \chi_0(t_1)\, dt_1 \cdot \int_{x_2-1/2}^{x_2+1/2} k_0(t_2 - x_2)k_0(t_2)\, dt_2. \end{aligned}$$

Observing that

$$\int_{x_i-1/2}^{x_i+1/2} \chi_0(t_i)\, dt_i = \int_{x_i-1/2}^{x_i+1/2} \chi_0(t_i)\chi_0(x_i - t_i)\, dt_i$$

for $i = 1, 2$ and that $k_0 * k_0 = h_0 * h_0$, we finally obtain

$$F(x_1, x_2) = (h_0 * h_0)(x_1)(\chi_0 * \chi_0)(x_2) - (\chi_0 * \chi_0)(x_1)(h_0 * h_0)(x_2),$$

as desired. $\square$

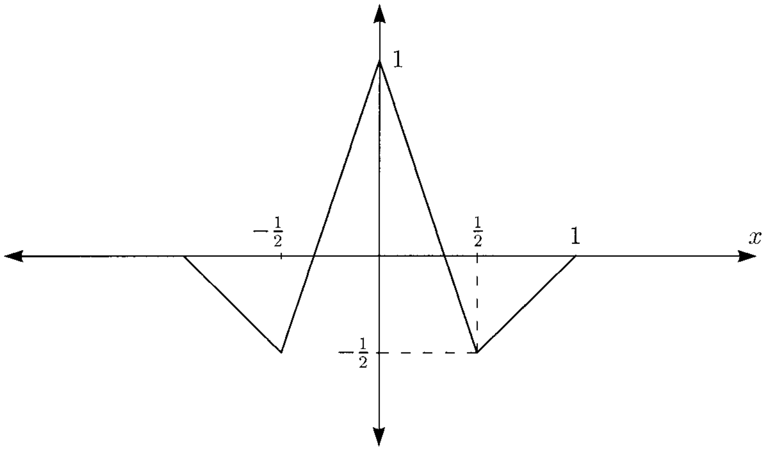

Figure 1 Graph of α

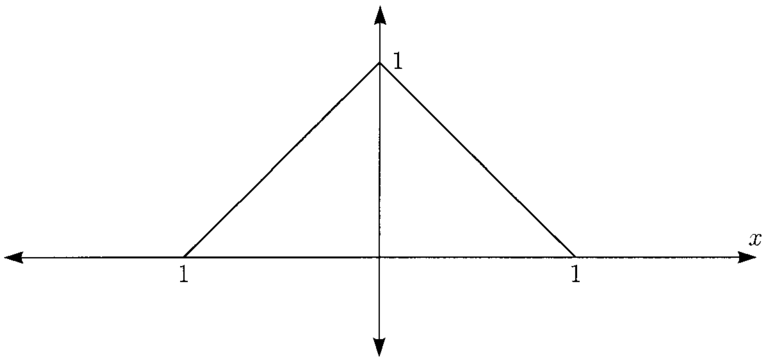

Figure 2 Graph of β

Graphs of functions α and β are shown as Figures 1 and 2, respectively.

Notice, as a corollary, that (1) also holds for the operator $\mathbb{E}\mathbb{P}$ in place of T_σ, since it held for all $\mathbb{P}_t$.

Instead of the unit grid we may consider a grid of squares with sides of an arbitrary length $\rho > 0$. Denote such a grid by $\mathcal{G}_t^\rho$ if $t \in \mathbb{R}^2$ is a vertex of one of its members. Henceforth we will call ρ the *size* of the grid and t its *reference point*. We obtain another family of operators, defined by

$$\mathbb{P}_t^\rho f := \sum_{Q \in \mathcal{G}_t^\rho} [\langle f, h_Q^I \rangle h_Q^I - \langle f, h_Q^J \rangle h_Q^J].$$

REMARK 1. In order to clarify some basic properties of these operators, we present the following observations (and omit the easy proofs).

For $t \in \mathbb{R}^2$, $\rho > 0$, and any function f on $\mathbb{R}^2$, define $Sf(x) = f(\rho x + t)$. If w is any weight, then S maps $L^2(w) \to L^2(Sw)$. We also have the identity

$$\mathbb{P}_t^\rho = S^{-1}\mathbb{P}_0 S.$$

More precisely, $\mathbb{P}_t^\rho|_{L^2(w)}$ is the composition of operators

$$L^2(w) \xrightarrow{S} L^2(Sw) \xrightarrow{\mathbb{P}_0} L^2(Sw) \xrightarrow{S^{-1}} L^2(w).$$

Since $\|Sf\|_{L^2(Sw)} = (1/\sqrt{\rho})\|f\|_{L^2(w)}$, it follows that $\|S\|_{L^2(w)\to L^2(Sw)} = 1/\sqrt{\rho}$; similarly $\|S^{-1}\|_{L^2(Sw)\to L^2(w)} = \sqrt{\rho}$.

If $w \in A_2$ then also $Sw \in A_2$ and $Q_{Sw,2} = Q_{w,2}$. Therefore, $\mathbb{P}_0\colon L^2(Sw) \to L^2(Sw)$ is a bounded operator that inherits the estimate for its norm from T_σ. That is, it satisfies the same inequality as T_σ does in (1), according to [9].

These facts combined tell us that every $\mathbb{P}_t^\rho$ is a bounded operator on $L^2(w)$ with $\|\mathbb{P}_t^\rho\|_{L^2(w)\to L^2(w)} \le CQ_{w,2}$, where C is an absolute constant, as usual.

Let us use the same procedure of averaging as that used earlier for $\mathbb{P}_t$, but now for $\mathbb{P}_t^\rho$; this yields $\mathbb{E}\mathbb{P}^\rho$. Applying a similar proof as for Proposition 1, we can show the following.

Proposition 2. *Choose $\rho > 0$. Then averaging operators $\mathbb{P}_t^\rho$ over the set $\Omega^\rho := \mathbb{R}^2/(\rho\mathbb{Z}^2)$ returns a convolution operator with the kernel*

$$F^\rho(x_1, x_2) := \frac{1}{\rho^2} F\left(\frac{x_1}{\rho}, \frac{x_2}{\rho}\right).$$

Here the set Ω^ρ is endowed with the normalized Lebesgue measure, $(1/\rho^2)\, dm_2$.

Thus we have found the kernel of the operator as a result of averaging over *all* grids of a *fixed* size. Our next step will be to average over all *sizes.* Let us explain what we mean by that.

Take $r > 0$. A *lattice of caliber r* is said to be a family of intervals (squares) obtained from the standard dyadic lattice $\mathcal{L}$ by dilating it by a factor r and translating by an arbitrary vector t. In other words, such a lattice (call it $\mathcal{L}_t^r$) is the union of *grids* of sizes $r \cdot 2^n$, $n \in \mathbb{Z}$, having t as their reference point. It is clear that the set of all possible calibers naturally corresponds to the interval $[1, 2)$. For our purpose, the most appropriate measure on this interval turns out to be dr/r. This makes all other possible choices of intervals (e.g., $[2^n, 2^{n+1})$) have the same measure (i.e., $\log 2$).

We introduce kernels

$$k^r := \sum_{n=-\infty}^{\infty} F^{r\cdot 2^n}.$$

This sum is well-defined because, as a function, it is equal (by our previous assertion) to

$$\sum_{n=-\infty}^{\infty} \frac{1}{r^2 \cdot 2^{2n}} F\left(\frac{\cdot}{r \cdot 2^n}\right) = \frac{1}{r^2} \sum_{n=-\infty}^{\infty} \frac{1}{4^n} F\left(\frac{\cdot}{r \cdot 2^n}\right).$$

Since the function F is bounded and of compact support, the series converges absolutely on $\mathbb{R}^2$ and uniformly outside any neighbourhood of the origin. It is easy to see that the series converges in the sense of distributions to k^r when understood as a distribution in the following sense. Let ϕ be a test function from the Schwartz class $\mathcal{S}$ and let $x = (x_1, x_2)$; then

$$(\phi, k^r) = \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \phi(x) k^r(x)\, dx.$$

We would like to show that k^r defines a bounded convolution operator—more precisely, that it is a strong limit of its partial sums.

That k^r* is a sum of operators obtained by averaging over grids of size $r \cdot 2^n$ hints at k^r* itself being an average, this time over *unions* of these grids (i.e., lattices of caliber r). Here we present what exactly we have in mind by that.

For $M \in \mathbb{Z}$, let the Mth partial sum of the series k^r be

$$k_M^r := \sum_{n=-\infty}^{M} F^{r\cdot 2^n}.$$

LEMMA 1. *Function k_M^r defines a bounded convolution operator on $L^2(w)$. The limit $k^r* := \lim_{M\to\infty} k_M^r*$ exists in the strong sense and also gives rise to a bounded operator on $L^2(w)$.*

Proof. For the sake of simplicity we will assume that $r = 1$; the proof does not change at all for general r. We start with a formal definition:

$$k_M^1 * f = \left(\sum_{n=-\infty}^{M} F^{2^n} \right) * f = \sum_{n=-\infty}^{M} \frac{1}{4^n} \int_{[0,2^n]^2} \mathbb{P}_t^{2^n} f\, dt \tag{5}$$

by Proposition 2. At this point we will need the following observation: *For any $n, M \in \mathbb{Z}$ with $n \le M$, we have*

$$\frac{1}{4^n} \int_{[0,2^n]^2} \mathbb{P}_t^{2^n} f\, dt = \frac{1}{4^M} \int_{[0,2^M]^2} \mathbb{P}_t^{2^n} f\, dt.$$

This time we simplify the proof by taking $M = 0$. Then the square $[0,1]^2$ consists of exactly 4^{-n} dyadic squares of size 2^n. The integral $\int \mathbb{P}_t^{2^n} f\, dt$ over each of them equals the integral over $[0, 2^n]^2$ owing to the invariance of the measure on Ω^{2^n} with respect to the map on $\Omega^{2^n} \equiv \mathbb{R}^2/2^n\mathbb{Z}^2$ that is induced by the shift on $\mathbb{R}^2$. Since the sum of integrals over these squares equals the integral over their union, which is $[0,1]^2$, we have proved this part of the statement.

This enables us to rewrite (5) as

$$\sum_{n=-\infty}^{M} \frac{1}{4^M} \int_{[0,2^M]^2} \mathbb{P}_t^{2^n} f\, dt,$$

which is equal to

$$\frac{1}{4^M} \int_{[0,2^M]^2} \sum_{n=-\infty}^{M} \mathbb{P}_t^{2^n} f\, dt.$$

Let us explain why we were able to exchange the order of integration and summation in the last step. Again we do this for $M = 0$ only.

We need to show that the $L^2(w)$ norm of the difference

$$\int_{[0,1]^2} \sum_{n=-\infty}^{0} \mathbb{P}_t^{2^n} f\, dt - \sum_{n=-N}^{0} \int_{[0,1]^2} \mathbb{P}_t^{2^n} f\, dt = \int_{[0,1]^2} \sum_{n<-N} \mathbb{P}_t^{2^n} f\, dt$$

is small if $N \in \mathbb{N}$ is large.

Lemma 2. *For every function $f \in L^2(w)$ with $w \in A_2$ and for every $t \in \mathbb{R}^2$,*

$$\lim_{N\to\infty} \Big\| \sum_{|n|\geq N} \mathbb{P}_t^{2^n} f \Big\|_{L^2(w)} = 0.$$

Proof. A collection $\{e_n \mid n \in \mathbb{Z}\}$ of vectors in a Hilbert space $\mathcal{H}$ is said to give rise to a *Riesz basis* of $\mathcal{H}$ if there are $M, m > 0$ such that, for any sequence $(\lambda_n)_{n\in\mathbb{Z}}$ of scalars, we have the double-sided inequality

$$m \sum_{n\in\mathbb{Z}} \|\lambda_n e_n\|^2 \leq \Big\| \sum_{n\in\mathbb{Z}} \lambda_n e_n \Big\|^2 \leq M \sum_{n\in\mathbb{Z}} \|\lambda_n e_n\|^2.$$

It is a known fact that

$$w \in A_2 \iff \text{Haar functions form a Riesz basis in } L^2(w). \tag{$*$}$$

For every $N \in \mathbb{N}$, define a measurable function $g_N : \mathbb{R}^2 \to \mathbb{C}$ by

$$g_N(t) = \Big\| \sum_{|n|\geq N} \mathbb{P}_t^{2^n} f \Big\|_{L^2(w)}.$$

We claim that, for any $t \in \mathbb{R}^2$,

$$\lim_{N\to\infty} g_N(t) = 0. \tag{6}$$

In order to prove this statement, choose $t \in \mathbb{R}^2$ and consider the translation operator S_t. Then

$$\sum_{|n|\geq N} \mathbb{P}_t^{2^n} = S_{-t} \sum_{|n|\geq N} \mathbb{P}_0^{2^n} S_t,$$

as we saw in Remark 1. Thus

$$g_N(t) = \Big\| \sum_{|n|\geq N} \mathbb{P}_0^{2^n}(S_t f) \Big\|_{L^2(S_t w)} \leq \Big\| \sum_{|n|\geq N} \mathbb{P}_0^{2^n} \Big\|_{B(L^2(S_t w))} \|S_t f\|_{L^2(S_t w)}.$$

We can estimate the first term on the right by using (1) to obtain that it is less than or equal to $CQ_{S_t w,2}$. Referring once more to Remark 1, we get $Q_{S_t w,2} = Q_{w,2}$ and so conclude that

$$g_N(t) \leq CQ_{w,2} \|f\|_{L^2(w)} \tag{7}$$

for all $t \in \mathbb{R}^2$ and all $N \in \mathbb{N}$. Taking $N = 0$, the statement $(*)$ implies that

$$\sum_{n\in\mathbb{Z}}\sum_{Q\in\mathcal{G}_t^{2^n}}(\|\langle f,h_Q^I\rangle h_Q^I\|_{L^2(w)}^2+\|\langle f,h_Q^J\rangle h_Q^J\|_{L^2(w)}^2)<\infty,$$

so

$$\lim_{N\to\infty}\sum_{|n|\geq N}\sum_{Q\in\mathcal{G}_t^{2^n}}(\|\langle f,h_Q^I\rangle h_Q^I\|_{L^2(w)}^2+\|\langle f,h_Q^J\rangle h_Q^J\|_{L^2(w)}^2)=0.$$

Again by $(*)$ it follows that

$$\lim_{N\to\infty}\Big\|\sum_{|n|\geq N}\mathbb{P}_t^{2^n}f\Big\|_{L^2(w)}=0,$$

which is exactly the statement in (6). This proves Lemma 2. □

By (7), the functions g_N are all bounded; hence the dominated convergence theorem implies that

$$\lim_{N\to\infty}\int_{[0,1]^2}g_N(t)\,dt=\int_{[0,1]^2}\lim_{N\to\infty}g_N(t)\,dt.$$

But now equation (6) yields that

$$\lim_{N\to\infty}\int_{[0,1]^2}\Big\|\sum_{|n|\geq N}\mathbb{P}_t^{2^n}f\Big\|_{L^2(w)}dt=0.$$

We can do the same for $\sum_{n<-N}$ instead of $\sum_{|n|\geq N}$, which is how we justify reversing the order of integration and summation (see p. 425). So we were indeed right to claim that

$$k_M^1*f=\frac{1}{4^M}\int_{[0,2^M]^2}\sum_{n=-\infty}^{M}\mathbb{P}_t^{2^n}f\,dt.$$

The integrands all satisfy (a two-dimensional version of) the inequality (1), so we may conclude that k_M^1* does, too. This proves the first part of Lemma 1.

Now take $f\in L^2(w)$. Let us verify that $\{k_M^1*f\mid M\in\mathbb{N}\}$ form a Cauchy sequence in $L^2(w)$. Choose $M,N\in\mathbb{N}$ with $N\leq M$ and compute the difference

$$\begin{aligned}(k_M^1-k_N^1)*f&=\frac{1}{4^M}\int_{[0,2^M]^2}\sum_{n=-\infty}^{M}\mathbb{P}_t^{2^n}f\,dt-\frac{1}{4^N}\int_{[0,2^N]^2}\sum_{n=-\infty}^{N}\mathbb{P}_t^{2^n}f\,dt\\&=\frac{1}{4^M}\int_{[0,2^M]^2}\sum_{n=-\infty}^{M}\mathbb{P}_t^{2^n}f\,dt-\frac{1}{4^M}\int_{[0,2^M]^2}\sum_{n=-\infty}^{N}\mathbb{P}_t^{2^n}f\,dt\\&=\frac{1}{4^M}\int_{[0,2^M]^2}\sum_{n=N+1}^{M}\mathbb{P}_t^{2^n}f\,dt.\end{aligned}$$

This difference is small if M and N are sufficiently large, which basically follows from Lemma 2 as well. On the other hand, by choosing f from the Schwartz class $\mathcal{S}$ we can use the fact that k_M^1 converges to k^1 in the sense of distributions, and hence we may conclude that $k^1_M * f \to k^1 * f$ pointwise. Here $k^1 * f$ is understood as the convolution of a distribution and a test function. In particular, for $f \in \mathcal{S}$ we have

$$k^1 * f = \lim_{M\to\infty} \frac{1}{4^M} \int_{[0,2^M]^2} \sum_{n=-\infty}^{M} \mathbb{P}_t^{2^n} f \, dt \tag{8}$$

and similarly for all other $r \in [1, 2)$. Reasoning in the same way as before, we see that (1) is fulfilled with k^1 in place of T_σ.

This finishes the proof of Lemma 1. □

REMARK 2. To understand distribution k^1 better, one can notice that its Fourier transform is a bounded function. In fact, using the previous reasoning with $w = 1$, we conclude that $f \mapsto k^1 * f$ is a bounded operator on $L^2(\mathbb{R}^2)$.

Note that the integrand in (8) is the sum

$$\sum_Q [\langle f, h^I_Q\rangle h^I_Q - \langle f, h^J_Q\rangle h^J_Q],$$

where Q runs over a "truncated lattice"—that is, the union of grids of sizes 2^n $(-\infty < n \le M)$ and with reference points in t. Since the square $[0, 2^M]^2$ represents all possible reference points for such unions, the expression under the limit sign in (8) means exactly averaging over these unions. But when $M \to \infty$, the "truncated" lattice becomes "complete". Thus we may understand the limit (8)—and, more generally, the convolution with k^r—as the result of averaging over all lattices of caliber r. Since our operators over lattices are bounded in the sense of (1), the same is true of k^r*.

Averaging operators k^r*—that is, integrating k^r with respect to dr/r (again in the strong sense, i.e., on any fixed test function f)—gives us a convolution operator once again. Call its kernel k. Then

$$\begin{aligned} k(x) &= \int_1^2 k^r(x)\frac{dr}{r} = \int_1^2 \sum_{n=-\infty}^{\infty} F^{r\cdot 2^n}(x)\frac{dr}{r} = \sum_{n=-\infty}^{\infty} \int_1^2 F^{r\cdot 2^n}(x)\frac{dr}{r} \\ &= \sum_{n=-\infty}^{\infty} \int_{2^n}^{2^{n+1}} F^s(x)\frac{ds}{s} = \int_0^\infty F^s(x)\frac{ds}{s}. \end{aligned} \tag{9}$$

Since the integral $\int_1^2 dr/r$ is finite and the estimate (1) holds for all k^r*, it also holds for $k*$. Thus we have been able to represent the operator $k*$ as a result of averaging our "brick" operators $\mathbb{P}_t^\rho$ over lattices of all calibers.

Observe that there is a map $m\colon S^1 \to \mathbb{C}$ such that

$$k(x) = \frac{m(x/|x|)}{|x|^2} \tag{10}$$

for all $x \in \mathbb{R}^2$. Taking unimodular vectors yields $m = k|_{S^1}$. For existence of such m it suffices to show that the function $|x|^2 k(x)$ depends only on the direction of x. So take $x = (x_1, x_2) \in \mathbb{R}^2$ and $\lambda > 0$. We use Proposition 2 to compute

$$\begin{aligned}k(\lambda x) &= \int_0^\infty F^\rho(\lambda x)\frac{d\rho}{\rho} = \int_0^\infty F\left(\frac{\lambda x}{\rho}\right)\frac{d\rho}{\rho^3} \\ &= \int_0^\infty \left[\alpha\left(\frac{\lambda x_1}{\rho}\right)\beta\left(\frac{\lambda x_2}{\rho}\right) - \beta\left(\frac{\lambda x_1}{\rho}\right)\alpha\left(\frac{\lambda x_2}{\rho}\right)\right]\frac{d\rho}{\rho^3}.\end{aligned}$$

Clearly, we must now introduce the variable $\mu := \rho/\lambda$. We obtain

$$k(\lambda x) = \frac{1}{\lambda^2}\int_0^\infty \left[\alpha\left(\frac{x_1}{\mu}\right)\beta\left(\frac{x_2}{\mu}\right) - \beta\left(\frac{x_1}{\mu}\right)\alpha\left(\frac{x_2}{\mu}\right)\right]\frac{d\mu}{\mu^3} = \frac{1}{\lambda^2}k(x),$$

which is what we wanted.

REMARK 3. Observe that k is an even function, because so are α and β. Therefore, m is even on S^1 as well.

We are able to compute function m. Namely, the equality just displayed implies that $m(e^{i\varphi}) = M(\cos\varphi, \sin\varphi) - M(\sin\varphi, \cos\varphi) = M(e^{i\varphi}) - M(e^{i(\pi/2-\varphi)})$. Computation shows that

$$M(\cos\varphi, \sin\varphi) = \frac{1}{\cos^2\varphi}\Phi(\cot\varphi),$$

where

$$\Phi(a) = \begin{cases} a^2/6 - a^3/4 & \text{if } 0 < a \le 1/2, \\ a^3/12 - a^2/6 - 1/(48a) + 1/12 & \text{if } 1/2 \le a \le 1, \\ 1/(16a) - 1/12 & \text{if } a \ge 1. \end{cases}$$

For $a < 0$, the function Φ is defined by the requirement that it be even. Finally, let $\Psi(a) = (1 + a^{-2})\Phi(a)$. We have thus acquired the formula

$$m(e^{i\varphi}) = \Psi(\cot\varphi) - \Psi(\tan\varphi).$$

Clearly, function m is continuous on the sphere S^1 and so k is continuous on $\mathbb{R}^2\backslash\{0\}$.

Recall that we are aiming at the operator T, which is the convolution with $1/z^2$. Its kernel can be written in polar coordinates as $e^{-2i\varphi}/r^2$. Comparing this to equation (10), where $k(re^{i\varphi}) = m(e^{i\varphi})/r^2$, we suspect that it would be useful to find a suitable way of transforming $m(e^{i\varphi})$ into the function $e^{-2i\varphi}$.

Denoting $h(\zeta) = \zeta^{-2}$ for $\zeta \in S^1$, we have

$$\begin{aligned}(m * h)(e^{i\varphi}) &= \int_{-\pi}^{\pi} m(e^{i\psi})e^{-2i(\varphi-\psi)}\,d\psi \\ &= e^{-2i\varphi}\int_{-\pi}^{\pi} m(e^{i\psi})e^{2i\psi}\,d\psi \\ &= h(e^{i\varphi})\int_{-\pi}^{\pi} m(e^{i\psi})[\cos 2\psi + i\sin 2\psi]\,d\psi.\end{aligned}$$

Because m is an even function on $[-\pi, \pi]$ (since α and β are even on $\mathbb{R}$), the integral $\int_{-\pi}^{\pi} m(e^{i\psi})\sin 2\psi\,d\psi$ is equal to zero. Our expression thus simplifies to

$$h(e^{i\varphi})\int_{-\pi}^{\pi} m(e^{i\psi})\cos 2\psi\,d\psi.$$

Hence

$$h = \frac{m * h}{c}, \tag{11}$$

where

$$c = \int_{-\pi}^{\pi} m(e^{i\psi}) \cos 2\psi \, d\psi.$$

Notice that (11) gives essentially the representation of the kernel $t(re^{i\phi})$ of the Ahlfors–Beurling operator as a linear combination of kernels $r^{-2}[\Psi(\cot(\theta-\varphi)) - \Psi(\tan(\theta-\varphi))]$ with coefficients $e^{-2i\varphi}/c$, where Ψ is defined (see Remark 3) by a slightly awkward piecewise expression.

5. The Main Calculation

We come to the main point of our work. It needs to be verified that $c \neq 0$. If this is so (and we will see that it is), then (11) represents our singular operator T as the "average" of martingale transforms T_σ (actually their analogues, built with the help of the $\mathbb{P}_t^\rho$).

Recall from (9) and (10) that

$$m(e^{i\psi}) = \int_0^\infty F^\rho(e^{i\psi}) \frac{d\rho}{\rho} = \int_0^\infty F\left(\frac{e^{i\psi}}{\rho}\right) \frac{d\rho}{\rho^3}.$$

Hence

$$c = \int_0^\infty \int_0^{2\pi} F\left(\frac{e^{i\psi}}{\rho}\right) \frac{\cos 2\psi}{\rho^3} \, d\psi \, d\rho.$$

Taking $r = 1/\rho$ yields

$$c = \int_0^\infty \int_0^{2\pi} F(re^{i\psi}) r \cos 2\psi \, d\psi \, dr.$$

In Cartesian coordinates, this integral reads

$$c = \int_{\mathbb{R}} \int_{\mathbb{R}} F(x, y) \frac{x^2 - y^2}{x^2 + y^2} \, dx \, dy.$$

In Proposition 1 we saw that

$$F(x, y) = \alpha(x)\beta(y) - \beta(x)\alpha(y),$$

which leads to expressing our integral as

$$c = 2 \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(x)\beta(y) \frac{x^2 - y^2}{x^2 + y^2} \, dx \, dy.$$

Since α and β are even functions supported on the interval $[-1, 1]$, we obtain

$$c = 8 \int_0^1 \int_0^1 \alpha(x)\beta(y) \frac{x^2 - y^2}{x^2 + y^2} \, dx \, dy.$$

We shall evaluate c by computing the inner integral first.

Figure 2 (see p. 422) shows that $\beta(y) = 1 - y$ on interval $[0, 1]$. We may combine this with the identity

$$\frac{x^2 - y^2}{x^2 + y^2} = 1 - \frac{2y^2}{x^2 + y^2}$$

and the observation (which follows from Figure 1) that $\int_0^1 \alpha(x)\,dx = 0$ to arrive at

$$c = 16 \int_0^1 y^2(y-1) \int_0^1 \frac{\alpha(x)}{x^2 + y^2}\,dx\,dy.$$

For $y > 0$, computation returns

$$\begin{aligned} C(y) &:= \int_0^1 \frac{\alpha(x)}{x^2 + y^2}\,dx \\ &= \frac{1}{y}\left(2 \arctan \frac{1}{2y} - \arctan \frac{1}{y}\right) \\ &\quad + \left(\frac{1}{2}\log(y^2 + 1) + 3 \log y + 4 \log 2 - 2 \log(4y^2 + 1)\right). \end{aligned}$$

We now need only evaluate the integral

$$\int_0^1 y^2(y-1) C(y)\,dy.$$

We can directly calculate this integral to find that it equals

$$\frac{1}{12}\left(\arctan 2 - 4 \arctan 0.5 + \frac{15}{8} \log 5 - 4 \log 2\right),$$

which is approximately -0.042.

This fact enables us to state our main result as follows.

THEOREM 1. *For any A_2 weight w, the operator T of convolution with kernel z^{-2} satisfies the boundedness condition*

$$\|T\|_{L^2(w) \to L^2(w)} \leq C Q_{w,2},$$

where the constant C does not depend on the weight.

Proof. For $\psi \in [-\pi, \pi]$, let $U_\psi : \mathbb{C} \to \mathbb{C}$ be defined by $U_\psi(\zeta) := \zeta e^{-i\psi}$ and denote $k_\psi := k \circ U_\psi$. If we denote by S_ψ the mapping $S_\psi f = f \circ U_{-\psi}$ and if K_ψ stands for the convolution operator with kernel k_ψ, then we can easily verify the similarity relation

$$K_\psi = S_\psi^{-1} K_0 S_\psi.$$

Since the operator K_0 of convolution with k is bounded in $L^2(S_\psi w)$, as we saw on page 427, and since each S_ψ is an isometry $L^2(w) \to L^2(S_\psi w)$, it follows that the operators K_ψ again belong to $B(L^2(w))$, whose norms can be uniformly estimated by the norm of K_0.

Choose $n \in \mathbb{N}$ and let $-\pi/2 = \psi_0 < \psi_1 < \cdots < \psi_n = \pi/2$ be a subdivision of the interval $[-\pi/2, \pi/2]$ such that $\Delta\psi_j := \psi_j - \psi_{j-1} < 2\pi/n$ for $j = 1, \ldots, n$. Put

$$T_n' := \frac{1}{c}\sum_{j=1}^{n} e^{-2i\psi_j}\Delta\psi_j K_{\psi_j}.$$

This definition, applied to all $n \in \mathbb{N}$, determines a *bounded* family of operators in $B(L^2(w))$. Hence there is a subsequence that converges weakly to some operator, call it T'.

Let f, g be two smooth functions whose supports are *disjoint* compact sets. We would like to show that $\langle Tf, g\rangle_w = \langle T'f, g\rangle_w$.

First we make the following computation:

$$\begin{aligned}\frac{1}{c}\int_{\mathbb{R}^2} f(x-s)\int_{-\pi}^{\pi} e^{-2i\psi}k_\psi(s)\,d\psi\,dm_2(s)\\ = \frac{1}{c}\int_{\mathbb{R}^2} f(x-s)\frac{1}{|s|^2}\int_{-\pi}^{\pi} e^{-2i\psi}m\left(\frac{s}{|s|}e^{-i\psi}\right)d\psi\,dm_2(s).\end{aligned}$$

From (11) we see that this is equal to

$$\frac{1}{c}\int_{\mathbb{R}^2} f(x-s)\frac{1}{|s|^2}\cdot c\cdot h\left(\frac{s}{|s|}\right)ds,$$

so in fact we have

$$\int_{\mathbb{R}^2} f(x-s)\frac{1}{z(s)^2}\,ds = \left(f * \frac{1}{z^2}\right)(x) = (Tf)(x).$$

After the change of variable we get that

$$(Tf)(x) = \frac{1}{c}\int_{\mathbb{R}^2} f(s)\int_{-\pi}^{\pi} e^{-2i\psi}k_\psi(x-s)\,d\psi\,dm_2(s).$$

Denoting by $\mathcal{F}$ and $\mathcal{G}$ the supports of f and g, respectively, this equality yields

$$\langle Tf, g\rangle_w = \int_{\mathcal{G}}\frac{1}{c}\int_{\mathcal{F}} f(s)\int_{-\pi}^{\pi} e^{-2i\psi}k_\psi(x-s)\,d\psi\,ds\,\overline{g(x)}w(x)\,dx.$$

On the other hand,

$$\langle T_n'f, g\rangle_w = \int_{\mathcal{G}}\frac{1}{c}\int_{\mathcal{F}} f(s)\sum_{j=1}^{n} e^{-2i\psi_j}\Delta\psi_j k_{\psi_j}(x-s)\,ds\,\overline{g(x)}w(x)\,dx.$$

Thus

$$\begin{aligned}&\langle Tf, g\rangle_w - \langle T_n'f, g\rangle_w\\ &\quad= \int_{\mathcal{G}}\frac{1}{c}\int_{\mathcal{F}} f(s)\Bigg[\int_{-\pi}^{\pi} e^{-2i\psi}k_\psi(x-s)\,d\psi\\ &\qquad\qquad - \sum_{j=1}^{n} e^{-2i\psi_j}\Delta\psi_j k_{\psi_j}(x-s)\Bigg]ds\,\overline{g(x)}w(x)\,dx\end{aligned}$$

and hence, for every $\varepsilon > 0$, we have

$$|\langle Tf, g\rangle_w - \langle T_n'f, g\rangle_w| \le \varepsilon\frac{1}{|c|}\int_{\mathcal{F}}|f(s)|\,ds\int_{\mathcal{G}}|g(x)|w(x)\,dx$$

if $n \in \mathbb{N}$ is sufficiently large. It follows that

$$\begin{aligned}|\langle Tf, g\rangle_w - \langle T_n' f, g\rangle_w| \\ \le \frac{\varepsilon}{|c|}\|f\|_{L^2(w)}\|g\|_{L^2(w)}\left(\int_{\mathcal{F}} w(s)^{-1}\,ds\right)^{1/2}\left(\int_{\mathcal{G}} w(x)\,dx\right)^{1/2}.\end{aligned}$$

Since $\mathcal{F}$ and $\mathcal{G}$ are compact sets, there is a square E such that $\mathcal{F} \cup \mathcal{G} \subset E$. This enables us to estimate

$$\begin{aligned}\left(\int_{\mathcal{F}} w(s)^{-1}\,ds\right)^{1/2}\left(\int_{\mathcal{G}} w(x)\,dx\right)^{1/2} &\le \left(\int_E w(s)^{-1}\,ds \cdot \int_E w(x)\,dx\right)^{1/2} \\ &= (|E|\langle w^{-1}\rangle_E |E|\langle w\rangle_E)^{1/2} \le |E| Q_{w,2}^{1/2}.\end{aligned}$$

We have proved that, for every $\varepsilon > 0$, there is an $n_0 \in \mathbb{N}$ such that $n \ge n_0$ implies

$$|\langle Tf, g\rangle_w - \langle T_n' f, g\rangle_w| \le \varepsilon \frac{|E|}{|c|}\|f\|_{L^2(w)}\|g\|_{L^2(w)} Q_{w,2}^{1/2},$$

which of course means that $\langle Tf, g\rangle_w = \langle T'f, g\rangle_w$ for any $f, g \in C^\infty$ with disjoint compact supports.

From here it is easy to see that $T - T' = M_\omega$, the multiplication operator by some $\omega \in L^\infty(\mathbb{R}^2)$. Note that our ω does not depend on w. Finally,

$$\|T\|_{L^2(w)\to L^2(w)} \le \|T'\|_{L^2(w)\to L^2(w)} + \|M_\omega\|_{L^2(w)\to L^2(w)}.$$

Operator T' is a weak limit of operators whose norms are uniformly bounded by $C \cdot Q_{w,2}$. Also,

$$\|M_\omega\|_{L^2(w)\to L^2(w)} \le \|\omega\|_\infty \le \|\omega\|_\infty Q_{w,2}$$

because, by Hölder's inequality, $Q_{w,2} \ge 1$ for all weights w. We conclude that there is a constant $C' > 0$ such that

$$\|T\|_{L^2(w)\to L^2(w)} \le C' \cdot Q_{w,2}$$

for all $w \in A_2$.

We have thus proved our main estimate. To finish the proof of Theorem 1, it is enough to show that $\omega = 0$. Fix a compact set K and let R be a large number. Consider ψ_R smooth with compact support, which is 1 on the disc of radius R centered at the origin, 0 outside of the disc of radius $R+1$, and has a bounded gradient (independent of R). It is easy to see that $T\psi_R(x) \to 0$ as R tends to infinity, and it is easy to see that $T_n'\psi_R(x) \to 0$ uniformly in n and $x \in K$. In particular, $\|T_n'\psi_R\|_{L^\infty} \le \varepsilon(R)$ and $\|T\psi_R\|_{L^\infty} \le \varepsilon(R)$. Fixing a measurable subset E of K and using the fact that $(T - T')\psi_R = \omega\psi_R = \omega$ on K, we can write

$$\left|\int \omega\chi_E\,dx\right| = |\langle (T - T')\psi_R, \chi_E\rangle| \le 2\varepsilon(R)|E|,$$

which tends to zero when R grows to infinity. Hence $\int_E \omega\,dx = 0$ for any measurable subset E of K. We conclude that $M_\omega = 0$. So $T = T'$, and Theorem 1 is completely proved. □

6. Sharpness

We still need to show that this estimate is sharp in the same sense that the estimate (1) was. For this purpose we shall need the following auxiliary result, which can be proved by direct simple calculation.

LEMMA 3. *Let* $|\alpha| < 2$ *and define* $w\colon \mathbb{R}^2 \to [0,\infty)$ *by* $w(x) = |x|^\alpha$. *Then there exist constants* $M, m > 0$, *not depending on* α, *such that*

$$\frac{m}{4-\alpha^2} \le Q_{w,2} \le \frac{M}{4-\alpha^2}.$$

Now we are ready to prove the sharpness. Calculations of this kind have already been made for other singular integral operators but not for T, so we include this calculation for the sake of completeness.

PROPOSITION 3. *Let* $\phi\colon \mathbb{R} \to (0,\infty)$ *grow more slowly than a linear function. Then there is a weight* $w \in A_2$ *and a function* $f \in L^2(w)$ *such that*

$$\|Tf\|_{L^2(w)} > \phi(Q_{w,2})\|f\|_{L^2(w)}. \tag{12}$$

Proof. For $|\alpha| < 2$, define $w(z) = |z|^\alpha$. (The restriction on α is needed if w is to satisfy the A_2 condition.) Furthermore, let $E = \{(r,\varphi) \mid 0 < r < 1,\ 0 < \varphi < \pi/2\}$, let $X = -E$, and let $f(z) = |z|^{-\alpha}\chi_E$.

We shall estimate the left and right sides (actually, their squares) of inequality (12). Thus we begin by

$$\begin{aligned}
\|Tf\|^2_{L^2(w)} &= \iint_{\mathbb{R}^2} |(f * z^{-2})(x,y)|^2 w(x,y)\,dx\,dy \\
&\ge \iint_X |(f * z^{-2})(x,y)|^2 w(x,y)\,dx\,dy \\
&= \iint_X \left| \iint_E \frac{(s^2+t^2)^{-\alpha/2}}{[(x-s)+i(y-t)]^2}\,ds\,dt \right|^2 (x^2+y^2)^{\alpha/2}\,dx\,dy.
\end{aligned} \tag{13}$$

We use the identity

$$\frac{1}{[(x-s)+i(y-t)]^2} = \frac{(x-s)^2-(y-t)^2}{[(x-s)^2+(y-t)^2]^2} - i\,\frac{2(x-s)(y-t)}{[(x-s)^2+(y-t)^2]^2}$$

to estimate the square of the modulus of the inner integral (i.e., the one over E) from above by the square of its imaginary part, that is, by

$$\left(\iint_E \frac{2(x-s)(y-t)}{[(x-s)^2+(y-t)^2]^2}(s^2+t^2)^{-\alpha/2}\,ds\,dt \right)^2.$$

Since $(x,y) \in X$ and $(s,t) \in E$, we have

$$(x-s)(y-t) \ge xy.$$

The bound for the denominator comes from the triangle inequality:

$$[(x-s)^2+(y-t)^2]^2 = |(x,y)-(s,t)|^4 \le (|(x,y)|+|(s,t)|)^4.$$

Therefore,

$$\begin{aligned}\left|\iint_E \frac{(s^2+t^2)^{-\alpha/2}}{[(x-s)+i(y-t)]^2}\,ds\,dt\right|^2 &\ge 4x^2y^2\left(\iint_E \frac{(s^2+t^2)^{-\alpha/2}}{(|(x,y)|+|(s,t)|)^4}\,ds\,dt\right)^2\\ &= \pi^2x^2y^2\left(\int_0^1 \frac{r^{-\alpha}}{(|(x,y)|+r)^4}\,r\,dr\right)^2.\end{aligned}$$

By taking $u = r/|(x,y)|$, we can continue with

$$\begin{aligned}&\pi^2x^2y^2\left(|(x,y)|^{-\alpha-2}\int_0^{1/|(x,y)|}\frac{u^{1-\alpha}}{(1+u)^4}\,du\right)^2\\ &\quad\ge \frac{\pi^2x^2y^2}{(x^2+y^2)^{\alpha+2}}\left(\int_0^1\frac{u^{1-\alpha}}{(1+1)^4}\,du\right)^2 = \frac{\pi^2}{256(2-\alpha)^2}\cdot\frac{x^2y^2}{(x^2+y^2)^{\alpha+2}}.\end{aligned}$$

Now the integral (13) can be estimated as

$$\ge \frac{\pi^2}{256\,(2-\alpha)^2}\iint_X x^2y^2(x^2+y^2)^{\alpha/2-\alpha-2}\,dx\,dy = \frac{C_l}{(2-\alpha)^3}, \tag{14}$$

where

$$C_l = \frac{\pi^2}{256}\int_0^{\pi/2}\cos^2\varphi\sin^2\varphi\,d\varphi.$$

The (square of the) right-hand side of (12) reads

$$\begin{aligned}\phi(Q_{w,2})^2\|f\|^2_{L^2(w)} &= \phi(Q_{w,2})^2\iint_{\mathbb{R}^2}|f(x,y)|^2w(x,y)\,dx\,dy\\ &= \phi(Q_{w,2})^2\iint_E (x^2+y^2)^{-\alpha}(x^2+y^2)^{\alpha/2}\,dx\,dy\\ &= \phi(Q_{w,2})^2\int_0^{\pi/2}\int_0^1 r^{-2\alpha+\alpha}\,r\,dr\,d\varphi\\ &= \phi(Q_{w,2})^2\cdot\frac{\pi}{2}\cdot\frac{1}{2-\alpha}.\end{aligned}$$

Our goal is to arrange such α that the expression on the right will be exceeded by that in (14). That is, we aim to solve the inequality

$$\frac{C_l}{(2-\alpha)^3} > \frac{\pi}{2}\phi(Q_{w,2})^2\cdot\frac{1}{2-\alpha}$$

or

$$\frac{1}{2-\alpha} > C'\phi(Q_{w,2})$$

for a given constant C'. It suffices to show that

$$\lim_{\alpha\to2^-}(2-\alpha)\phi(Q_{w,2}) = 0$$

or, equivalently, that

$$\lim_{\alpha\to2^-}(4-\alpha^2)\phi(Q_{w,2}) = 0.$$

By Lemma 3,

$$Q_{w,2} \asymp \frac{1}{4-\alpha^2}.$$

Combining this with the assumption on the growth of ϕ completes the proof. $\square$

References

[1] K. Astala, T. Iwaniec, and E. Saksman, *Beltrami operators in the plane,* Duke Math. J. 107 (2001), 27–56.

[2] R. Bañuelos and P. J. Méndez-Hernández, *Space-time Brownian motion and the Beurling–Ahlfors transform*, Indiana Univ. Math. J. (to appear).

[3] R. Bañuelos and G. Wang, *Sharp inequalities for martingales with applications to the Beurling–Ahlfors and Riesz transforms,* Duke Math. J. 80 (1995), 575–600.

[4] T. Iwaniec, *Extremal inequalities in Sobolev spaces and quasiconformal mappings,* Z. Anal. Anwendungen 1 (1982), 1–16.

[5] F. Nazarov and A. Volberg, *Heating of the Ahlfors–Beurling operator and the estimates of its norms,* St. Petersburg Math. J. (to appear).

[6] S. Petermichl, S. Treil, and A. Volberg, *Why the Riesz transforms are averages of the dyadic shifts,* Proceedings of the 6th harmonic analysis conference (El Escorial), Publ. Mat. (to appear).

[7] S. Petermichl and A. Volberg, *Heating of the Ahlfors–Beurling operator: Weakly quasiregular maps on the plane are quasiregular,* Duke Math. J. 112 (2002), 281–305.

[8] W. Rudin, *Functional analysis,* McGraw-Hill, New York, 1991.

[9] J. Wittwer, *A sharp estimate on the norm of the martingale transform,* Math. Res. Lett. 7 (2000), 1–12.

O. Dragičević
Department of Mathematics
Michigan State University
East Lansing, MI 48824

dragicev@msu.edu

A. Volberg
Department of Mathematics
Michigan State University
East Lansing, MI 48824

volberg@math.msu.edu