Publ. Mat. (2002), 209–228
Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations. El Escorial, 2000.

WHY THE RIESZ TRANSFORMS ARE AVERAGES OF THE DYADIC SHIFTS?

S. Petermichl, S. Treil* and A. Volberg*

Abstract

The first author showed in [**18**] that the Hilbert transform lies in the closed convex hull of dyadic singular operators —so-called dyadic shifts. We show here that the same is true in any $\mathbb{R}^n$ —the Riesz transforms can be obtained as the results of averaging of dyadic shifts. The goal of this paper is almost entirely methodological: we simplify the previous approach, rather than presenting the new one.

Contents

1. The "simplest" operator whose average is the Hilbert transform

Let $\mathcal{L}$ denote a dyadic *lattice* in $\mathbb{R}$. By $\mathcal{L}(k)$ we understand the dyadic *grid* of intervals from $\mathcal{L}$ having length 2^{-k}, $k \in \mathbb{Z}$. For the convenience we would like to use the notations $\mathcal{D} =: \mathcal{L}(0)$. We consider first such a dyadic lattice that the grid $\mathcal{D}$ has the point 0 as one of the end-points of its intervals. To emphasize that we write $\mathcal{D}_0$. Later we will have $\mathcal{D}_t$ —the point t plays the role of 0.

2000 *Mathematics Subject Classification.* Primary: 47B37; Secondary: 47B40.
Key words. Riesz transforms, Haar functions, dyadic shift, dyadic lattice, rectifiable measures, Menger's curvature.
*Authors are partially supported by NSF grant DMS 9970395.

Let us consider the following linear operation

$$f \to \varphi(x) := \Sigma_{I\in\mathcal{D}_0}(f, h_I)\chi_I(x).$$

Here h_I denotes the Haar function of the interval I, that is

$$h_I(x) = \begin{cases} \dfrac{-1}{|I|^{1/2}}, & \text{for } x \in I_- \\ \dfrac{1}{|I|^{1/2}}, & \text{for } x \in I_+, \end{cases}$$

and I_-, I_+ are left and right halves of the interval I correspondingly. Symbol χ_I as usual stands for the characteristic function of the interval I.

This linear operation is not even a bounded operator in $L^2(\mathbb{R})$, but it will be our main building block, so it deserves a name —$\mathbb{P}$. Actually, we will call it $\mathbb{P}_0$, thus $\varphi_0(x) := \mathbb{P}_0 f := \Sigma_{I\in\mathcal{D}_0}(f, h_I)\chi_I(x)$. Index 0 indicates the end-point of one of the intervals from $\mathcal{D}_0$. So similarly we consider

$$\varphi_t(x) := \mathbb{P}_t f$$

defined exactly as before, but with respect to the grid $\mathcal{D}_t$ of unit intervals such that the end-point of one of them is in $t \in \mathbb{R}$.

Notice that the family of grids $\mathcal{D}_t$, $t \in \mathbb{R}$, can be naturally provided with the sructure of probability space. This space is $(\mathbb{R}/\mathbb{Z}, dt) = ((-1, 0], dt)$. As usual we can use the letter ω for a point from $(-1, 0]$, and $dP(\omega)$ denotes the probability —in this case just Lebesgue measure on the interval $(-1, 0]$.

We want to fix $x \in \mathbb{R}$ and to write a nice formula for

$$\mathbb{E}\left(\varphi_\omega(x)\, dP(\omega)\right).$$

So we want to average operators $\mathbb{P}_\omega$. It can be noticed immediately that $\mathbb{E}\mathbb{P}_\omega$ is a convolution operator. In fact, let us denote by L_a the shift operator: $L_a(f)(x) = f(x + a)$. Then obviously

$$\mathbb{P}_{t-a} L_a = L_a \mathbb{P}_t.$$

Applying averaging (and the fact that our $dP(\omega)$ is invariant with respect to the natural shift on $\mathbb{R}/\mathbb{Z}$ induced by the shift on $\mathbb{R}$) we immediately get

$$\mathbb{E}\mathbb{P}_\omega L_a = L_a \mathbb{E}\mathbb{P}_\omega. \tag{1.1}$$

So the average operator $\mathbb{EP}_\omega$ is a convolution operator, we will write this as follows

$$\mathbb{E}\left(\varphi_\omega(x)\, dP(\omega)\right) = \mathbb{EP}_\omega f(x) = F_0 * f(x). \tag{1.2}$$

It is easy to compute F_0. By the definition of $\varphi_t(x)$ one can write (see Figure 1)

$$\varphi_t(x) = \int f(s) h^{t-\frac{1}{2}}(s)\, ds, \quad x - \frac{1}{2} < t - \frac{1}{2} < x + \frac{1}{2}, \tag{1.3}$$

where

$$h^t(s) = \begin{cases} -1, & s \in (t - 1/2, t) \\ +1, & s \in (t, t + 1/2). \end{cases}$$

But $h^t(s) = k_0(t-s)$, where

$$k_0(s) = \begin{cases} +1, & s \in (-1/2, 0) \\ -1, & s \in (0, 1/2). \end{cases}$$

So (1.3) can be rewritten as follows

$$\varphi_{t+\frac{1}{2}}(x) = \int f(s) k_0(t-s)\, ds, \quad x - \frac{1}{2} < t < x + \frac{1}{2}. \tag{1.4}$$

Thus comparing this with (1.2) (and using again the shift invariance of $d\mathbb{P}(\omega)$) we get

$$\begin{aligned} F_0 * f(x) &= \mathbb{E}\left(\varphi_\omega(x)\, dP(\omega)\right) \\ &= \mathbb{E}\left(\varphi_{\omega+\frac{1}{2}}(x)\, dP(\omega)\right) = \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} \left(\int f(s) k_0(t-s)\, ds\right) dt. \end{aligned}$$

From which we get the formula for F_0:

$$F_0(x) = \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} k_0(t)\, dt = k_0 * \chi_0(x), \tag{1.5}$$

where χ_0 is the characteristic function of the unit interval $(-1/2, 1/2)$. On Figure 2 one can see the graph of F_0.

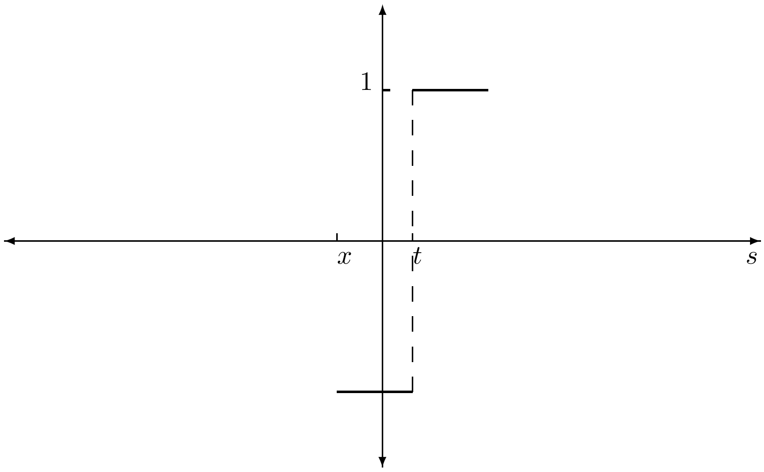

FIGURE 1. Function $h^t(s)$

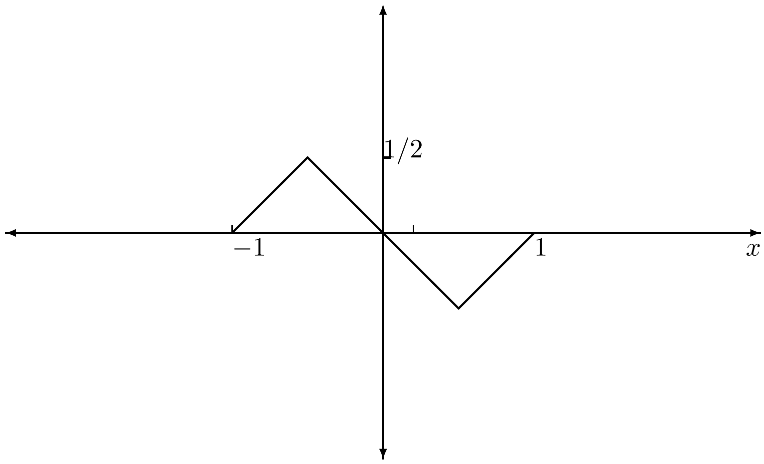

FIGURE 2. Function F_0

Let us start over the beginning of this section with one slight difference —we rescale all our operators, and now $\mathbb{P}^\rho_t$, φ^ρ_t, F^ρ_0, k^ρ_0 are *precisely* as above, but when the unit length intervals are replaced by intervals of length $\rho > 0$. We just change the scale —nothing else. In particular,

$$\varphi^\rho_0(x) := \mathbb{P}^\rho_0 f := \Sigma_{I\in\mathcal{D}^\rho_0}(f, h_I)\chi_I(x)/\sqrt{\rho}$$

where $\mathcal{D}_0^\rho$ is the grid of intervals of length ρ such that 0 is the end-point of two intervals from this grid. We want to remind that h_I here is always normalized in L^2.

Again we have a natural probability space of all grids of intervals of size ρ: $\left(\mathbb{R}/\rho\mathbb{Z};\ \frac{1}{\rho}\, dt \mid (-\rho, 0]\right)$.

$$\varphi_t^\rho(x) := \mathbb{P}_t^\rho f := \Sigma_{I\in\mathcal{D}_t^\rho}(f, h_I)\chi_I(x)/\sqrt{\rho}.$$

Averaging over all grids of intervals of size ρ makes $\mathbb{P}_t^\rho$ a convolution operator —there is no difference with our reasoning above. It is easy to see that this is the convolution operator with the kernel

$$F_0^\rho(x) := \frac{1}{\rho}\int_{x-\frac{\rho}{2}}^{x+\frac{\rho}{2}} \frac{1}{\rho} k_0\left(\frac{t}{\rho}\right) dt = \frac{1}{\rho} F_0\left(\frac{x}{\rho}\right). \tag{1.6}$$

The first $\frac{1}{\rho}$ is because of the form our probability has. The second $\frac{1}{\rho}$ because we should average a function normalized in L^1.

Let us now consider *all* convolution operators with kernels F_0^ρ. Let us fix $r \in [1, 2)$ and let us take a look at the convolution operator with kernel

$$F_r = \Sigma_{n=-\infty}^{\infty} F_0^{2^n r}. \tag{1.7}$$

The grids $\mathcal{D}_t^{2^n r}$ (t is fixed) can be united into *a "dyadic" lattice* $\mathcal{L}_t^r$. Here t means the reference point —one of the end-point of intervals from our lattice, and r means the length of one of the intervals of the lattice— let us call r the *calibre* of the lattice. Obviously the convolution operator with the kernel F_r is the averaging over *all "dyadic" lattices* (not grids!) $\mathcal{L}_t^r$ of fixed calibre r of the operators given by

$$\mathcal{P}_{\mathcal{L}_t^r} f = \Sigma_{I\in\mathcal{L}_t^r}(f, h_I)\chi_I(x)/\sqrt{|I|}$$
$$F_r * f = \mathbb{E}\mathcal{P}_{\mathcal{L}_t^r} f.$$

This is just because the kernel F_r is the sum of kernels, each of which appeared as averaging of the grid opearators assigned to grids of size $2^n r$, $n = 0, \pm 1, \pm 2, \dots$, where we summed up over the grids, and the lattice of calibre r is the union of such grids.

Now let us finally average over $r \in [1, 2)$:

$$F(x) := \int_1^2 F_r(x)\frac{dr}{r}.$$

Now we have from one side

$$F * f = (\text{Average}\, \mathcal{P}_{\mathcal{L}})\, f, \tag{1.8}$$

where averaging is performed over all lattices $\mathcal{L}_t^r$.

On the other hand it is easy to compute Φ.

$$\begin{aligned} F(x) &= \int_1^2 F_r(x) \frac{dr}{r} \\ &= \int_1^2 \Sigma_{n=-\infty}^{\infty} F_0^{2^n r} \frac{dr}{r} = \int_0^\infty F_0^\rho \frac{d\rho}{\rho} = \int_0^\infty F_0\left(\frac{x}{\rho}\right) \frac{d\rho}{\rho^2}. \end{aligned} \tag{1.9}$$

We used (1.6) here. Finally we have (see Figure 2)

$$F(x) = -\frac{1}{x} \int_0^\infty F_0(t)\, dt = \frac{1}{4} \frac{1}{x}. \tag{1.10}$$

Theorem 1.1. *Averaging of operators $\mathcal{P}_{\mathcal{L}_t^r}$ over both parameters t and r is equal to one quarter of the Hilbert transform.*

Proof: Just compare (1.8) and (1.10). □

We have a good thing:

The Hilbert transform is the averaging over the family of lattices of very simple operators

and a bad thing:

These simple operators are not bounded in L^2.

We want the best of the both worlds: a) the Hilbert transform is the averaging over the family of lattices of very simple operators; b) these simple operators have to be bounded in L^2.

2. What is the dyadic shift?

The function that generated everything in the first section was function F_0 —the kernel of the convolution operator which is the averaging of grid operators $\mathbb{P}_t$. It is easy to see that $F_0(x \pm 1)$ are also kernels of the convolution operators which are the averagings of some grid operators. Given f, let us consider $\varphi_t(x)$ as above and also

$$\varphi_t(x+1) = \Sigma_{I \in \mathcal{D}_t} (f, h_{I-1}) \chi_I(x) = \Sigma_{I \in \mathcal{D}_t} (f, h_I) \chi_{I+1}(x) =: \mathbb{P}_t^+(f)$$

$$\varphi_t(x-1) = \Sigma_{I \in \mathcal{D}_t} (f, h_{I+1}) \chi_I(x) = \Sigma_{I \in \mathcal{D}_t} (f, h_I) \chi_{I-1}(x) =: \mathbb{P}_t^-(f).$$

So we test f on h_I and put the result on $I \pm 1$.

What if we average these operators? Repeating (1.2) we get

$$\left(\int_0^1 \mathbb{P}_t^{\pm}\,dt\right) f = F_0(x \mp 1) * f. \tag{2.1}$$

Consider

$$S(x) := F_0(x) - \frac{1}{2}[F_0(x+1) + F_0(x-1)]. \tag{2.2}$$

Supposedly S is a kernel of a convolution operator corresponding to averaging over grids of a certain grid operator (we will show which one). If we build S^ρ as before for all calibres, we can consider again $S_r := \Sigma_{n=-\infty}^{\infty} S^{2^n r}$. Operators S_r are averagings over all *lattices* of calibre r of the operators which are sums of our hypothetical grid operators. Averaging over $r \in [1,2)$ with respect to the measure dr/r, we will get the operator with kernel (see Figure 3 and (1.9))

$$\int_1^2 S_r(x)\frac{dr}{r} = \int_0^\infty S\left(\frac{x}{\rho}\right)\frac{d\rho}{\rho^2} = \frac{1}{x}\int_0^\infty S(t)\,dt = \frac{1}{4}\frac{1}{x}. \tag{2.3}$$

So we are left to invent a simple "grid" operator, whose average will give us $S(x)$.

Theorem 2.1. *Let $\mathcal{D}_t^{(2)}$ be a grid of intervals of length 2 such that t is the end-point. Consider operators*

$$f \to \Sigma_{J\in\mathcal{D}_t^{(2)}}(f, h_{J_-})\chi_{J_+}$$

$$f \to \Sigma_{J\in\mathcal{D}_t^{(2)}}(f, h_{J_+})\chi_{J_-}$$

$$f \to \Sigma_{J\in\mathcal{D}_t^{(2)}}(f, h_{J_-})\chi_{J_-}$$

$$f \to \Sigma_{J\in\mathcal{D}_t^{(2)}}(f, h_{J_+})\chi_{J_+}.$$

The averaging over t of the first operator gives a convolution with kernel $\frac{1}{2}F_0(x-1)$, the averaging over t of the second operator gives a convolution with kernel $\frac{1}{2}F_0(x+1)$, and the averaging over t of the third and the fourth operator gives a convolution with kernel $\frac{1}{2}F_0(x)$ each.

Proof: Let us call the first operator H_t, and let us average $\mathbb{E}H_t$ it over its probability space $\big(\mathbb{R}/2\mathbb{Z};\, \frac{1}{2}\,dt \mid (-2,0]\big)$. Instead of considering the grid of intervals of length 2 let us consider the grid of intervals of length 1 —we call it $\mathcal{D}_t^1$. Consider operators $A_t\colon f \to \Sigma_{I \text{ is odd},\, I\in\mathcal{D}_t^{(1)}}(f, h_I)\chi_{I+1}$, $B_t\colon f \to \Sigma_{I \text{ is even},\, I\in\mathcal{D}_t^{(1)}}(f, h_I)\chi_{I+1}$. Clearly $A_{t+1} = B_t$. Also it is clear that $A_t + B_t = \mathbb{P}_t^+$, where the last operator is our grid operator from the beginning of Section 2.

$$\mathbb{E}H_t = \frac{1}{2}\int_0^1 (A_t + A_{t+1}) = \frac{1}{2}\int_0^1 (A_t + B_t) = \frac{1}{2}\int_0^1 \mathbb{P}_t^+ \, dt.$$

From (2.1) we get that

$$\mathbb{E}H_t = \frac{1}{2}F_0(x-1) * .$$

Similarly, if we call the second operator G_t we get from (2.1)

$$\mathbb{E}G_t = \frac{1}{2}F_0(x+1) * .$$

Using (1.2) and (1.5) we show that averagings of the third and the fourth operators give us convolution operator with kernel $\frac{1}{2}F_0$. Theorem 2.1 is proved. □

Theorem 2.2. *Let us consider the following grid operator*

$$f \to \Sigma_{J\in\mathcal{D}_t^{(2)}}(f, h_{J_+} - h_{J_-})h_J, \quad t \in \left(\mathbb{R}/2\mathbb{Z};\, \frac{1}{2}\,dt \mid (-2,0]\right).$$

Then its averaging is the convolution operator with kernel $\frac{1}{\sqrt{2}}S(x)$.

Proof: We write h_J as $\frac{1}{\sqrt{2}}(-\chi_{J_-} + \chi_{J_+})$. Then it is an obvious algebraic remark that

$$\begin{aligned}\sqrt{2}\text{ our operator} &= \text{third operator of Theorem 2.2}\\ &\quad + \text{fourth operator of Theorem 2.2}\\ &\quad - \text{first operator of Theorem 2.2}\\ &\quad - \text{second operator of Theorem 2.2.}\end{aligned}$$

Averaging this and using Theorem 2.1 finishes the proof. □

As in the previous section, given the lattice $\mathcal{L} = \mathcal{L}_t^r$, we can consider the lattice operator

$$\text{Ш}^{\mathcal{L}} f := \Sigma_{J\in\mathcal{L}}(f, h_{J_+} - h_{J_-})h_J$$

amalgamated from the grid operators of Theorem 2.2.

This operator is called *the dyadic shift*. It has been proved in [**18**] that averaging of dyadic shifts over all lattices gives us operator which is proportional to the Hilbert transform (we certainly mean that coefficient of proportionality is not zero).

Let us reproduce this result. Fixing r and averaging over lattices with fixed calibre r (we leave for the reader to invent the natural probability space of all lattices with fixed calibre r) we get the convolution operator with the kernel

$$\frac{1}{\sqrt{2}} \sum_{n=-\infty}^{\infty} \frac{1}{2^n r} S\left(\frac{x}{2^n r}\right) =: \frac{1}{\sqrt{2}} S_r(x).$$

Averaging convolution operators with kernels $\frac{1}{\sqrt{2}} S_r$ over $\left([1,2); \frac{dr}{r}\right)$, we get (see (2.3)) the operator with the kernel $\frac{1}{4}\frac{1}{\sqrt{2}}\frac{1}{x}$. So we get

(2.4) Averaging of the shift operators over all lattices of all calibres

$$= \frac{1}{4\sqrt{2}} \text{ the Hilbert transform.}$$

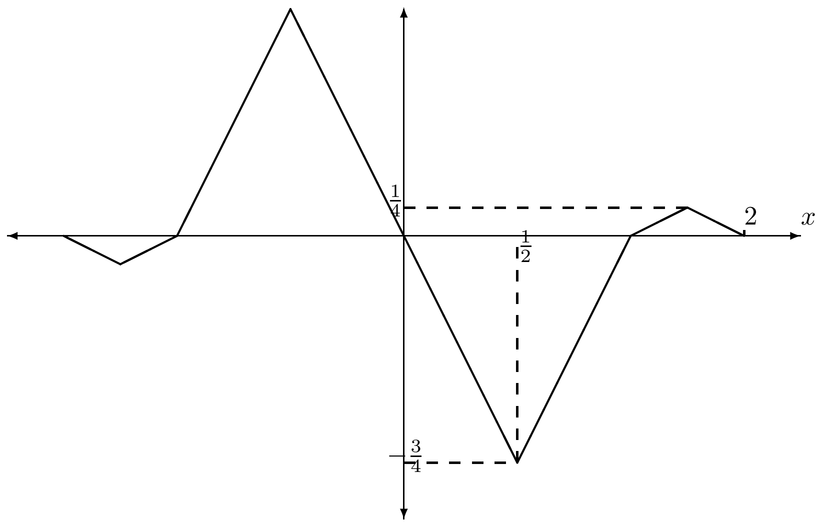

FIGURE 3. Function S

3. Planar case

We can and will reason by analogy. We have lattices $\mathcal{L}_t^\rho$ of squares, where t now is in $\Omega^\rho := \mathbb{R}^2/\rho\mathbb{Z}^2$ with normalized Lebesgue measure (Lebesgue measure on the torus Ω^ρ divided by ρ^2). We have the main grid operator

$$\mathbb{P}_t f := \Sigma_{Q\in\mathcal{D}_t}(f, h_Q)\chi_Q$$

where $\mathcal{D}_t$ is a grid of unit squares such that $t \in \mathbb{R}^2$ is a vertex for 4 of them, where

$$h_Q(x) := \begin{cases} \frac{-1}{\sqrt{|Q|}}, & \text{for } x \in Q_l \\ \frac{1}{\sqrt{|Q|}}, & \text{for } x \in Q_r \\ 0, & \text{otherwise.} \end{cases}$$

Here Q_l, Q_r are left and right halves of Q, function h_Q is normalized in L^2. We consider the same type of grid operators for grids $\mathcal{D}_t^\rho$ of squares of side ρ —the only change is that we divide χ_Q by ρ to make it normalized in L^2.

Let us denote by k_0 the function $-h_{Q_0}$, where Q_0 is the unit square centered at 0. Also χ_0 denotes the characteristic function of this square.

Consider

$$\Phi_0 := \chi_0 * k_0, \quad \Phi_0^\rho(x) := \frac{1}{\rho^2}\frac{1}{\rho^2}\chi_0\left(\frac{\cdot}{\rho}\right) * k_0\left(\frac{\cdot}{\rho}\right) = \frac{1}{\rho^2}\Phi_0\left(\frac{x}{\rho}\right).$$

Exactly as before (in one dimensional case) function Φ_0 is the kernel of the convolution operator, which appears as averaging of $\mathbb{P}_t$ over Ω^1. Function Φ_0^ρ is the kernel of the convolution operator, which appears as averaging of $\mathbb{P}_t^\rho$ over Ω^ρ.

Again, we can consider kernel

$$k(x) := \int_0^\infty \Phi_0^\rho(x)\frac{d\rho}{\rho} = \frac{\omega\left(\frac{x}{|x|}\right)}{|x|^2}.$$

And it is very easy to see that ω is an odd *non-zero* function on the unit circle. Literally as before we can see that k is the convolution operator which is the average with respect to measure $\frac{dr}{r} \mid [1,2)$ of the convolution operators with kernels

$$k_r(x) := \sum_{n=-\infty}^{\infty} \Phi_0^{r\cdot 2^n}(x).$$

In its turn, k_r is the average of the lattice operators which are sums of corresponding grid operators, here are those lattice operators:

$$\mathcal{P}_{\mathcal{L}^r} := \Sigma_{Q\in\mathcal{L}^r}(f, h_Q)\chi_Q/\sqrt{|Q|}.$$

Here r is fixed and denotes the calibre of the lattice. The averaging over the lattices of this fixed calibre gives us the convolution operator with kernel k_r. So the averaging over the calibres $(= \int_1^2 \dots \frac{dr}{r})$ gives us the averaging over all lattices, over all calibres. As a result we get the convolution operator with kernel $k = \frac{\omega\left(\frac{x}{|x|}\right)}{|x|^2}$.

Again we would like to repeat all this but with slightly different lattice operators —just because there are nicer ones and because $\mathcal{P}_{\mathcal{L}^r}$ are not L^2 bounded. Another problem we face now is that k is not necessarily a kernel of a Riesz transform. So we will need to work a bit more than in the one-dimensional case to obtain the Riesz transform kernel.

For a square Q consider its partition to 4 equal squares and let us call them Q^{nw}, Q^{ne}, Q^{sw}, Q^{se} according to northwest, northeast, ... Let us consider the following grid operator

$$f \to \Sigma_{Q\in\mathcal{D}_t^{(2)}}(f, h_{Q^{ne}} + h_{Q^{se}} - h_{Q^{nw}} - h_{Q^{sw}})h_Q,$$
$$t \in \Omega^{(2)} := \left(\mathbb{R}^2/2\mathbb{Z}^2;\ \frac{1}{4}\text{ Lebesgue measure}\right).$$

Consider also the function $(x = (x_1, x_2))$

$$\begin{aligned} S(x_1,x_2) &= \Phi_0(x_1,x_2) - \frac{1}{2}\Phi_0(x_1+1,x_2) - \frac{1}{2}\Phi_0(x_1-1,x_2) \\ &+ \frac{1}{2}\Phi_0(x_1,x_2+1) - \frac{1}{4}\Phi_0(x_1+1,x_2+1) - \frac{1}{4}\Phi_0(x_1-1,x_2+1) \\ &+ \frac{1}{2}\Phi_0(x_1,x_2-1) - \frac{1}{4}\Phi_0(x_1+1,x_2-1) - \frac{1}{4}\Phi_0(x_1-1,x_2-1). \end{aligned} \tag{3.1}$$

Theorem 3.1. *The averaging of the grid operator above over $\Omega^{(2)}$ gives the convolution operator with kernel $\frac{1}{2}S(x)$.*

The proof is literally the same as the proof of Theorem 2.2.

Let us start with one observation about (3.1). Function Φ_0 is the convolution $\chi_0 * k_0$. But both functions χ_0 and k_0 are products of functions of one variable —-$\Phi_0(x_1,x_2) = f_0(x_2)\cdot F_0(x_1)$. Moreover, function f_0 is nonnegative. Actually $f_0(x_2)$ is a convolution square of the characteristic function of the unit interval centered at 0. Formula (3.1) now looks like

$$\begin{aligned} S(x_1,x_2) = \left(f_0(x_2) + \frac{1}{2}f_0(x_2+1) + \frac{1}{2}f_0(x_2-1)\right) \\ \times \left(F_0(x_1) - \frac{1}{2}F_0(x_1+1) - \frac{1}{2}F_0(x_1-1)\right). \end{aligned}$$

For the future purposes we can say what happens in $n > 2$ case easily. We get $S_n(x) = S_n(x_1, x_2, \dots, x_n)$ and

$$\text{(3.2)} \quad S_n(x) = \left(F_0(x_1) - \frac{1}{2}F_0(x_1+1) - \frac{1}{2}F_0(x_1-1)\right) \times \prod_{i=2}^{n} \left(f_0(x_i) + \frac{1}{2}f_0(x_i+1) + \frac{1}{2}f_0(x_i-1)\right).$$

As in the previous section this S generates kernel s by formula

$$s(x) = \int_0^\infty \frac{1}{\rho^n} S\left(\frac{x}{\rho}\right) \frac{d\rho}{\rho} = \frac{\xi_n\left(\frac{x}{|x|}\right)}{|x|^n}.$$

And it is very easy to see that ξ_n is an odd *non-zero* function on the unit sphere. We will show it below. Literally as before we can see that s is the convolution operator which is the average with respect to measure $\frac{dr}{r} \mid [1,2)$ of the convolution operators with kernels

$$s_r(x) := \sum_{n=-\infty}^{\infty} S_0^{r\cdot 2^n}(x).$$

In its turn, s_r is the average of the lattice operators which are sums of corresponding grid operators, here are those lattice operators:

$$\text{(3.3)} \qquad \mathcal{S}_{\mathcal{L}^r} := \Sigma_{Q \in \mathcal{L}^r}(f, h_{Q^{ne}} + h_{Q^{se}} - h_{Q^{nw}} - h_{Q^{sw}}) h_Q.$$

Here r is fixed and denotes the calibre of the lattice. The averaging over the lattices of this fixed calibre gives us the convolution operator with kernel s_r. So the averaging over the calibres $(= \int_1^2 \dots \frac{dr}{r})$ gives us the averaging over all lattices, over all calibres. As a result we get the convolution operator with kernel $s = \frac{\xi_n\left(\frac{x}{|x|}\right)}{|x|^n}$.

Let S^{n-1} denote as always the boundary sphere of the n-dimensional unit ball. Denote by S_+^{n-1} the right half sphere —the half that lies in $\{x \in \mathbb{R}^n : x_1 > 0\}$. Let e_1 be a unit vector in the direction of coordinate axis x_1. Let $\langle \cdot \rangle$ denote the scalar product in $\mathbb{R}^n$. Let σ denote Lebesgue measure of S^{n-1}. It would be important to prove

$$\text{(3.4)} \qquad \int_{S_+^{n-1}} \xi_n(\omega) \langle \omega, e_1 \rangle \, d\sigma(\omega) < 0.$$

For $n = 2$ we can just prove that $\xi_2(\omega) < 0$ for any $\omega \in S_+^1$. Then (3.4) follows immediately. To do this we use formula (3.2) and notice that $f_0(x) + \frac{1}{2}f_0(x+1) + \frac{1}{2}f_0(x-1) = (1 - \frac{1}{2}x)_+$. Then the fact that $\xi_2(\omega) < 0$ follows from the following lemma.

Lemma 3.2. *For any $k \in [0,\infty)$ we have*

$$\int_0^2 \left(1-\frac{1}{2}kx\right)_+ \left(F_0(x)-\frac{1}{2}F_0(x-1)\right) x\,dx < 0.$$

Proof: If $k \geq 2$ then the first factor vanishes everywhere where the second factor is positive. So we are done for such k. For $0 \leq k \leq 1$ we have $(1-\frac{1}{2}kx)_+ = (1-\frac{1}{2}kx)$ on $[0,2]$, and we can make an easy calculation of the integral. For the range $1 < k < 2$ the calculation becomes unpleasant, but still straightforward, we skip it just to avoid direct and simple calculations. □

For $n=2$, ω can be identified with a point of $[-\pi,\pi)$. Under this identification the kernel ξ_2 becomes an even function skew symmetric on $[0,\pi]$ with respect to the point $\pi/2$. Rotation of the kernel $\xi_2(\omega)$ means just the new kernel $\xi_2(\omega-\phi)$. Then

$$\begin{aligned}(\xi_2 * \cos)(\phi) &= \cos\phi \cdot \left(\int_{-\pi}^{\pi} \xi_2(s)\cos s\,ds\right)\\ &= \cos\phi\cdot\left(\int_{S^1}\xi_n(\omega)\langle\omega,e_1\rangle\,d\sigma(\omega)\right) = c_2\cos\phi,\end{aligned} \tag{3.5}$$

and constant $c_2 = \int_{S^1}\xi_n(\omega)\langle\omega,e_1\rangle\,d\sigma(\omega) \neq 0$ because of (3.4).

Consider $A_2 := \frac{\int_{-\pi}^{\pi}|\cos s|\,ds}{|c_2|}$. Notice that rotation of kernel ξ_2 corresponds to rotation of dyadic lattices on the plane. We have just proved the following theorem.

Theorem 3.3. *The Riesz transform $\frac{x_1}{|x|^3}*$ is the operator integral $c_2^{-1}\int\cos\psi\frac{\xi_2\left(U_\psi\frac{x}{|x|}\right)}{|x|^3} * d\psi$. In particular, this means that operator with the kernel $A_2^{-1}\frac{x_1}{|x|^3}$ lies in the closed convex hull (in the weak operator topology) of the planar dyadic shifts. Thus, uniform boundedness of dyadic shift operators in any Banach space implies the boundedness of the Riesz transform in the same space.*

For the case $n>2$ we again start with (3.4). Let us average ξ_n with respect to all rotations that leave e_1 fixed. We get a new function $\eta_n(\omega) = f(\langle\omega,e_1\rangle)$. Obviously,

$$\int_{S_+^{n-1}} f(\langle\omega,e_1\rangle)\langle\omega,e_1\rangle\,d\sigma(\omega) < 0. \tag{3.6}$$

Let SO is the group of orthogonal rotations of S^{n-1}.

Let us calculate $c_n = \int_{SO} f(\langle Ue_1, e_1\rangle)\langle Ue_1, e_1\rangle\, dU$. Obviously,

$$c_n = \int_{S^{n-1}} f(\langle \omega, e_1\rangle)\langle \omega, e_1\rangle\, d\sigma(\omega) \neq 0,$$

because of (3.6). Now let us consider the rotated functions $f(\langle U\omega, e_1\rangle)$. Consider

$$g(\omega) = \int_{SO} f(\langle U\omega, e_1\rangle)\langle Ue_1, e_1\rangle\, dU.$$

Then it is clear that $g(R\omega) = g(\omega)$ for every $R \in SO$ that fixes e_1. In fact,

$$\begin{aligned} g(R\omega) &= \int_{SO} f(\langle UR\omega, e_1\rangle)\langle Ue_1, e_1\rangle\, dU \\ &= \int_{SO} f(\langle V\omega, e_1\rangle)\langle VR^* e_1, e_1\rangle\, dV \\ &= \int_{SO} f(\langle V\omega, e_1\rangle)\langle Ve_1, e_1\rangle\, dV = g(\omega). \end{aligned}$$

On the other hand, it easy to see that

$$g(\omega) = \int_{S^{n-1}} f(\langle \omega, \xi\rangle)\langle \xi, e_1\rangle\, d\sigma(\xi). \tag{3.7}$$

Such a function (as we saw) depends only on $\langle \omega, e_1\rangle$. But moreover, it can be written as $\int_{S^{n-1}} f(\langle e_1, \xi\rangle)\langle \xi, \omega\rangle\, d\sigma(\xi)$. This is a restriction of a linear polynomial onto the sphere. This linear polynomial depends on $\langle \omega, e_1\rangle$ only, and, thus, is $c \cdot \langle \omega, e_1\rangle$. The constant c is just our c_n. One can see that by plugging $\omega = e_1$ into our formula (3.7) for $g(\omega)$.

Consider $A_n := \frac{\int_{SO} |\langle Ue_1, e_1\rangle|\, dU}{|c_n|}$. Notice that rotation of kernel ξ_n corresponds to rotation of dyadic lattices on the plane. We have just proved the following theorem.

Theorem 3.4. *The Riesz transform $\frac{x_1}{|x|^{n+1}} *$ is the operator integral*

$$c_n^{-1} \int_{SO} \langle Ue_1, e_1\rangle \frac{\eta_n\left(U\frac{x}{|x|}\right)}{|x|^{n+1}} * dU.$$

In particular, this means that operator with the kernel $A_n^{-1}\frac{x_1}{|x|^{n+1}}$ lies in the closed convex hull (in the weak operator topology) of the planar dyadic shifts. Thus, uniform boundedness of dyadic shift operators in any Banach space implies the boundedness of the Riesz transform in the same space.

Remind that we have proved (3.4) inequality only for the case $n = 2$ so far (see Lemma 3.2). We do not know how to compute the integral in (3.4) for the $n > 2$ dimensional analog of the operator in (3.3). However, we are going to introduce the operator similar to the one in (3.3) that will give us (3.4) immediately. Here is the description of this operator. For every cube Q of a dyadic lattice $\mathcal{L}$ in $\mathbb{R}^n$ we denoted by h_Q the function supported by Q and such that it is equal to $-\frac{1}{|Q|^{1/2}}$ on the left half Q_l of Q and is equal to $\frac{1}{|Q|^{1/2}}$ on the right half Q_r of Q. This just one of the Haar functions. We are going to choose where this function should be mapped by our dyadic shift by using the following considerations. The image must be the combination of Haar functions of the previous generation. That means that it should be supported by the father $\bar{Q}$ of Q, should be constant on each son of $\bar{Q}$ (including Q), and the sum of these costants must be zero. So the only choice is the choice of constants c_B, where B is either Q or one of $2^n - 1$ of its brothers. When $n = 1$ we made a correct choice by using the rule $c_Q = 1$, $c_B = -1$ for the only brother B of Q. One of the natural choices now would be

$$c_Q = \frac{2^n - 1}{|Q|^{1/2}}, \quad c_B = \frac{-1}{|Q|^{1/2}}, \quad B \text{ is the brother of } Q.$$

The operator which sends each h_Q, $Q \in \mathcal{L}$, to the corresponding function on $\bar{Q}$ is called $\text{II}^{\mathcal{L}}$ (we should say that it maps all other Haar functions to zero). And the operator, which does this for a dyadic grid G will be called P^G. Let us consider all dyadic grids of cubes $\bar{Q}$ with sidelength 2. And let us consider the averaging $\mathcal{P} = \mathbb{E}P^G$ over a probability space of all such dyadic grids. Operator $\mathcal{P}$ is of course a convolution operator. Let us call its kernel p. Remind that $S^n_+ = \{\omega \in \mathbb{R}^n : |\omega| = 1, \langle \omega, e_1\rangle > 0\}$ is the right half sphere. Obviously the next Lemma 3.5 proves (3.4), and we finish the proof that the Riesz transform R_1 can be “decomposed” into the dyadic shifts.

Lemma 3.5. *For any* $\omega \in S^n_+$, $\int_0^\infty p(\frac{\omega}{\rho})\frac{1}{\rho^n}\frac{d\rho}{\rho} > 0$.

Proof: For any given G of cubes $\bar{Q}$ of sidelength 2 let us split P^G into two operators. The first will be called V^G and it maps $h_{\bar{Q}_l}$ into $2^n\chi_{\bar{Q}_l}$, $h_{\bar{Q}_r}$ into $2^n\chi_{\bar{Q}_r}$. In other words, if Q is a unit cube that happens a son of a cube from G, then $V^G(h_Q) = 2^n\chi_Q$. And it maps all other Haar function to zero. The rest will be called W^G. In other words, if Q is a unit cube that happens a son of a cube $\bar{Q}$ from G, then $W^G(h_Q) = \sum_{B \text{ is the son of } \bar{Q}} \chi_Q = \chi_{\bar{Q}}$. As always, let us consider all dyadic grids of cubes $\bar{Q}$ with sidelength 2. And let us consider the averagings $\mathcal{V} = \mathbb{E}V^G$,

$\mathcal{W} = \mathbb{E}W^G$ over a probability space of all such dyadic grids. Operators $\mathcal{V}$, $\mathcal{W}$ are of course convolution operators. Let us call their kernels v, w. Denote $w_2(x) = 2^n w(2x)$. Notice that $\int_0^\infty w_2(\frac{\omega}{\rho})\frac{1}{\rho^n}\frac{d\rho}{\rho} = \int_0^\infty w(\frac{\omega}{\rho})\frac{1}{\rho^n}\frac{d\rho}{\rho}$. Remind that $p = v - w$. Now we can see that to prove the lemma we obviously just need to show that for any $\omega \in S_+^n$

$$\int_0^\infty v\left(\frac{\omega}{\rho}\right)\frac{1}{\rho^n}\frac{d\rho}{\rho} - \int_0^\infty w_2\left(\frac{\omega}{\rho}\right)\frac{1}{\rho^n}\frac{d\rho}{\rho} < 0. \tag{3.8}$$

We will just see that

$$v(x) - w_2(x) \leq 0. \tag{3.9}$$

It is very easy to see that inequality (3.9) gives a strict inequality in (3.8). So let us see why (3.9) holds by just computing the kernels. We will use the notation F_0 from the first section and f_0 is the convolution of the characteristic function $\chi_{[-1/2,1/2]}$ with itself. Then it is easy to see that

$$\begin{aligned} v(x) &= 2^n F_0(x_1) f_0(x_2) \cdot \ldots \cdot f_0(x_n) \\ w(x) &= 2^{-n}(2F_0(x_1) + F_0(x_1 - 1) + F_0(x_1 + 1))(2f_0(x_2) + f_0(x_2 - 1) \\ &\qquad + f_0(x_2 + 1)) \cdot \ldots \cdot (2f_0(x_n) + f_0(x_n - 1) + f_0(x_n + 1)). \end{aligned}$$

Let $a(t) := 2F_0(t)+F_0(t-1)+F_0(t+1)$, $b(t) := 2f_0(t)+f_0(t-1)+f_0(t+1)$. It is easy to check that

$$4F_0(t) \leq 2a(2t), \quad 4f_0(t) = 2b(2t).$$

Thus, $2^n F_0(x_1)f_0(x_2) \cdot \ldots \cdot f_0(x_n) - 2^n w(2x) \leq 0$. Inequality (3.9) is completely proved, and this proves the lemma. $\square$

3.1. Dimension tuning.

Let us consider a variant of dyadic shift, but slightly rescaled. Namely, let $\mathcal{L}$ be any "dyadic" (actually r times dyadic) lattice dropped on the plane. We have introduced dyadic planar shifts $\text{Ш}_{\mathcal{L}} = \Sigma_{Q\in\mathcal{L}} sh_Q$, here sh_Q is a rank one operator described in the previous section (or, for that matter, any other "dyadic shift operator", for example, the one from (3.3)). Let us call them dyadic planar shifts of order 2. The dyadic planar shift of order d is

$$\text{Ш}_{\mathcal{L}}^d := \Sigma_{Q\in\mathcal{L}}|Q|^{\frac{2-d}{2}} sh_Q.$$

As before by averaging over lattices we get kernels $p^d(x)$, $\zeta_\alpha^d(x)$.

With no changes we get

Theorem 3.6. *The planar Riesz transform with kernel $\frac{x_1}{|x|^{1+d}}$ is the operator integral $\int g(\alpha)\zeta_\alpha^d *$, in particular, the A^{-1} multiple of this operator is equal to a certain averaging of dyadic planar shifts of order d. This means that operator with the kernel $A^{-1}\frac{x_1}{|x|^{1+d}}$ lies in the closed convex hull (in the weak operator topology) of the planar dyadic shifts of order d. Thus, uniform boundedness of dyadic shift operators of order d in any Banach space implies the boundedness of the Riesz transform in the same space.*

4. Geometric application

The most interesting case is $d = 1$. Then we know what are the measure μ on the plane such that operators $\frac{x_1}{|x|^2}$, $\frac{x_2}{|x|^2}$ are bounded in $L^2(\mu)$. See [**13**], [**20**]. Description of such measures was the solution of the famous problem. For arclength measure on curves it has been found by Guy David [**4**]. They are the same as those for which the Cauchy transform is bounded. We will call them here *rectifiable measures*. One can find the explanation for this name in the groundbreaking article [**11**]. We have the following theorem in which $\ell(Q)$ denotes the side length of the square Q.

Theorem 4.1. *If dyadic shift operators $\text{Ш}^1_{\mathcal{L}} := \Sigma_{Q\in\mathcal{L}}\ell(Q)sh_Q$ of order 1 are all bounded as operators from $L^2(\mu)$ to itself, then μ is rectifiable.*

One can prove easily

Theorem 4.2. *The dyadic shift operators $\text{Ш}^1_{\mathcal{L}} := \Sigma_{Q\in\mathcal{L}}|Q|^{\frac{1}{2}}sh_Q$ of order 1 are all bounded as operators from $L^2(\mu)$ to itself if and only if*

$$\mu(Q) \le C_1\ell(Q), \quad \forall\, Q \tag{4.1}$$

and for any lattice $\mathcal{L}$ and any square $R \in \mathcal{L}$ the following oscillation criterion for the measure μ is satisfied

$$\sum_{Q\in\mathcal{L},\, Q\subset R} \frac{(\mu(Q_l) - \mu(Q_r))^2}{\ell(Q)} \le C_2\mu(R). \tag{4.2}$$

Unfortunately, the condition (4.2) is too strong —it is not satisfied even for Lebesgue measure on a straight segment! We can propose a much weaker condition on μ which corresponds to boundedness of "dyadic operators" in average versus their uniform boundedness.

Here is this condition: for any dyadic lattice $\mathcal{L}$

$$\sum_{Q\in\mathcal{L},\, Q\subset R} \frac{\mu(Q)(\mu(Q_l) - \mu(Q_r))^2}{\ell(Q)^2} \le C_2\mu(R). \tag{4.3}$$

One can compare it with curvature condition in [**12**] and may wonder whether they are equivalent. If yes, it is getting interesting because (4.3) can be obviously extended to $\mathbb{R}^n$, $n > 2$ —and it is a very interesting problem what replaces the curvature condition for the case $n > 2$. Our (4.3) or its modification can be a curious candidate.

References

[1] J. BOURGAIN, Some remarks on Banach spaces in which martingale difference sequences are unconditional, *Ark. Mat.* **21(2)** (1983), 163–168.

[2] J. BOURGAIN, Vector-valued singular integrals and the H^1-BMO duality, in: *"Probability theory and harmonic analysis"* (Cleveland, Ohio, 1983), Monogr. Textbooks Pure Appl. Math. **98**, Dekker, New York, 1986, pp. 1–19.

[3] D. L. BURKHOLDER, A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions, in: *"Conference on harmonic analysis in honor of Antoni Zygmund"*, Vol. I, II (Chicago, Ill., 1981), Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983, pp. 270–286.

[4] G. DAVID, Opérateurs intégraux singuliers sur certaines courbes du plan complexe, *Ann. Sci. École Norm. Sup. (4)* **17(1)** (1984), 157–189.

[5] J. B. GARNETT, *"Bounded analytic functions"*, Pure and Applied Mathematics **96**, Academic Press, Inc., New York-London, 1981.

[6] T. A. GILLESPIE, S. POTT, S. TREIL AND A. VOLBERG, Logarithmic growth for matrix martingale transforms, *J. London Math. Soc. (2)* **64(3)** (2001), 624–636.

[7] T. A. GILLESPIE, S. POTT, S. TREIL AND A. VOLBERG, The transfer method in estimates for vector Hankel operators, (Russian), *Algebra i Analiz* **12(6)** (2000), 178–193; translation in *St. Petersburg Math. J.* **12(6)** (2001), 1013–1024.

[8] N. H. KATZ, Matrix valued paraproducts, in: *"Proceedings of the conference dedicated to Professor Miguel de Guzmán"* (El Escorial, 1996), *J. Fourier Anal. Appl.* **3**, Special Issue (1997), 913–921.

[9] M. T. LACEY AND C. M. THIELE, L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$, *Ann. of Math. (2)* **146(3)** (1997), 693–724.

[10] F. LUST-PIQUARD, Opérateurs de Hankel 1-sommant de $l^1(\mathbb{N})$ dans $l^\infty(\mathbb{N})$ et multiplicateurs de $H^1(\mathbb{T})$, *C. R. Acad. Sci. Paris Sér. I Math.* **299(18)** (1984), 915–918.

[11] P. Mattila, M. S. Melnikov and J. Verdera, The Cauchy integral, analytic capacity, and uniform rectifiability, *Ann. of Math. (2)* **144(1)** (1996), 127–136.

[12] M. S. Melnikov and J. Verdera, A geometric proof of the L^2 boundedness of the Cauchy integral on Lipschitz graphs, *Internat. Math. Res. Notices* **1995(7)** (1995), 325–331.

[13] F. Nazarov, S. Treil and A. Volberg, Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces, *Internat. Math. Res. Notices* **1997(15)** (1997), 703–726.

[14] F. Nazarov, S. Treil and A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces, *Internat. Math. Res. Notices* **1998(9)** (1998), 463–487.

[15] F. Nazarov, S. Treil and A. Volberg, Accretive system *Tb* theorem of M. Christ for non-homogeneous spaces, *Duke Math. J.* (to appear).

[16] F. Nazarov, S. Treil and A. Volberg, Counterexample to the infinite-dimensional Carleson embedding theorem, *C. R. Acad. Sci. Paris Sér. I Math.* **325(4)** (1997), 383–388.

[17] F. Nazarov, S. Treil and A. Volberg, The Bellman functions and two-weight inequalities for Haar multipliers, *J. Amer. Math. Soc.* **12(4)** (1999), 909–928.

[18] S. Petermichl, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol, *C. R. Acad. Sci. Paris Sér. I Math.* **330(6)** (2000), 455–460.

[19] X. Tolsa, Cotlar's inequality without the doubling condition and existence of principal values for the Cauchy integral of measures, *J. Reine angew. Math.* **502** (1998), 199–235.

[20] X. Tolsa, L^2-boundedness of the Cauchy integral operator for continuous measures, *Duke Math. J.* **98(2)** (1999), 269–304.

[21] X. Tolsa, Curvature of measures, Cauchy singular integral, and analytic capacity, Thesis, Department of Mathematics, Universitat Autònoma de Barcelona (1998).

[22] S. Treil and A. Volberg, Wavelets and the angle between past and future, *J. Funct. Anal.* **143(2)** (1997), 269–308.

[23] S. Treil and A. Volberg, Completely regular multivariate stationary processes and the Muckenhoupt condition, *Pacific J. Math.* **190(2)** (1999), 361–382.

[24] A. Volberg, A_p weights via S-functions, *J. Amer. Math. Soc.* **10(2)** (1997), 445–466.

S. Petermichl:
Institute for Advanced Studies
Princeton, NJ 08540
U.S.A.
E-mail address: `stefanie@math.msu.edu`

S. Treil:
Department of Mathematics
Brown University
Providence, RI
U.S.A.
E-mail address: `treil@math.brown.edu`

A. Volberg:
Department of Mathematics
Michigan State University
East Lansing, Michigan 48824
U.S.A.
and
Equipe d'Analyse
Université Paris VI
4, place Jussieu
75252 Paris Cedex 05
France
E-mail address: `volberg@math.msu.edu`