Lecture notes on the mini-course *Martingales and Fourier multipliers. What's new in this old marriage?* 3rd lecture

Oliver Dragičević

Albuquerque, October 12, 2007

This exposition is based on several recent results involving the Ahlfors-Beurling operator T. It is defined, for a test function $f:\mathbb{C}\to\mathbb{C}$, as

$$Tf(z) = -\frac{1}{\pi}\,\text{p.v.}\int_{\mathbb{C}}\frac{f(\zeta)}{(z-\zeta)^2}\,d\zeta\,.$$

From the very onset we will also need the Riesz transforms R_j. For $n\in\mathbb{N}$ and $j\in\{1,\dots,n\}$, set

$$\widehat{R_j f}(\xi) = i\frac{\xi_j}{|\xi|}\hat{f}(\xi)\,;\qquad \xi\in\mathbb{R}^n\,.$$

We will consider various results due to F. Nazarov, S. Petermichl, A. Volberg and the above author concerning estimates of

a) $\|T\|_{L^p(\mathbb{C})}$

b) $\|T\|_{L^p(\mathbb{C},w)}$ with w from the Muckenhoupt class A_p

c) $\|T^n\|_{L^p(\mathbb{C})}$.

There will be two main methods utilized in our proofs:

1.) Bellman function

2.) averaging (of martingale transforms).

Both methods have one thing in common- Burkholder's sharp estimate for differentially subordinated martingales.

Motivation

Let $\|\mu\|_\infty < 1$. Consider the *Beltrami equation* on the plane

$$\bar{\partial} f = \mu \cdot \partial f \,.$$

Known:

- $f \in W^{1,2}_{loc} \wedge () \Rightarrow f \in W^{1,p}_{loc}$ for $p > 2$, $p \approx 2$
- $f \in W^{1,q}_{loc}$ for $q < 2$, $q \approx 2$, $\wedge$ $() \Rightarrow f \in W^{1,2}_{loc}$.

A question was asked (Bojarski, 1957) about the best p, q for which the above implications still hold. It was subsequently conjectured (Gehring, Reich 1966) that $p = 1 + k^{-1}$ and $q = 1 + k$. The conjecture was confirmed by Astala (1994). Prior to the proof of the Gehring-Reich conjecture it was noticed that it would immediately follow if we knew that the $\|T\|_p = p - 1$ for $p \geqslant 2$. This was in turn conjectured by T. Iwaniec (1982).

Bellman function

In the continuation we will assume that $1 < p < \infty$, q is its conjugate coefficient, and p^* is the maximum of p and q.

Theorem 1 (Nazarov, Volberg 2000). *Let*

$$\Omega = \{(\zeta, \eta, Z, H) \in \mathbb{C} \times \mathbb{C} \times \mathbb{R} \times \mathbb{R}\,;\, |\zeta|^p < Z, |\eta|^q < H\}\,.$$

Then $\exists\; B : \Omega \to \mathbb{R}$*, such that*

(i) $0 \leqslant B(\zeta, \eta, Z, H) \leqslant (p^* - 1) Z^{1/p} H^{1/q}$

(ii) $\langle -d^2 B(v)\, dv, dv\rangle \geqslant 2|d\zeta||d\eta|\,;\;\; v = (\zeta, \eta, Z, H),\; dv = (d\zeta, d\eta, dZ, dH).$

Strictly speaking, the function need not be differentiable, but we may achieve that (arbitrarily close to the boundary of Ω) by convolving it with the standard mollifiers. The first property changes just marginally. We nevertheless state the result as above, for the sake of clarity.

We present the main ingredients of the proof, since it is this result that connects the Burkholder's theorem and singular integrals in our approach.

Sketch of the proof. Take any interval $J \subseteq \mathbb{R}^2$. Let J_- and J_+ be its left and right half, respectively. By h_J denote the Haar function associated to J, i.e. the function which is identically $-c$ and $+c$ on J_- and J_+, respectively, the constant $c > 0$ being chosen so as to make h_J normalized in L^2. Let $\sigma = \{\sigma_I\,; I^{\text{dyad}} \subseteq J\,,\ |\sigma_I| \leqslant 1\}$ and let T_σ be a martingale transform on J, i.e. if $f = \sum_{I^{\text{dyad}} \subseteq J} \langle f, h_I\rangle h_I$, then $T_\sigma f = \sum_{I^{\text{dyad}} \subseteq J} \sigma_I \langle f, h_I\rangle h_I$.

Burkholder's crucial result says that

$$\sup_\sigma \|T_\sigma\|_{L^p} = p^* - 1\,.$$

This answered - affirmatively - a question raised by Pełczyński. One way of writing (a part of) this result is the following. Take two test functions f, g. Then

$$|\langle T_\sigma f, g\rangle| \leqslant (p^*-1)\|f\|_p \|g\|_q = (p^*-1)\langle |f|^p\rangle_J^{1/p} \langle |g|^q\rangle_J^{1/q} |J|$$

Here $\langle \phi\rangle_K = \frac{1}{|K|}\int_K \phi$.

On the other hand

$$|\langle T_\sigma f, g\rangle| = \Big|\sum \sigma_I \langle f, h_I\rangle \overline{\langle f, h_I\rangle}\Big|\,.$$

For the appropriate choice of σ_I this is the same as

$$\sum |\langle f, h_I\rangle \overline{\langle g, h_I\rangle}|\,.$$

Consequently

$$\frac{1}{4|J|}\sum_{I^{\text{dyad}} \subseteq J} |\langle f\rangle_{I_+} - \langle f\rangle_{I_-}||\langle g\rangle_{I_+} - \langle g\rangle_{I_-}||I| \leqslant (p^*-1)\langle |f|^p\rangle_J^{1/p} \langle |g|^q\rangle_J^{1/q}\,.$$

If you consider the left-hand side of the inequality above and take the supremum over all f, g such that $\langle f\rangle_J = \zeta$, $\langle g\rangle_J = \eta$, $\langle |f|^p\rangle_J = Z$, $\langle |g|^q\rangle_J = H$, you get, by definition, $B(\zeta, \eta, Z, H)$. The property (i) follows immediately. As for (ii), first we note that the invariance of B under scaling implies

$$B\Big(\frac{a_+ + a_-}{2}\Big) - \frac{B(a_+) + B(a_-)}{2} \geqslant \Big|\frac{\zeta_+ - \zeta_-}{2}\Big| \cdot \Big|\frac{\eta_+ - \eta_-}{2}\Big|$$

for any $a_\pm = (\zeta_\pm, \eta_\pm, Z_\pm, H_\pm) \in \overline{\Omega}$. This is a dyadic analogue of (ii). To get the non-dyadic inequality take convolutions with smoothing kernels. $\square$

Heat extensions

Let $\widetilde{\varphi}(x,t) = (h_t * \varphi)(x)$ (the convolution being taken with respect to the measure $\frac{dm}{2\pi}$), where

$$h_t(x) = \frac{1}{2t}\, e^{-\frac{\|x\|^2}{4t}}\,, \qquad x \in \mathbb{R}^2\,.$$

Such function solves the *heat equation*

$$\Delta u = \frac{\partial u}{\partial t}$$

in $\mathbb{R}^3_+$, whereas on the boundary $\mathbb{R}^2 \times \{0\}$ it coincides with φ.

Then, for $f, g \in C_0^\infty(\mathbb{R}^2)$ and $k, j \in \{1,2\}$,

$$\iint_{\mathbb{R}^2} R_k R_j f(y)\, g(y)\, dy = -2 \int_0^\infty \int_{\mathbb{R}^2} \frac{\partial \widetilde{f}}{\partial x_k}(y,t) \frac{\partial \widetilde{g}}{\partial x_j}(y,t)\, dy\, dt\,. \tag{1}$$

This can be proven by utilizing the Fourier transform.

Next we put these concepts together. Define, for $(x,t) \in \mathbb{R}^2 \times (0,\infty)$,

$$v = v(x,t) = (\widetilde{f}(x,t), \widetilde{g}(x,t), \widetilde{|f|^p}(x,t), \widetilde{|g|^q}(x,t))$$

and

$$b = B \circ v\,.$$

The idea is to estimate

$$\int_{\mathbb{R}^2 \times (0,\infty)} \Big(\frac{\partial}{\partial t} - \Delta\Big) b$$

from below and above.

The upper estimate is $\leqslant (p^* - 1)\|f\|_p \|g\|_q$ (it follows by combining property (i) and an argument involving the Green function).

For the lower estimate first note that

$$\left(\frac{\partial}{\partial t} - \Delta\right) b(x,t) = \sum_{j=1,2} \langle -d^2 B(v) \frac{\partial v}{\partial x_j}, \frac{\partial v}{\partial x_j}\rangle\,.$$

Indeed, for every function v we have

$$\left(\frac{\partial}{\partial t} - \Delta\right) b(x,t) = \langle \nabla B(v), \big(\partial/\partial t - \Delta\big) v\rangle - \sum_{j=1,2} \langle d^2 B(v) \frac{\partial v}{\partial x_j}, \frac{\partial v}{\partial x_j}\rangle\,.$$

But if components of v solve the heat equation, then the first term on the right disappears.

Thus, together with (ii),

$$\left(\frac{\partial}{\partial t}-\Delta\right) b(x,t) \geqslant 2 \sum_{j=1,2} \left|\frac{\partial \widetilde{f}}{\partial x_j}\right|\left|\frac{\partial \widetilde{g}}{\partial x_j}\right| \tag{2}$$

and so

$$|\langle (R_2^2-R_1^2)f,g\rangle| \leqslant 2\int \sum_{j=1,2} \left|\frac{\partial \widetilde{f}}{\partial x_j}\right|\left|\frac{\partial \widetilde{g}}{\partial x_j}\right| \leqslant (p^*-1)\|f\|_p\|g\|_q\,.$$

The same for $2R_1R_2$, of course. Since $T=(R_2^2-R_1^2)+2iR_1R_2$, this proves that $\|T\|_p \leqslant 2(p^*-1)$ (Nazarov, Volberg 2000).

This result was subsequently improved, in the asymptotical sense.

Theorem 2 (Volberg, D. 2002).

$$\limsup_{p\to\infty}\frac{\|T\|_p}{p-1} \leqslant \sqrt{2} \qquad \text{and} \qquad \limsup_{p\to\infty}\frac{\|T\|_{L^p_{real}\to L^p}}{p-1} \leqslant 1\,.$$

In order to prove it we need a certain Littlewood-Paley-type inequality.

Theorem 3.

$$2\int_0^\infty\int_{\mathbb{R}^2}\Big(\Big|\frac{\partial \tilde{f}}{\partial x_1}\Big|^2+\Big|\frac{\partial \tilde{f}}{\partial x_2}\Big|^2\Big)^{1/2}\Big(\Big|\frac{\partial \tilde{g}}{\partial x_1}\Big|^2+\Big|\frac{\partial \tilde{g}}{\partial x_2}\Big|^2\Big)^{1/2} dy\,dt \leqslant (p^*-1)\|f\|_p\|g\|_q\,.$$

For obtaining this result we need, in turn, another one.

Lemma 1. *Let $\mathcal{H}$ be a finite-dimensional real Euclidean space, $\mathcal{H}_{1,2}$, are two mutually orthogonal subspaces of $\mathcal{H}$ and $P_{1,2}$ are the corresponding orthogonal projections. Let Q be a self-adjoint operator on $\mathcal{H}$ such that*

$$\langle Qh,h\rangle \geqslant 2\|P_1h\|\|P_2h\|$$

for all $h\in\mathcal{H}$. Then there exists $\tau=\tau(Q)>0$, satisfying

$$\langle Qh,h\rangle \geqslant \tau\|P_1h\|^2+\frac{1}{\tau}\,\|P_2h\|^2$$

for all $h\in\mathcal{H}$.

Although this result suffices for our needs, we quote its generalization due to S. Treil, since it can be stated in a beautiful and compact way, covering the "original" result as a mere special case. Also, the proof given by Treil was shorter.

Theorem 4 (S. Treil 2003). *Let V be a vector space and σ_0, σ_1, σ_2 non-negative, non-zero, quadratic forms on V. Assume*

$$\sigma_0 \geqslant 2\sqrt{\sigma_1\sigma_2}$$

pointwise on V. Then there is $\tau > 0$ such that

$$\sigma_0 \geqslant \tau\sigma_1^2 + \tau^{-1}\sigma_2^2 \,.$$

In order to prove Theorem 3 apply Lemma 1 and simply "correct" (2) as

$$\left(\frac{\partial}{\partial t} - \Delta\right) b(x,t) \geqslant \sum_{j=1,2} \left(\tau\left|\frac{\partial\widetilde{f}}{\partial x_j}\right|^2 + \tau^{-1}\left|\frac{\partial\widetilde{g}}{\partial x_j}\right|^2\right).$$

Factor out τ and τ^{-1} whereever possible and apply the inequality between the arithmetic and geometric mean. The coefficient τ disappears and one is left with the integrand from Theorem 3. □

Lemma 2. *Let $T_\vartheta := (R_1^2 - R_2^2)\cos\vartheta + 2R_1R_2\sin\vartheta$. Then $\|T_\vartheta\|_p \leqslant p^* - 1$.*

Proof. By (1) we can write

$$\langle T_\vartheta f, g\rangle = -2\int_0^\infty \int_{\mathbb{R}^2} \left\langle \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix}, \begin{pmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \end{pmatrix} \right\rangle_{\mathbb{C}^2} dx\, dt$$

The desired inequality follows from Theorem 3 and the Cauchy-Schwarz inequality. □

Here is the last step towards the proof of Theorem 2.

Lemma 3. *Let μ be a positive measure on space X and let A, B be operators acting on $L^p_{real}(X,\mu)$. Denote*

$$C(p) := \max_{\vartheta\in[0,2\pi)} \|A\cos\vartheta + B\sin\vartheta\|_p \,.$$

Then

$$\left\|\begin{pmatrix} A \\ B \end{pmatrix} : L^p_{real} \to L^p(l^2_{real})\right\| \leqslant \tau_p C(p), \quad \text{where} \quad \lim_{p\to\infty} \tau_p = 1 \,.$$

Proof. Take $f \in L^p_{real}$, $x \in X$ and write temporarily $a = Af(x)$, $b = Bf(x)$. Since a and b are real (by the assumption), there is $\delta = \delta(x) \in [0, 2\pi)$ such that

$$(a, b) = \sqrt{a^2 + b^2}\,(\cos\delta,\, \sin\delta)\,.$$

It follows that

$$Af(x)\cos\vartheta + Bf(x)\sin\vartheta = \sqrt{|Af|^2 + |Bf|^2}(x)\cos(\vartheta - \delta(x))$$

for all $\vartheta \in [0, 2\pi)$. Consequently,

$$\int_X \sqrt{|Af(x)|^2 + |Bf(x)|^2}^{\,p}\, |\cos(\vartheta - \delta(x))|^p\, d\mu(x) \leqslant C(p)^p \|f\|_p^p\,.$$

Integrate this inequality with respect to the normalized Lebesgue measure $(2\pi)^{-1} d\vartheta$ on $[0, 2\pi)$. By denoting

$$\tau_p := \|\cos\|^{-1}_{L^p(0,2\pi)}$$

we get what we wanted. □

Finally, apply this lemma with $A = \Re T = R_2^2 - R_1^2$ and $B = \Im T = 2R_1R_2$ and use Lemma 2. The second part of Theorem 2 is proved. The proof of the first part is a bit more technical, but follows essentially the same path. □

Weighted estimates; the averaging

Astala, Iwaniec and Saksman (2001) showed that the (lower) boundary case of the Gehring-Reich conjecture reduces to finding sharp estimates of $\|T\|_{B(L^p(w))}$, where the weight w belongs to the so-called Muckenhoupt class. It is defined as follows.

Recall that by

$$\langle f\rangle_Q = \frac{1}{|Q|}\int_Q f\, dm$$

we denoted the average of a function f over the set Q. Typically, Q's will be squares in the plane. A positive measurable function w is said to belong to the A_p class if the quantity

$$Q_{w,p} := \sup_{Q \subset \mathbb{R}^2} \langle w\rangle_Q \big\langle w^{-\frac{1}{p-1}}\big\rangle_Q^{p-1}$$

is finite. The supremum is taken over squares in $\mathbb{R}^2$. In particular,

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

Theorem 5. *There is $C > 0$ such that for all $w \in A_2$,*

$$\|T\|_{B(L^2(w))} \leqslant C \cdot Q_{w,2}$$

and this is sharp.

One can extend this result by the use of the extrapolation technique.

Corollary 1.

$$\|T\|_{B(L^p(w))} \leqslant C(p) Q_{w,p} \,, \qquad p \geqslant 2 \,.$$

The first proof of the Theorem 5, due to Petermichl and Volberg (2002) utilized a Bellman function argument. We describe contours of a subsequent proof (Volberg, D. 2003) which can be viewed as a basic (though not first) example of the averaging technique.

The following result is due to J. Wittwer (2000). It addresses the sharp behaviour of L^p-norms of martingale transforms on the weighted space.

Theorem 6.

$$\|T_\sigma\|_{B(L^2(w))} \leqslant C \cdot Q_{w,2} \,.$$

Since this estimate is exactly of the kind we wanted for T, one might ask how to pass from T_σ to T without changing the norm too much. First we consider two-dimensional martingale transforms.

The term *dyadic lattice* and the symbol $\mathfrak{L}$ 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. The main difference with the one-dimensional case is that on $\mathbb{R}^2$ we assign to each square Q three Haar functions:

$$\begin{aligned} h_Q^1(s,t) &= \chi_I(s) h_J(t) |I|^{-1/2} \\ h_Q^2(s,t) &= h_I(s) \chi_J(t) |J|^{-1/2} \\ h_Q^3(s,t) &= h_I(s) h_J(t) \,. \end{aligned}$$

Symbolically,

$$h_Q^1 \equiv \begin{array}{|c|}\hline + \\ \hline - \\ \hline \end{array} \qquad h_Q^2 \equiv \begin{array}{|c|c|}\hline - & + \\ \hline \end{array} \qquad h_Q^3 \equiv \begin{array}{|c|c|}\hline - & + \\ \hline + & - \\ \hline \end{array}$$

Thus we can define the two-dimensional martingale transform.

$$T_\sigma f := \sum_{Q\in\mathcal{D}} \sum_{k=1}^{3} \sigma_Q^k \langle f, h_Q^k\rangle\, h_Q^k$$

As before, $|\sigma_Q^k| \leqslant 1$. It turns out that the Wittwer's proof (also heavily involving Bellman functions!) still works and gives the same sharp (linear) estimate. So now the right question is about passage to the Ahlfors-Beurling operator:

$$T_\sigma \overset{?}{\longrightarrow} T$$

Theorem 7 (Volberg, D. 2003).

$$T = C \cdot \text{“\textit{average of 2-dim. mart. transforms}”},$$

where

$$C = \frac{12\log 2}{16\pi + 32\log 2 - 15\log 5 - 40\arctan 2} \approx 2{\cdot}07\,.$$

It was noted earlier that Burkholder proved that all martingale transforms on the line have their norm majorized by $p^* - 1$. Thus one is tempted to apply Burkholder's theorem and obtain the estimate $\|T\|_p \leqslant 2{\cdot}07(p^* - 1)$.

There are, however, two problems:

i) The constant has to be decreased. (At the time when this work was done, the best known constant was 2, i.e. just slightly better than 2·07. Nowadays it is, as far as this author knows, 1·575. It was obtained by R. Bañuelos and P. Janakiraman.)

ii) Burkholder's proof cannot be applied in a direct way to higher-dimensional martingale transforms. Roughly speaking, the reason is that if we consider a $\sigma-$algebra generated by squares of a fixed size, then there is only one $\sigma-$algebra "between" this one and the one generated by squares of the next smaller size. So we have two types of $\sigma-$algebras, but three different Haar functions. This means that partial sums of T_σ's form either martingales which do not satisfy the differential subordination condition or else differentially subordinated stochastic processes which, however, are not martingales (even weak or very weak ones). In any case,

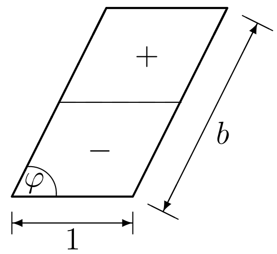


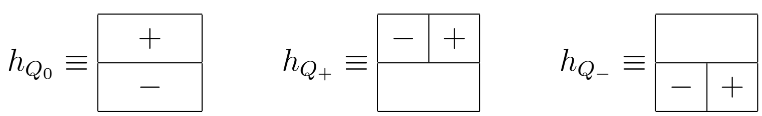
Figure 1: General h_{Q_0}

Burkholder's lemma cannot be applied. We have one type of functions too many. Of course, we could split these types (i.e. our operators) in two parts, estimate each separately and use the triangle inequality, but the constant would worsen by a factor of 2.

A way of getting around the second problem was suggested by Guy David. It simply says that one should consider a different family of functions as our "Haar" basis. This is the family:

$$\begin{aligned} h_{Q_0} &:= h_Q^1 \\ h_{Q_+} &:= \frac{1}{\sqrt{2}}(h_Q^2 + h_Q^3) \\ h_{Q_-} &:= \frac{1}{\sqrt{2}}(h_Q^2 - h_Q^3)\,. \end{aligned} \tag{3}$$

Symbolically,



The big advantage of this system over the "old" one is that two of the three functions have disjoint supports, thus we may take them simultaneously when building our stochastic process. This is how we get rid of the "excessive" step. It is immediate that martingale transforms associated with this system also admit $p^* - 1$−estimates. Another very useful fact is that when performing the averaging (i.e. repeating Theorem 7 with this new system), the constant does not change, owing to some unexpected but convenient cancellation.

It is left to decrease the constant of 2·07. The idea is to induce further changes in the Haar systems, but this time to keep the same structure of functions and change their supports geometrically. More precisely, instead of squares we consider parallelograms, as deipcted in Figure 1.

The minimum (of all constants) over all such parallelograms is attained at $\varphi = \pi/2$ and $b = \sqrt{2}$. Note that this is the only such object which preserves ratio of sides as well as inclination when cut in half. In this case

$$
\begin{aligned}
C &= \frac{2\log 2}{\frac{1}{\sqrt{2}}\left(\frac{5\pi}{6} + 9\arctan\frac{1}{\sqrt{2}} - 7\arctan\sqrt{2}\right) + \frac{8}{3}\log 2 - 2\log 3} \\
&\approx 2,00714\,.
\end{aligned}
$$

It turns out that the optimal coefficients are $\sigma_0 = -1$, $\sigma_+ = \sigma_- = 1$. This effectively means that we are not really applying the full Burkholder's result. Instead we only need L^p estimate of one single concrete martingale transform, namely $-P_0 + P_+ + P_-$, where $P_* :=$ is the projection onto $\overline{\mathrm{Lin}}\,\{h_{Q_*}\,; Q \in \mathcal{D}\}$. We do not know whether the norm of this operator is (significantly) better than $p^* - 1$.

There were other attempts (triangles, L-shaped objects, ...), but the constant 2˙007 remained undiminuished. However, this approach eventually did produce new results.

Estimates for T^n, $n \in \mathbb{Z}$

Motivation: theory of quasiregular mappings in higher dimensions. T. Iwaniec in G. Martin (1996) reduced the estimates of arising singular operators (on spaces of arbitrary dimensions) to the L^p-estimates of integer powers of a "twodimensional" operator - $\sqrt{T}$. The latter was introduced by Vekua as an operator with an odd kernel:

$$
-\frac{1}{2\pi}\,\frac{e^{-i\varphi}}{r^2}\,.
$$

Of course, even powers of $\sqrt{T}$ are powers of T.

Theorem 8 (Petermichl, Volberg, D. 2004). *For every $n \in \mathbb{Z}$ we have*

$$
T^n = C(n) \cdot T'
$$

where T' is a result of a certain averaging process, adjusted to T^n. We can estimate $|C(n)| \leqslant Cn$.

The proof of this theorem is much like the one of Theorem 7. One needs to be careful when calculating the kernel of the operator resulting from the averaging, since this is where the information about the constant appears. It turns out one really needs rectangles to conclude about the linear behaviour; squares alone would not do.

Consequently, $\|T^n\|_p \leqslant Cn(p^* - 1)$. For large n,

$$C \approx e\,. \tag{4}$$

The growth of norms is sharp in p and also in n *for large* p. But while this is the best estimate of the form $C(n)D(p)$ that one can get, it is not optimal, since T is an isometry on L^2. A candidate for the optimal behaviour of the norms $\|T^n\|_p$ emerged from the lower estimates:

$$\|T^n\|_p \geqslant \kappa_n(p^*)\,, \tag{5}$$

where

$$\kappa_n(p) = \prod_{k=0}^{n-1} \frac{k - 1/p + 1}{k + 1/p}\,.$$

We believe that actually $\|T^n\|_p = \kappa_n(p^*)$. One can show that, for $p \geqslant 2$,

$$0{\cdot}964 \leqslant \frac{\kappa_n(p)}{n^{1-2/p}(p-1)} \leqslant 1\,.$$

The example which gives (5) is a somewhat generalized example of Lehto (1965) which gave $\|T\|_p \geqslant p^* - 1$. See also the expository article by Baernstein and Montgomery-Smith.

It turns out that the lower estimate is the optimal one. In other words,

$$\|T^n\|_p \sim n^{1-2/p^*}(p^* - 1)\,.$$

In order to prove this we used (following the suggestion by M. Christ) a powerful result on weak-type estimates of singular integrals with rough kernels.

Theorem 9 (Christ, Rubio de Francia; Hofmann 1988). *Let* $\Omega \in L^q(S^1)$ *for some* $q > 1$, $\int_{S^1} \Omega = 0$. *If* Ω *is also homogeneous of degree* 0, *then*

$$Tf(x) = \text{p.v.} \int f(x-y) \frac{\Omega(y)}{|y|^2}\, dy$$

defines an operator which is of weak type $(1,1)$. *Its weak norm depends linearly on* $\|\Omega\|_q$.

By applying the Riesz-Thorin and Marcinkiewicz theorems one obtains precisely

$$\|T^n\|_p \leqslant C n^{1-2/p^*}(p^* - 1)\,. \tag{6}$$

A small advantage of the averaging approach is that it gives some information about the constant, see (4). We cannot trace what the constant in (6), arising from the proofs by Christ, Rubio de Francia and Hofmann, would be.