Haar multipliers
meet Bellman functions

María Cristina Pereyra

Dedicado a la memoria de mi maestro y amigo Mischa Cotlar

Abstract

Using Bellman function techniques, we obtain the optimal dependence of the operator norms in $L^2(\mathbb{R})$ of the Haar multipliers T_w^t on the corresponding RH_2^d or A_2^d characteristic of the weight w, for $t = 1, \pm 1/2$. These results can be viewed as particular cases of estimates on homogeneous spaces $L^2(v d\sigma)$, for σ a doubling positive measure and $v \in A_2^d(d\sigma)$, of the weighted dyadic square function S_σ^d. We show that the operator norms of such square functions in $L^2(v d\sigma)$ are bounded by a linear function of the $A_2^d(d\sigma)$ characteristic of the weight v, where the constant depends only on the doubling constant of the measure σ. We also show an inverse estimate for S_σ^d. Both results are known when $d\sigma = dx$. We deduce both estimates from an estimate for the Haar multiplier $(T_v^\sigma)^{1/2}$ on $L^2(d\sigma)$ when $v \in A_2^d(d\sigma)$, which mirrors the estimate for $T_w^{1/2}$ in $L^2(\mathbb{R})$ when $w \in A_2^d$. The estimate for the Haar multiplier adapted to the σ measure, $(T_v^\sigma)^{1/2}$, is proved using Bellman functions. These estimates are sharp in the sense that the rates cannot be improved and be expected to hold for all σ, since the particular case $d\sigma = dx$, $v = w$, correspond to the estimates for the Haar multipliers $T_w^{1/2}$ proven to be sharp.

1. Introduction

The *Haar multipliers* considered in these paper, are operators of the form

$$T_w^t f(x) = \sum_{I \in \mathcal{D}} \left(\frac{w(x)}{m_I w} \right)^t \langle f, h_I \rangle h_I(x);$$

where $\mathcal{D}$ denotes the dyadic intervals; $\{h_I\}_{I\in\mathcal{D}}$ the Haar functions normalized in $L^2(\mathbb{R})$, i.e. $h_I(x) = |I|^{-1/2}\big(\chi_{I_r}(x) - \chi_{I_l}(x)\big)$, I_l, I_r the left and right halves of I; $\langle\cdot,\cdot\rangle$ denotes the inner product in $L^2(\mathbb{R})$; w is a weight, $m_I w$ denotes the average of w on the dyadic interval I, and $t \in \mathbb{R}$.

Necessary and sufficient conditions for boundedness of T_w^t on $L^p(\mathbb{R})$, $1 < p < \infty$, are known in most cases, see [28], [17], [29].

When $tp > 1$, these operators are bounded in $L^p(\mathbb{R})$ if and only if w satisfies the *dyadic Reverse Hölder q condition*, RH_q^d, where $q = tp$, namely, there exists a constant $C > 0$ such that for all $I \in \mathcal{D}$,

$$\left(\frac{1}{|I|}\int_I w^q\right)^{1/q} \leq \frac{C}{|I|}\int_I w.$$

The smallest constant on the right hand side is denoted by $[w]_{RH_q^d}$, and it is called the RH_q^d-*characteristic* of the weight w.

When $tp < 0$, these operators are bounded in $L^p(\mathbb{R})$ if and only if w satisfies the following *dyadic A_q^d condition*, where $q = 1 - \frac{1}{pt}$,

$$\sup_{I\in\mathcal{D}}\left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I w^{-\frac{1}{q-1}}\right)^{q-1} < \infty.$$

The left-hand-side is called the A_q^d-*characteristic* of the weight w, and is denoted $[w]_{A_q^d}$.

When $0 < tp \leq 1$, and if we assume that $w \in \bigcup_{p>1} RH_p^d$ then the corresponding operators are bounded in $L^p(\mathbb{R})$ for $1 < p < \infty$.

A weight w is *dyadic doubling* if

$$\sup_{I\in\mathcal{D}}\frac{w(\tilde{I})}{w(I)} < \infty,$$

where $\tilde{I}$ denotes the parent of I, and $w(I) = \int_I w(x)\,dx$. The left-hand-side is called the *dyadic doubling constant* of the weight w, and is denoted by $D(w)$.

In the classical non-dyadic theory, $w \in A_p$ implies doubling, and $\bigcup_{p>1} A_p = \bigcup_{p>1} RH_p$, see [7]. In the dyadic theory, $w \in A_p^d$ implies dyadic doubling, but $w \in RH_p^d$ does not. In this case, $\bigcup_{p>1} A_p^d$ is a strict subset of $\bigcup_{p>1} RH_p^d$, however if we consider the dyadic doubling weights that belong to $\bigcup_{p>1} RH_p^d$, then we recover $\bigcup_{p>1} A_p^d$, see [5], [17].

In this paper we are interested in studying the dependence of the L^p-bounds of T_w^t on the corresponding characteristic of the weight w, and sometimes also on the dyadic doubling constant of w. We will concentrate on the cases $p = 2$, and $t = 1, 1/2, -1/2$.

This work was inspired by a string of papers that have appeared in the wake of this millennium, where sharp linear bounds in $L^2(w)$ for classical operators (square function, martingale transform, Beurling transform, Hilbert transform, and Riesz transforms) on weighted Lebesgue spaces have been obtained, see [16], [37], [38], [35], [34], [10], [11], [31], [32]. Very recently the same linear bound has been proved to hold also for the dyadic paraproduct, see [3]. All these results use Bellman function techniques introduced by Nazarov, Treil and Volberg [24], [25]. In [25] necessary and sufficient conditions for two weighted estimates for the martingale transform (Haar multiplier with symbol ± 1) were found. These results were in turn used by J. Wittwer [37] who considered one weight estimates but noticed that the Bellman function method provided optimal (linear) estimates in terms of the A_2-characteristic of the weight. See [26] for a very lucid review on the connections between the original Bellman functions (solutions of the Bellman differential equation) in stochastic control theory and the Bellman functions in harmonic analysis. These sharp estimates are not just a mathematical curiosity, people use them in a variety of settings, for example in the theory of quasiconformal maps ([1], [2], [36], [34]), and when considering L^p-solvability of elliptic problems ([13], [9]).

We use Bellman function techniques to study the sharp dependence of the operator norm of the Haar multipliers $T_w^{\pm 1/2}$ and T_w^1 in $L^2(\mathbb{R})$ on $[w]_{A_2^d}$ and $[w]_{RH_2^d}$ respectively.

We will prove the following theorems:

Theorem 1. *Let $w \in A_2^d$ then there exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{R})$,*
$$\|T_w^{-1/2} f\|_{L^2(\mathbb{R})} \le C[w]_{A_2^d}\|f\|_{L^2(\mathbb{R})}.$$

Denote $T_w = T_w^1$.

Theorem 2. *Let $w \in RH_2^d$ and dyadic doubling, then there exists a constant $C > 0$ depending on the dyadic doubling constant of w such that for all $f \in L^2(\mathbb{R})$,*
$$\|T_w f\|_{L^2(\mathbb{R})} \le C[w]_{RH_2^d}^2\|f\|_{L^2(\mathbb{R})}.$$

The following theorem involves the case $t = 1/2$ and $w \in A_2^d$.

Theorem 3. *Let $w \in A_2^d$ then there exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{R})$,*
$$\|T_w^{1/2} f\|_{L^2(\mathbb{R})} \le C[w]_{A_2^d}^{1/2}\|f\|_{L^2(\mathbb{R})}.$$

The results are optimal in the sense that we cannot get a slower decaying function of the corresponding characteristics of the weights.

The proof of Theorem 3 can be found embedded in the proof of an inverse estimate for the square function that is not stated in the language of Haar multipliers [33]. What is interesting, is that we can deduce Theorem 1 as a corollary of Theorem 3. This is very much in line with the beautiful operator theory argument presented in [33].

The proof of the first two theorems will lead naturally to the study of weighted square functions on homogeneous spaces. In fact, the results for the weighted square functions will provide yet another proof of these results.

Verifying the boundedness of T_w^t in $L^2(\mathbb{R})$ with a particular bound is equivalent to verifying the boundedness of its adjoint $(T_w^t)^*$ in $L^2(\mathbb{R})$ with the same bound. The advantage of the adjoints is that it will be very easy to compute L^2-norms. In the first case, $t = -1/2$, the L^2-norm of $(T_w^{-1/2})^* f$ is equal to the $L^2(w)$-norm of the dyadic square function $S^d(w^{-1/2}f)$ defined below by (1.3) when $d\sigma = dx$. But it has been shown, by Hukovic, Treil and Volberg [16], that

$$(1.1) \qquad \|S^d f\|_{L^2(w)} \le C[w]_{A_2^d} \|f\|_{L^2(w)},$$

and this result is optimal. This leads to a quick proof of Theorem 1. But it could be used, and we will, in the other direction, if we can prove Theorem 1 independently of the square function estimate, then we will get (1.1), this idea was used in [33, Corollary 3.2].

It is well known, see [28], that if w is dyadic doubling, the boundedness of T_w in $L^2(\mathbb{R})$ is equivalent to the boundedness of the weighted dyadic square S_w^d function in $L^2(\mathbb{R})$, defined below by (1.3) when $d\sigma = wdx$, and either of these events happens if and only if $w \in RH_2^d$. It will be clear from the proof of Theorem 2 that

$$(1.2) \qquad \|S_w^d f\|_{L^2(\mathbb{R})} \le C[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})},$$

where C depends on the dyadic doubling constant of w.

We claim that these square functions estimates, (1.1) and (1.2), are of the same nature, to be explained subsequently. Let σ be a positive dyadic doubling[1] measure. We will say that $v \in A_2^d(d\sigma)$ if

$$\sup_{I \in \mathcal{D}} \left(\frac{1}{\sigma(I)} \int_I v \, d\sigma\right) \left(\frac{1}{\sigma(I)} \int_I v^{-1} \, d\sigma\right) < \infty.$$

The quantity on the left hand side is denoted by $[v]_{A_2^d(d\sigma)}$. When $d\sigma = dx$ we simply write $[v]_{A_2^d}$.

---

[1]A positive measure σ is called *dyadic doubling* if there exists $C > 0$ such that $\sigma(\tilde{I})/\sigma(I) \le C$, for all $I \in \mathcal{D}$, $\tilde{I}$ being I's parent, and $\sigma(I) = \int_I d\sigma$. Denote by $D(\sigma)$ the smallest such constant, which we will call the *dyadic doubling constant* of σ. Note that $D(\sigma) \ge 2$. Given a weight w let $d\sigma = wdx$, then σ is dyadic doubling if and only if the weight w is dyadic doubling, moreover $D(\sigma) = D(w)$.

Define the σ-dyadic square function by

$$(1.3) \qquad S_\sigma^d f(x) := \left(\sum_{j \in \mathbb{Z}} |\Delta_j^\sigma f(x)|^2 \right)^{1/2},$$

where the j-th σ-difference $\Delta_j^\sigma := E_{j+1}^\sigma - E_j^\sigma$, and the j-th σ-expectation is given by

$$E_j^\sigma f(x) := \frac{1}{\sigma(I)} \int_I f \, d\sigma =: m_I^\sigma f, \quad x \in I \in \mathcal{D}_j,$$

where $\mathcal{D}_j$ denotes the dyadic intervals of length 2^{-j}.

Define the σ-dyadic maximal function by

$$(1.4) \qquad M_\sigma^d f(x) := \sup_{j \in \mathbb{Z}} E_j^\sigma |f|(x).$$

When $d\sigma = dx$ we write E_j, Δ_j, S^d, and M^d and when $d\sigma = w dx$ we write E_j^w, Δ_j^w, S_w^d, and M_w^d. It is well known that M_σ^d and S_σ^d are bounded in $L^2(v d\sigma)$, if and only if $v \in A_2^d(d\sigma)$, see [7]. In this paper we prove the following estimate that generalizes (1.1),

Theorem 4. *Let σ be a positive dyadic doubling measure, and $v \in A_2(d\sigma)$ then there exists a constant C depending only on the dyadic doubling constant of σ such that for all $f \in L^2(v d\sigma)$,*

$$\|S_\sigma^d f\|_{L^2(v d\sigma)} \leq C[v]_{A_2^d(d\sigma)} \|f\|_{L^2(v d\sigma)}.$$

It is clear that (1.1) is a corollary of this result where $d\sigma = dx$ and $v = w$.

We claim that (1.2) is also a corollary of this result where this time we choose $d\sigma = w dx$ and $v = w^{-1}$. There is, apparently, a discrepancy in the nature of the constants, until one realizes the following tautology,

$$w \in RH_2^d \Leftrightarrow w^{-1} \in A_2^d(w dx), \quad \text{moreover} \quad [w^{-1}]_{A_2^d(w dx)} = [w]_{RH_2^d}^2.$$

It is worth mentioning the following optimal inverse result due to S. Petermichl and S. Pott [33]: assume $w \in A_2^d$, then

$$(1.5) \qquad \|f\|_{L^2(w)} \leq C[w]_{A_2^d}^{1/2} \|S^d f\|_{L^2(w)}.$$

Estimate (1.5) can be seen to be equivalent to Theorem 3, this observation can be traced back to [33]. For the Lusin square function a similar inverse estimate was known to R. Fefferman and J. Pipher (see comment at the bottom of page 359 in [14]). It is natural to conjecture that a similar lower bound to (1.5) can be found in the general case, that is the content of the next theorem.

Theorem 5. *Let σ be a positive doubling measure, and $v \in A_2^d(d\sigma)$, let $d\mu = vd\sigma$. There exists a constant C depending only on the dyadic doubling constant of σ, such that for all $f \in L^2(vd\sigma)$,*

$$\|f\|_{L^2(vd\sigma)} \le C[v]_{A_2^d(d\sigma)}^{1/2}\|S_\sigma^d f\|_{L^2(vd\sigma)}.$$

Furthermore one can deduce this result from an appropriate $d\sigma$-analogue of Theorem 3.

Theorem 6. *Let σ be a positive doubling measure, and $v \in A_2^d(d\sigma)$, then there exists a constant C depending only on the dyadic doubling constant of σ, such that for all $f \in L^2(d\sigma)$,*

$$\|(T_v^\sigma)^{1/2}f\|_{L^2(d\sigma)} \le C[v]_{A_2^d(d\sigma)}^{1/2}\|f\|_{L^2(d\sigma)}.$$

Where

$$(T_v^\sigma)^{1/2}f := \sum_{I \in \mathcal{D}} \left(\frac{v(x)}{m_I^\sigma v}\right)^{1/2} \langle f, h_I^\sigma\rangle_\sigma h_I^\sigma(x),$$

the functions $\{h_I^\sigma\}_{I \in \mathcal{D}}$ form an orthonormal basis in $L^2(d\sigma)$, called the weighted Haar system, and $\langle \cdot, \cdot\rangle_\sigma$ denotes the inner product in $L^2(d\sigma)$.

To prove Theorem 6 we follow the argument that Petermichl and Pott [33] used in the case $d\sigma = dx$. In fact we will deduce Theorem 4 as a corollary of Theorem 6. We can also deduce as a corollary the following result for $(T_v^\sigma)^{-1/2}$ that is analogous to the corresponding result stated in Theorem 1

Theorem 7. *Let σ be a positive doubling measure, and $v \in A_2^d(d\sigma)$ then there exists a constant C depending only on the dyadic doubling constant of σ, such that for all $f \in L^2(d\sigma)$,*

$$\|(T_v^\sigma)^{-1/2}f\|_{L^2(d\sigma)} \le C[v]_{A_2^d(d\sigma)}\|f\|_{L^2(d\sigma)}.$$

Where

$$(T_v^\sigma)^{-1/2}f := \sum_{I \in \mathcal{D}} \left(\frac{v(x)}{m_I^\sigma v}\right)^{-1/2} \langle f, h_I^\sigma\rangle_\sigma h_I^\sigma(x).$$

The paper is organized as follows. In Section 2 we prove Theorem 1 by showing that it is equivalent to the dyadic square function estimate on weighted spaces. In Section 3 we prove Theorem 2, by reducing the proof to proving a precise weighted Carleson estimate that is shown to hold in two steps: first a so called Sawyer's estimate that is handled using Bellman functions, and second a weighted estimate that jump-starts the Sawyer estimate, and which turns out to be trivial in this case. This argument works

for dyadic doubling weights and the numerical constant depends linearly on the dyadic doubling constant of w. We also prove similar estimates for the dyadic weighted maximal function M_w^d. In Section 4 we reduce Theorem 3 to proving the Petermichl-Pott estimate (1.5), in fact the two estimates are equivalent. We then deduce as corollaries of Theorem 3: Theorem 1, and the Hukovic-Treil-Volberg estimate (1.1). In Section 5, we set up the scene so that it is clear one can deduce from Theorem 6 (square root estimate for $(T_v^\sigma)^{1/2}$) the linear estimate for $(T_v^\sigma)^{-1/2}$ (Theorem 7), the linear estimate for S_σ^d (Theorem 4), and the inverse estimate for S_σ^d (Theorem 5), exactly in the same way as it was done in the case $d\sigma = dx$ in Section 4. In Section 6 we prove Theorem 6. The estimate is reduced to proving a weighted Carleson estimate. To achieve that, a Sawyer's estimate is needed, it turns out that the Bellman function required for this estimate is the same one used in the proof of Theorem 2, but this time the weight lemma, necessary to jump-start it, is not trivial and requires a proof, which we achieve using again Bellman functions. In Section 7 we explain how to prove Theorem 2 by applying a Bellman function argument directly to the adjoint problem for T_w, and we use some homogeneity considerations to show the sharpness of the estimate, by showing that no Bellman function with faster decay can exist. This proof also requires w to be dyadic doubling, not only in RH_2^d. We learned about this argument from Fedja Nazarov [23], the homogeneity trick that reduces the number of variables of the Bellman function one is searching is an idea that appeared first in [6]. In Section 8 we present some final remarks.

2. Sharp bound for $T_w^{-1/2}$

The formal adjoints of T_w^t are operators of the form

$$(2.1) \qquad (T_w^t)^* f(x) = \sum_{I \in \mathcal{D}} \frac{\langle fw^t, h_I \rangle}{(m_I w)^t} h_I(x),$$

hence one can compute the L^2-norm of the output by a direct application of Plancherel Theorem. More precisely,

$$(2.2) \qquad \|(T_w^t)^* f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} \frac{|\langle fw^t, h_I \rangle|^2}{(m_I w)^{2t}}.$$

Proof of Theorem 1. Set $t = -1/2$ in (2.2) to obtain,

$$\|(T_w^{-1/2})^* f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} m_I w |\langle w^{-1/2} f, h_I \rangle|^2.$$

Remember that the dyadic square function is given by

$$S^d g(x) = \left(\sum_{I \in \mathcal{D}} \frac{|\langle g, h_I \rangle|^2}{|I|} \chi_I(x) \right)^{1/2},$$

and its $L^2(w)$-norm can be calculated directly by

(2.3)
$$\|S^d g\|_{L^2(w)}^2 = \sum_{I \in \mathcal{D}} m_I w |\langle g, h_I \rangle|^2.$$

Hence,
$$\|(T_w^{-1/2})^* f\|_{L^2(\mathbb{R})}^2 = \|S^d(w^{-1/2} f)\|_{L^2(w)}^2.$$

Now we can use the sharp estimate (1.1) for $g = w^{-1/2} f$,

$$\begin{aligned}
\|S^d(w^{-1/2} f)\|_{L^2(w)} = \|S^d g\|_{L^2(w)} &\leq C[w]_{A_2^d} \|g\|_{L^2(w)} \\
&= C[w]_{A_2^d} \|w^{-1/2} f\|_{L^2(w)}.
\end{aligned}$$

Finally observing that $\|w^{-1/2} f\|_{L^2(w)} = \|f\|_{L^2(\mathbb{R})}$, we obtain the desired inequality for the adjoint operator and hence for the operator,

$$\|T_w^{-1/2} f\|_{L^2(\mathbb{R})} \leq C[w]_{A_2^d} \|f\|_{L^2(\mathbb{R})}.$$

This result must be sharp otherwise estimate (1.1) for the dyadic square function would not be sharp. In fact the two estimates are equivalent.   ∎

3. Bounds for T_w and S_w^d

In this section we assume that $w \in RH_2^d$ and that it is dyadic doubling. We denote by $D(w)$ the dyadic doubling constant of w. We prove that the operator norm of T_w in $L^2(\mathbb{R})$ is less than or equal than a constant that depends linearly on the doubling constant of w times $[w]_{RH_2^d}^2$, more precisely there exists $C > 0$ such that for all $f \in L^2(\mathbb{R})$,

(3.1)
$$\|T_w f\|_{L^2(\mathbb{R})} \leq C D(w) [w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})}.$$

Proof of Theorem 2. By (2.2) in the case $t = 1$ we obtain that

$$\|T_w^* f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} \left| \frac{\langle wf, h_I \rangle}{m_I w} \right|^2 = \sum_{I \in \mathcal{D}} |I| \left| \frac{m_I(fw)}{m_{\tilde{I}} w} - \frac{m_{\tilde{I}}(fw)}{m_{\tilde{I}} w} \right|^2.$$

Where $\tilde{I}$ is the parent of I.

The weighted dyadic square function is given by

$$S_w^d f(x) = \left(\sum_{I \in \mathcal{D}} \chi_I(x) \left| \frac{m_I(fw)}{m_I w} - \frac{m_{\tilde{I}}(fw)}{m_{\tilde{I}} w} \right|^2 \right)^{1/2}.$$

Notice that it coincides with $S_\sigma^d f(x)$ given by (1.3) in the case $d\sigma = wdx$. Its L^2-norm is given by,

$$(3.2) \qquad \|S_w^d f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} |I| \left| \frac{m_I(fw)}{m_I w} - \frac{m_{\tilde{I}}(fw)}{m_{\tilde{I}} w} \right|^2.$$

It is not hard to believe now, that T_w and S_w^d are bounded simultaneously in $L^2(\mathbb{R})$, and, for dyadic doubling weights, this occurs if and only if $w \in RH_2^d$. See [28] for details. We will sketch the argument so that it becomes obvious where do we need careful estimates to pin down the dependence on $[w]_{RH_2^d}$ and on $D(w)$. Adding and subtracting $\frac{m_I fw}{m_{\tilde{I}} w}$ inside the absolute value in the summands in (3.2), using that $(a+b)^2 \leq 2(a^2 + b^2)$, and that

$$\left(\frac{m_I fw}{m_I w} \right)^2 \leq D(w)^2 \left(\frac{m_{\tilde{I}} fw}{m_{\tilde{I}} w} \right)^2,$$

we can see that

$$\|S_w^d f\|_{L^2(\mathbb{R})}^2 \leq 2\|T_w^* f\|_{L^2(\mathbb{R})}^2 + D^2(w)Q(w, f),$$

where the Carleson Embedding term $Q(w, f)$, is given by

$$(3.3) \qquad Q(w, f) = \sum_{I \in \mathcal{D}} \frac{m_I^2(fw)}{m_I^2 w} \left(\frac{\langle w, h_I \rangle}{m_I w} \right)^2.$$

On the other hand, see [28, p. 654–656],

$$\|T_w^* f\|_{L^2(\mathbb{R})}^2 \leq \|M_w^d f\|_{L^2(\mathbb{R})}^2 + D^2(w)Q(w, f) + D(w)Q^{1/2}(w, f)\|T_w^* f\|_{L^2(\mathbb{R})},$$

where M_w^d is the weighted dyadic maximal function, i.e.

$$M_w^d f(x) := \sup_j E_j^w |f| (x), \quad \text{and} \quad E_j^w f(x) = \frac{m_I(fw)}{m_I w}, \quad x \in I \in \mathcal{D}_j.$$

We will check in Section 3.3 that

Lemma 1. *If $w \in RH_2^d$ then there exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{R})$,*

$$\|M_w^d f\|_{L^2(\mathbb{R})} \leq C[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})}.$$

All together we conclude that,

$$\|T_w^* f\|_{L^2(\mathbb{R})}^2 \leq C\left([w]_{RH_2^d}^4 \|f\|_{L^2(\mathbb{R})}^2 + D^2(w)Q(w,f)\right.$$

(3.4)
$$\left. + D(w)Q^{1/2}(w,f)\|T_w^* f\|_{L^2(\mathbb{R})}\right).$$

The Carleson Embedding term $Q(w,f)$ is bounded by a constant depending on w times $\|f\|_{L^2(\mathbb{R})}^2$. That is sufficient to ensure the boundedness of T_w and therefore of S_w^d in $L^2(\mathbb{R})$ by a bootstrapping argument. However if we get control on the constant in terms of the fourth power of the RH_2^d-characteristic of the weight w, then the bootstrapping will give the quadratic bound for T_w.

Lemma 2 (Carleson's Embedding Lemma). *Assume $w \in RH_2^d$, then for all $f \in L^2(\mathbb{R})$,*

$$Q(w,f) = \sum_{I \in \mathcal{D}} \frac{m_I^2(fw)}{m_I^2 w}\left(\frac{\langle w, h_I\rangle}{m_I w}\right)^2 \leq 4[w]_{RH_2^d}^4 \|f\|_{L^2(\mathbb{R})}^2.$$

Inserting this estimate in (3.4) we get,

$$\|T_w^* f\|_{L^2(\mathbb{R})}^2 \leq C\left(D^2(w)[w]_{RH_2^d}^4 \|f\|_{L^2(\mathbb{R})}^2\right.$$

$$\left. + D(w)[w]_{RH_2^d}^2 \|T_w^* f\|_{L^2(\mathbb{R})}\|f\|_{L^2(\mathbb{R})}\right).$$

Then bootstrapping[2] the above inequality we conclude that for some other constants $c, c' > 0$, independent of w,

$$\|T_w^* f\|_{L^2(\mathbb{R})} \leq cD(w)[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})},$$
$$\|S_w^d f\|_{L^2(\mathbb{R})} \leq c'D(w)[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})}.$$

This implies (3.1).   ■

In Section 7 we argue about the optimality of the quadratic bound on the RH_2^d-characteristic of the weight w. The argument presented there also features the appearance of the weight's doubling constant through the extension of the domain. At this point it is not completely clear to the author whether one could push these arguments without the appearance of the doubling constant. In particular, if the weight is in RH_2^d but is not dyadic

---

[2]All we are doing here is using the fact that if $A, B, C > 0$, and $A^2 \leq CB^2 + CAB$ then $A \leq cB$, for some $c > 0$ depending only on C. In our case, $A = \|T_w^* f\|_{L^2(\mathbb{R})}$, and $B = D(w)[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})}$.

doubling, what is the optimal dependence on $[w]_{RH_2^d}$? We know that T_w is bounded in that case, see [17] for a stopping-time proof (without keeping track of the nature of the constant) that works in the non-dyadic doubling case, unlike the original proof in [28] that assumes dyadic doubling, and which is the basis for the proof in this section.

We will prove the Carleson's Embedding (Lemma 2) in two steps, one is the so-called Sawyer's Estimate, the other is an appropriate weight estimate that turns out to be trivial in this case. The weight estimate is necessary to jump-start Sawyer's estimate.

Lemma 3 (Sawyer's Estimate). *Given w a weight, and $\{\lambda_I\}_{I \in \mathcal{D}}$ a sequence of positive numbers such that there is a constant $Q > 0$ such that for all $J \in \mathcal{D}$,*

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \frac{m_I^2(w^2)}{m_I^2 w} \lambda_I \leq Q m_J(w^2),$$

then for all dyadic intervals J,

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \frac{m_I^2(fw)}{m_I^2 w} \lambda_I \leq 4 Q m_J(f^2).$$

Proof of Lemma 2. Choosing $w \in RH_2^d$, and $\lambda_I = \left(\frac{\langle w, h_I \rangle}{m_I w}\right)^2$ we can jump-start Sawyer's Estimate with $Q = [w]_{RH_2^d}^4$. In fact, the following weight estimate holds,

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \frac{m_I^2(w^2)}{m_I^2 w} \left(\frac{\langle w, h_I \rangle}{m_I w}\right)^2 \leq [w]_{RH_2^d}^4 \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} |\langle w, h_I \rangle|^2$$

$$\leq [w]_{RH_2^d}^4 m_J(w^2),$$

where we are using the hypothesis $w \in RH_2^d$ to obtain the first inequality, and the fact that the collection of Haar functions $\{h_I\}_{I \in \mathcal{D}(J)}$ is an orthonormal set in $L^2(J)$ to obtain the last inequality.

Applying Sawyer's Estimate and letting J grow we conclude that

$$\sum_{I \in \mathcal{D}} \frac{m_I^2(fw)}{m_I^2 w} \left(\frac{\langle w, h_I \rangle}{m_I w}\right)^2 \leq 4[w]_{RH_2^d}^4 \|f\|_{L^2(\mathbb{R})}^2.$$

■

3.1. Sawyer's Estimate

We will prove Sawyer's Estimate using the technique of Bellman functions.

Lemma 4. *Suppose there exists a real-valued function of 5 variables, $B(\mathbf{s}) = B(x, y, w, v, M)$ whose domain D is given by those $\mathbf{s} = (x, y, w, v, M) \in \mathbb{R}^5$ such that*

$$x, y, w, v, M \geq 0,$$
$$M \leq v,$$
$$y^2 \leq xv;$$

whose range is given by $0 \leq B(\mathbf{s}) \leq x$, $\mathbf{s} \in D$, and such that the following convexity *property holds: for all $\mathbf{s}, \mathbf{s}_{\pm} \in D$ such that $\mathbf{s} - \frac{\mathbf{s}_+ + \mathbf{s}_-}{2} = (0, 0, 0, 0, \alpha)$ then*

(3.5)
$$B(\mathbf{s}) - \frac{B(\mathbf{s}_+) + B(\mathbf{s}_-)}{2} \geq \frac{1}{4} \left(\frac{y}{v} \right)^2 \alpha.$$

Then Sawyer's Estimate (Lemma 3) holds.

Proof. Without loss of generality can assume $f \geq 0$. Fix a dyadic interval J. Let $\mathbf{s_J} = (x_J, y_J, w_J, v_J, M_J)$, where $x_J = m_J(f^2)$, $y_J = m_J(fw)$, $w_J = m_J w$, $v_J = m_J(w^2)$, and

$$M_J = \frac{1}{Q} \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left(\frac{m_I(w^2)}{m_I w} \right)^2 \lambda_I.$$

Clearly for each $J \in \mathcal{D}$, $\mathbf{s_J}$ belongs to the domain, these are all positive quantities, $M_J \leq v_J$ is the hypothesis of Lemma 3, and $y_J^2 \leq x_J v_J$ is nothing more than Cauchy-Schwarz inequality. Let now $\mathbf{s}_{\pm} = \mathbf{s_{J_{\pm}}} \in D$. By definition, $\mathbf{s_J} - \frac{\mathbf{s_{J_+}} + \mathbf{s_{J_-}}}{2} = (0, 0, 0, 0, \alpha_J)$, where $\alpha_J = \frac{1}{Q|J|} \left(\frac{m_J(w^2)}{m_J w} \right)^2 \lambda_J$.

We are assuming a function B exists, such that $B(\mathbf{s_J}) \leq m_J(f^2)$, and such that the convexity property (3.5) is satisfied, namely

$$B(\mathbf{s_J}) - \frac{B(\mathbf{s_{J_+}}) + B(\mathbf{s_{J_-}})}{2} \geq \frac{1}{4} \left(\frac{y_J}{v_J} \right)^2 \alpha_J = \frac{1}{4Q|J|} \left(\frac{m_J(fw)}{m_J w} \right)^2 \lambda_J.$$

Hence,

$$|J| m_J(f^2) \geq |J| B(\mathbf{s_J})$$
$$\geq \frac{1}{4Q} \left(\frac{m_J(fw)}{m_J w} \right)^2 \lambda_J + |J| \frac{B(\mathbf{s_{J_+}}) + B(\mathbf{s_{J_-}})}{2}$$
$$= \frac{1}{4Q} \left(\frac{m_J(fw)}{m_J w} \right)^2 \lambda_J + |J_+| B(\mathbf{s_{J_+}}) + |J_-| B(\mathbf{s_{J_-}}).$$

Applying the convexity property (3.5) to $B(\mathbf{s_{J_\pm}})$, and iterating the argument, we obtain that

$$|J|m_J(f^2) \geq \frac{1}{4Q} \sum_{I \in \mathcal{D}(J)} \left(\frac{m_I(fw)}{m_I w} \right)^2 \lambda_I.$$

We are done proving Sawyer's estimate, provided such function B exists. ∎

3.2. Existence of Bellman function

The function

(3.6)
$$B(x, y, w, v, M) = x - \frac{y^2}{v + M}$$

is defined on the domain D, satisfies range property, $0 \leq B \leq x$, and furthermore the following differential properties hold,

$$\frac{\partial B}{\partial M} \geq \frac{y^2}{4v^2}, \qquad -\mathrm{d}^2 B \geq 0.$$

The boundedness and differential properties of the given function B on the domain are left as an exercise for the reader, see Lemma 11 for similar computations.

The convexity of the domain together with the infinitesimal differential properties, imply the discrete convexity property (3.5).

Notice that for $\mathbf{s} = (x, y, v, w, M)$, $\mathbf{s}_\pm = (x_\pm, y_\pm, v_\pm, w_\pm, M_\pm)$, and $\frac{\mathbf{s}_+ + \mathbf{s}_-}{2} = \mathbf{s} - (0, 0, 0, 0, \alpha) = (x, y, v, w, M - \alpha)$, by the Mean Value Theorem and the Fundamental Theorem of Calculus,

$$B(\mathbf{s}) - \frac{B(\mathbf{s}_+) + B(\mathbf{s}_-)}{2} = [B(\mathbf{s}) - B(s_0)] + \left[B(s_0) - \frac{B(\mathbf{s}_+) + B(\mathbf{s}_-)}{2} \right]$$

(3.7)
$$= \frac{\partial B}{\partial M}(x, y, v, w, M')\alpha - \int_{-1}^{1} (1 - |t|)b''(t)dt,$$

where $s_0 = \frac{\mathbf{s}_+ + \mathbf{s}_-}{2}$, $M' = M(T) = (1 - T)M + T(M - \alpha)$ for some $0 < T < 1$, and

$$b(t) = B\big(\mathbf{s}(t)\big), \quad \mathbf{s}(t) = \frac{1 + t}{2}\mathbf{s}_+ + \frac{1 - t}{2}\mathbf{s}_-, \quad -1 \leq t \leq 1.$$

Notice that $\mathbf{s}(1) = \mathbf{s}_+$, $\mathbf{s}(-1) = \mathbf{s}_-$, and $\mathbf{s}(0) = s_0 = \frac{\mathbf{s}_+ + \mathbf{s}_-}{2}$.

The differential properties together with (3.7) imply (notice that $b''(t) = d^2 B(\mathbf{s}(t))$).

$$B(\mathbf{s}) - \frac{B(\mathbf{s}_+) + B(\mathbf{s}_-)}{2} \geq \frac{y^2}{4v^2}\alpha,$$

which is what we wanted to prove.

We are entitled to use (3.7) as long as $\mathbf{s}, \mathbf{s}_\pm \in D$ imply that (i) $\big(x, y, v, w,$ $M(T)\big) \in D$ for all $0 \leq T \leq 1$, and (ii) $\mathbf{s}(t) \in D$ for all $-1 \leq t \leq 1$. These are the convexity properties that the domain must satisfy, and they are not difficult to prove. We leave the proof as an exercise in convex analysis for the reader. Similar calculations have been done in detail in other papers, see for example [26], [16], [37], [31]. A similar argument will be used in Section 6.4 and in Section 7. In the first case, the non-convex domain is distorted according to the doubling constant $D(\sigma)$ of an underlying doubling measure σ, and the differentiability domain is enlarged by a parameter depending on $D(\sigma)$. In the second case, the domain is non-convex, and the differentiability domain is enlarged by a parameter depending on $D(w)$, the doubling constant of the weight w.

3.3. Weighted maximal function

It is well known that the dyadic weighted maximal function M_w^d is bounded in $L^q(\mathbb{R})$ if and only if $w \in RH_p^d$, $\frac{1}{p} + \frac{1}{q} = 1$, see [30] and [28]. We will show in this section Lemma 1, that is the estimate,

$$\|M_w^d f\|_{L^2(\mathbb{R})} \leq C[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})}.$$

The following related optimal estimate for the Hardy-Littlewood maximal function on weighted L^p-spaces due to S. Buckley (see [4]) is well known:

$$\|Mf\|_{L^p(w)} \leq C[w]_{A_p}^{\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p(w)}.$$

See [27] and [20] for corresponding L^p-estimates for M_w.

Proof of Lemma 1. For the proof we will use a sharp weak boundedness result of Muckenhoupt [21] and Gehring's self-improvement theorem for Reverse Hölder classes [15]. More precisely, Muckenhoupt proved the following weak (q, q) estimate, for $1 \leq q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$,

$$(3.8) \qquad |\{x \in \mathbb{R} : M_w f(x) > \lambda\}| \leq \left(\frac{[w]_{RH_p}}{\lambda} \|f\|_{L^q(\mathbb{R})}\right)^q,$$

this estimate is sharp in the sense that $\phi(t) = t^q$, for $t = [w]_{RH_p}$, cannot be replaced by a faster decaying function of t. Notice that this sharp estimate automatically gives at least a linear bound in terms of $[w]_{RH_2}$ for the operator norm of the weighted maximal function, that is it shows that if $\|M_w f\|_{L^2(\mathbb{R})} \leq C[w]_{RH_2}^\alpha \|f\|_{L^2(\mathbb{R})}$, then $\alpha \geq 1$. Same result holds in the dyadic setting.

Gehring's self-improvement result states that $w \in RH_p$ implies that there exists an $0 < \epsilon_0 \sim [w]_{RH_p}^{-p}$, such that for all $\epsilon < \epsilon_0$, $w \in RH_{p+\epsilon}$.

Furthermore, there exists a constant $C > 0$ such that $[w]_{RH_{p+\epsilon}} \leq C[w]_{RH_p}$ whenever $\epsilon \sim [w]_{RH_p}^{-p}$.

It is a simple consequence of Hölder's inequality that if $w \in RH_p$ then for all $\epsilon < p - 1$, $w \in RH_{p-\epsilon}$, and $[w]_{RH_{p-\epsilon}} \leq [w]_{RH_p}$ (with constant 1). Same results hold in the dyadic setting.

We will now concentrate in the case $p = 2$.

We are going to interpolate Muckenhoupt's weak bounds with end-points (q_1, q_1), (q_2, q_2), with q_1, q_2 dual exponents of $p_1 = 2 + \epsilon$, $p_2 = 2 - \epsilon$, $\epsilon \sim [w]_{RH_2^d}^{-2}$, i.e. $q_1 = \frac{2+\epsilon}{1+\epsilon} < 2$, $q_2 = \frac{2-\epsilon}{1-\epsilon} > 2$. Interpolation guarantees that if our operator obeys weak bouds of the type

$$\left| \{ x \in \mathbb{R} : M_w^d f(x) > \lambda \} \right| \leq \left(\frac{B_i}{\lambda} \|f\|_{L^{q_i}(\mathbb{R})} \right)^{q_i}, \quad i = 1, 2;$$

then M_w^d is bounded in $L^2(\mathbb{R})$ ($q_1 < 2 < q_2$), furthermore,

$$\|M_w^d f\|_{L^2(\mathbb{R})} \leq C_t B_1^{1-t} B_2^t \|f\|_{L^2(\mathbb{R})},$$

where $\frac{1}{2} = \frac{1-t}{q_1} + \frac{t}{q_2}$, and $C_t^2 = \frac{2q_2}{q_2-2} + \frac{2q_1}{2-q_1}$. We can write all these variables in terms of ϵ,

$$t = \frac{2 - \epsilon}{4}, \quad C_t = \frac{2\sqrt{2}}{\sqrt{\epsilon}}.$$

We are assuming that $\epsilon \sim [w]_{RH_2^d}^{-2}$, hence $C_t \sim [w]_{RH_2^d}$, and by Muckenhoupt's weak bounds we know that $B_1 = [w]_{RH_{2+\epsilon}^d} \leq C[w]_{RH_2^d}$, and $B_2 = [w]_{RH_{2-\epsilon}^d} \leq [w]_{RH_2^d}$. All together we conclude that

$$\|M_w^d f\|_{L^2(\mathbb{R})} \leq C[w]_{RH_2^d}^2 \|f\|_{L^2(\mathbb{R})},$$

which is what we wanted to show. ∎

What we have shown so far is that if $\|M_w f\|_2 \leq C[w]_{RH_2^d}^\alpha \|f\|_{L^2(\mathbb{R})}$, then $1 \leq \alpha \leq 2$. The quadratic upper bound is the sharp bound, as Kabe Moen very recently proved [20].

4. Sharp bound for $T_w^{1/2}$ and corollaries

Verifying the boundedness of $T_w^{1/2}$ in $L^2(\mathbb{R})$ is equivalent to verifying the boundedness of its adjoint, furthermore the operator norms are the same. With this in mind, our problem is to obtain the following estimate,

$$(4.1) \qquad \|(T_w^{1/2})^* f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} \frac{|\langle w^{1/2} f, h_I \rangle|^2}{m_I w} \leq C[w]_{A_2^d} \|f\|_{L^2(\mathbb{R})}^2.$$

Let $g = w^{1/2}f$, and observe that $f \in L^2(\mathbb{R})$ if and only if $g \in L^2(w^{-1})$, furthermore, $\|f\|_{L^2(\mathbb{R})} = \|g\|_{L^2(w^{-1})}$, therefore the estimate we are seeking is equivalent to verifying that there exists a constant $C > 0$ such that for all $g \in L^2(w^{-1})$

$$(4.2) \qquad \sum_{I \in \mathcal{D}} \frac{|\langle g, h_I \rangle|^2}{m_I w} \leq C[w]_{A_2^d} \|g\|_{L^2(w^{-1})}^2.$$

This is exactly what Petermichl and Pott proved [33, see equation (3.2)].

Let us first introduce some operator notation for multiplication in space, and dyadic multiplication, the building blocks of the Haar multipliers. Denote by $\mathcal{M}_w^t$, $t \in \mathbb{R}$, w a weight, the linear operator (possibly unbounded) of *multiplication* by w^t, and D_w^t the linear operator of *dyadic multiplication* also possible unbounded, defined by its action on Haar functions, $h_I \to (m_I w)^t\, h_I$. That is,

$$\mathcal{M}_w^t f = w^t f, \quad D_w^t f = \sum_{I \in \mathcal{D}} (m_I w)^t \langle f, h_I \rangle h_I.$$

When denoting an operator T followed by multiplication by w^t we will often write directly $w^t T$ instead of $\mathcal{M}_w^t T$, similarly we will often write $T w^t$ instead of $T\mathcal{M}_w^t$ when this will not cause any confusion. Notice that with this notation, the Haar multipliers we have been studying are given by

$$T_w^t = \mathcal{M}_w^t D_w^{-t} = w^t D_w^{-t}.$$

Also notice that these operators, $\mathcal{M}_w^t$ and D_w^t, are positive operators, formally selfadjoint, and formally invertible with formal inverses given by $(\mathcal{M}_w^t)^{-1} = \mathcal{M}_w^{-t}$ and $(D_w^t)^{-1} = D_w^{-t}$. Also notice that formally we can compute adjoints and inverses for the Haar multipliers,

$$(T_w^t)^* = D_w^{-t} w^t, \quad (T_w^t)^{-1} = (D_w^{-t})^{-1}(\mathcal{M}_w^t)^{-1} = D_w^t w^{-t} = (T_w^{-t})^*.$$

Note that if $w \in A_2^d$, then both $T_w^{1/2}$ and $T_w^{-1/2}$ are bounded operators in $L^2(\mathbb{R})$, so are their adjoints, and the inverse of $T_w^{1/2}$ is $(T_w^{-1/2})^*$.

To be safe when handling the possibly unbounded multiplication and dyadic multiplication operators, we consider the following truncated weights: for $n \in \mathbb{N}$, $w_n(x) = \min\{\max\{w(x), 1/n\}, n\}$. If $w \in A_2^d$, then $w_n \in A_2^d$ (this is true of any weight bounded away from zero and from infinity, except that the A_2^d-characteristic can depend on the upper and lower bounds), and furthermore $[w_n]_{A_2^d} \leq C[w]_{A_2^d}$. The multiplication and dyadic multiplication corresponding to these truncations are positive, selfadjoint, invertible and bounded operators (the inverses are also bounded), although not uniformly

on n. However the Haar multipliers $T_{w_n}^{\pm 1/2}$ will be uniformly bounded (by an appropriate power of the $[w]_{A_2^d}$), and for g in the dense subspace of finite linear combinations of Haar functions we know that $T_{w_n}^{\pm 1/2} g \to T_w^{\pm 1/2} g$ in $L^2(\mathbb{R})$, therefore by the uniform boundedness principle this holds for all $f \in L^2(\mathbb{R})$ and moreover the same uniform bound that works for the truncated Haar multipliers will hold for the non-truncated Haar multipliers. We will prove all estimates with bounds independent of the truncation parameter, then take appropriate limits as $n \to \infty$. To ease the notation, we will not write down the truncation parameter.

The following observation will be useful:

Lemma 5. *Given $w \in A_2^d$ then the operator $w^{1/2} S^d D_w^{-1/2}$ is an isometry in $L^2(\mathbb{R})$.*

Proof. A direct calculation shows that

$$\|w^{1/2} S^d D_w^{-1/2} f\|_{L^2(\mathbb{R})}^2 = \int_{\mathbb{R}} \sum_{I \in \mathcal{D}} \frac{|\langle f, h_I \rangle|^2}{m_I w} \frac{\chi_I(x)}{|I|} w(x)\, dx$$

$$= \sum_{I \in \mathcal{D}} \frac{|\langle f, h_I \rangle|^2}{m_I w} m_I w = \|f\|_{L^2(\mathbb{R})}^2$$

$\blacksquare$

Assuming (1.5), that is the inverse estimate for the square function, we can now present a proof of Theorem 3. In fact we will show that both estimates are equivalent.

Lemma 6 (Theorem 3 is equivalent to (1.5))**.** *There exists a constant $C > 0$ such that for all $w \in A_2^d$,*

$$(4.3) \qquad \|T_w^{1/2} f\|_{L^2(\mathbb{R})} \le C[w]_{A_2^d}^{1/2} \|f\|_{L^2(\mathbb{R})}$$

if and only if there exists a constant $C > 0$ such that for all $w \in A_2^d$,

$$(4.4) \qquad \|f\|_{L^2(w)} \le C[w]_{A_2^d}^{1/2} \|S^d f\|_{L^2(w)}.$$

Proof. ($\Rightarrow$) The first observation is that (4.4) is equivalent to showing

$$(4.5) \qquad \|g\|_{L^2(\mathbb{R})} \le C[w]_{A_2^d}^{1/2} \|w^{1/2} S^d w^{-1/2} g\|_{L^2(\mathbb{R})}.$$

Let $g = T_w^{1/2} f$, then $f = (T_w^{1/2})^{-1} g = D_w^{1/2} w^{-1/2} g$, and substitute into (4.3) to get,

$$\|g\|_{L^2(\mathbb{R})} \le C[w]_{A_2^d}^{1/2} \|D_w^{1/2} w^{-1/2} g\|_{L^2(\mathbb{R})}.$$

Now we use Lemma 5 to force the square function into the right hand side,

$$\|g\|_{L^2(\mathbb{R})} \leq C[w]_{A_2^d}^{1/2} \|w^{1/2} S^d D_w^{-1/2} D_w^{1/2} w^{-1/2} g\|_{L^2(\mathbb{R})}$$
$$= C[w]_{A_2^d}^{1/2} \|w^{1/2} S^d w^{-1/2} g\|_{L^2(\mathbb{R})}.$$

This is exactly what we wanted to prove.

($\Leftarrow$) Let $f = (T_w^{1/2})^{-1} g$, then $g = T_w^{1/2} f = w^{1/2} D_w^{-1/2} f$, substituting into (4.5) we get

$$\|T_w^{1/2} f\|_{L^2(\mathbb{R})} \leq C[w]_{A_2^d}^{1/2} \|w^{1/2} S w^{-1/2} w^{1/2} D_w^{-1/2} f\|_{L^2(\mathbb{R})}$$
$$= C[w]_{A_2^d}^{1/2} \|w^{1/2} S D_w^{-1/2} f\|_{L^2(\mathbb{R})}$$
$$= C[w]_{A_2^d}^{1/2} \|f\|_{L^2(\mathbb{R})}$$

Where we used the isometry in Lemma 5 for the last equality. ∎

We will now show how to deduce Theorem 1, and the Hukovic-Treil-Volberg inequality (1.1) as simple corollaries of Theorem 3. Of course at some point there should be a proof of either the Petermichl-Pott Theorem or of Theorem 3 independent of each other, this was done in [33]. We will present a generalization of their argument for the σ-case in Section 6.

Corollary 1 (Theorem 1). *Given $w \in A_2^d$ then*

$$\|T_w^{-1/2} f\|_{L^2(\mathbb{R})} \leq C[w]_{A_2^d} \|f\|_{L^2(\mathbb{R})}.$$

Proof. A direct calculation for the adjoint, using the fact that $m_I w \, m_I(w^{-1}) \leq [w]_{A_2^d}$ shows that,

$$\|(T_w^{-1/2})^* f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} |\langle w^{-1/2} f, h_I \rangle|^2 m_I w$$
$$\leq \sum_{I \in \mathcal{D}} |\langle w^{-1/2} f, h_I \rangle|^2 \frac{[w]_{A_2^d}}{m_I(w^{-1})}.$$
$$= [w]_{A_2^d} \|(T_{w^{-1}}^{1/2})^* f\|_{L^2(\mathbb{R})}^2$$
$$\leq C[w]_{A_2^d}^2 \|f\|_{L^2(\mathbb{R})}^2.$$

for the last inequality we used Theorem 3 applied to w^{-1} instead of w, and the fact that $w \in A_2^d$ implies $w^{-1} \in A_2^d$ with the same A_2^d-characteristic, $[w^{-1}]_{A_2^d} = [w]_{A_2^d}$. This gives us the estimate for the adjoint, hence the same estimate holds for the operator. ∎

Corollary 2 (Hukovic-Treil-Volberg inequality (1.1)). *Given $w \in A_2^d$ then*

$$\|S^d f\|_{L^2(w)} \le C[w]_{A_2^d} \|f\|_{L^2(w)}.$$

Proof. The first observation is that what we want to prove is equivalent to showing

$$\|w^{1/2} S^d w^{-1/2} g\|_{L^2(\mathbb{R})} \le C[w]_{A_2^d} \|g\|_{L^2(\mathbb{R})}.$$

We now estimate the left-hand side taking advantage of the isometry in Lemma 5, and the linear estimate already proven for $T_w^{-1/2}$ in Corollary 1,

$$
\begin{aligned}
\|w^{1/2} S^d w^{-1/2} g\|_{L^2(\mathbb{R})} &= \|w^{1/2} S^d D_w^{-1/2} D_w^{1/2} w^{-1/2} g\|_{L^2(\mathbb{R})} \\
&= \|D_w^{1/2} w^{-1/2} g\|_{L^2(\mathbb{R})} \\
&= \|(T_w^{-1/2})^* g\|_{L^2(\mathbb{R})} \\
&\le C[w]_{A_2^d} \|g\|_{L^2(\mathbb{R})}.
\end{aligned}
$$

This is exactly what we wanted to prove. $\blacksquare$

A few more observations are in order. First of all, we can verify that the following equalities hold,

$$\|f\|_{L^2(w)}^2 = \langle \mathcal{M}_w f, f \rangle, \quad \|S^d f\|_{L^2(w)}^2 = \langle D_w f, f \rangle.$$

In the language of operators we can rephrase, as Petermichl and Pott did in [33], the direct and inverse estimates for the square function as follows,

$$D_w \le C[w]_{A_2^d}^2 \mathcal{M}_w, \quad \mathcal{M}_w \le c[w]_{A_2^d} D_w,$$

where it is understood that for two operators, $A \le B$ if and only if $\langle Af, f \rangle \le \langle Bf, f \rangle$, for all $f \in L^2(w)$. Theorem 3 or its equivalent formulation (4.2) can also be restated in operator language,

$$D_w^{-1} \le C[w]_{A_2^d} \mathcal{M}_w^{-1}.$$

In fact we have just shown that for this very special pair of operators, D_w and $\mathcal{M}_w$ (positive, selfadjoint, invertible) that

$$\mathcal{M}_w \le C[w]_{A_2^d} D_w \text{ if and only if } D_w^{-1} \le C[w]_{A_2^d} \mathcal{M}_w^{-1}.$$

This was the departure point in [33]. This statement always holds for any pair of positive, self-adjoint and invertible operators. If the operators commute this is trivial, if they don't, like in our case, it requires a small argument. This is a standard result in the theory of C^*-algebras, see for example [22, Thm 2.2.5]), for a proof using spectral theory, in particular the Gelfand representation.

5. Direct and inverse bounds for S^d_σ in $L^2(vd\sigma)$

In this section we will see that one can deduce Theorem 4, Theorem 5 and Theorem 7 as corollaries of Theorem 6, following the same scheme as in Section 4.

We will assume in this section that $v \in A^d_2(d\sigma)$, and σ is a doubling measure.

Recall that for a doubling measure σ, we define the σ-dyadic square function by

$$S^d_\sigma f(x) = \left(\sum_{j \in \mathbb{Z}} |\Delta^\sigma_j f(x)|^2 \right)^{1/2} = \left(\sum_{I \in \mathcal{D}} |m^\sigma_I f - m^\sigma_{\tilde{I}} f|^2 \chi_I(x) \right)^{1/2},$$

where the j-th σ-difference is $\Delta^\sigma_j = E^\sigma_{j+1} - E^\sigma_j$, and the σ-expectation is given by

$$E^\sigma_j f(x) = \frac{1}{\sigma(I)} \int_I f \, d\sigma = m^\sigma_I f, \quad x \in I \in \mathcal{D}_j.$$

It is well known that S^d_σ is bounded in $L^2(vd\sigma)$ if and only if $v \in A_2(d\sigma)$. Let $d\mu = vd\sigma$, it is easy to check that

$$\|S^d_\sigma f\|^2_{L^2(d\mu)} = \sum_{I \in \mathcal{D}} \mu(I)|m^\sigma_I f - m^\sigma_{\tilde{I}} f|^2.$$

Given a doubling measure σ, one can construct Haar functions $\{h^\sigma_I\}_{I \in \mathcal{D}}$ that form an orthonormal basis in $L^2(d\sigma)$, sometimes called *weighted Haar system*, see [8]. Such functions are step functions similar to the Haar functions except that the weights on each half are different. More precisely,

$$h^\sigma_I(x) = a^\sigma_{I_+} \chi_{I_+}(x) - a^\sigma_{I_-} \chi_{I_-}(x),$$

where $a^\sigma_I = \sqrt{\frac{\sigma(I^*)}{\sigma(I)\sigma(\tilde{I})}}$, $\tilde{I}$ is the parent of I, and I^* is the sibling of I. With this choice, it is not hard to check that,

$$(5.1) \qquad \Delta^\sigma_j f(x) = \langle f, h^\sigma_{\tilde{I}} \rangle_\sigma h^\sigma_{\tilde{I}}(x) = m^\sigma_I f - m^\sigma_{\tilde{I}} f, \quad x \in I \subset \tilde{I} \in \mathcal{D}_j.$$

For $d\mu = vd\sigma$, we can now compute the $L^2(d\mu)$-norm of $S^d_\sigma f$ in terms of the system $\{h^\sigma_I\}_{I \in \mathcal{D}}$,

$$\|S^d_\sigma f\|^2_{L^2(d\mu)} = \sum_{I \in \mathcal{D}} \mu(I)|\langle f, h^\sigma_{\tilde{I}} \rangle_\sigma h^\sigma_{\tilde{I}}(x_I)|^2$$

$$= \sum_{\tilde{I} \in \mathcal{D}} \left(\frac{\mu(I)\sigma(I^*)}{\sigma(\tilde{I})\sigma(I)} + \frac{\mu(I^*)\sigma(I)}{\sigma(\tilde{I})\sigma(I^*)} \right) |\langle f, h^\sigma_{\tilde{I}} \rangle_\sigma|^2,$$

where x_I denotes a point in I.

But σ is doubling, and if we denote by $D(\sigma)$ its doubling constant, then $D^{-1}(\sigma) \leq \frac{\sigma(I^*)}{\sigma(I)} \leq D(\sigma)$, and we conclude that

$$D^{-1}(\sigma) m_{\tilde{I}}^\sigma v \leq \frac{\mu(I)\sigma(I^*)}{\sigma(\tilde{I})\sigma(I)} + \frac{\mu(I^*)\sigma(I)}{\sigma(\tilde{I})\sigma(I^*)} \leq D(\sigma) m_{\tilde{I}}^\sigma v.$$

Therefore

$$(5.2) \quad D^{-1}(\sigma) \sum_{I \in \mathcal{D}} m_I^\sigma v |\langle f, h_I^\sigma \rangle_\sigma|^2 \leq \|S_\sigma^d f\|_{L^2(d\mu)}^2 \leq D(\sigma) \sum_{I \in \mathcal{D}} m_I^\sigma v |\langle f, h_I^\sigma \rangle_\sigma|^2.$$

Denote by $(D_v^\sigma)^{-1/2}$ the *discrete multiplication*, a possibly unbounded operator densely defined to map h_I^σ into $(m_I^\sigma v)^{-1/2} h_I^\sigma$.

Now we can reproduce almost verbatim what we did in Section 4. Estimate (5.2) applied to $(D_v^\sigma)^{-1/2} f$ implies the analogue of Lemma 5 in this context, namely

Lemma 7. *Let $v \in A_2(d\sigma)$ then*

$$D^{-1/2}(\sigma)\|f\|_{L^2(d\sigma)} \leq \|v^{1/2} S_\sigma^d (D_v^\sigma)^{-1/2} f\|_{L^2(d\sigma)} \leq D^{1/2}(\sigma)\|f\|_{L^2(d\sigma)}.$$

The corollaries and their proofs are very similar to the corresponding ones we did in Section 4 (we omit their proofs), modulo the appearance of a dependence of constants on the doubling constant of σ each time we use Lemma 7, and each time we use Theorem 6, in particular the precise estimate (6.1) to be proved in Section 6. We obtain the following direct and inverse estimates for S_σ^d in $L^2(vd\sigma)$,

Corollary 3 (Theorem 4). *Given σ a doubling measure, $v \in A_2^d(d\sigma)$ then there exists a constant $C > 0$, independent of v and σ, such that for all $f \in L^2(vd\sigma)$,*

$$(5.3) \qquad \|S_\sigma^d f\|_{L^2(vd\sigma)} \leq C D^2(\sigma) [v]_{A_2^d(d\sigma)} \|f\|_{L^2(vd\sigma)}.$$

Corollary 4 (Theorem 5). *Given σ a doubling measure, $v \in A_2^d(d\sigma)$ then there exists a constant $C > 0$, independent of v and σ, such that for all $f \in L^2(vd\sigma)$,*

$$(5.4) \qquad \|f\|_{L^2(vd\sigma)} \leq C D^2(\sigma) [v]_{A_2^d(d\sigma)}^{1/2} \|S_\sigma^d f\|_{L^2(vd\sigma)}.$$

Both for the direct and the inverse estimates the constants depend not only on the $A_2^d(d\sigma)$ characteristic of the weight v (at the rates advertised in Theorem 4 and Theorem 5), but also on $D(\sigma)$. The dependence on $D(\sigma)$ comes from Theorem 6, more precisely (6.1) (responsible for a power $3/2$ in both estimates) and from Lemma 7 (responsible for a power $1/2$ in both estimates). An intermediate estimate needed to get Corollary 3 is the following estimate, which also follows almost verbatim the proof of Corollary 1.

Corollary 5 (Theorem 7). *Given σ a doubling measure, $v \in A_2^d(d\sigma)$ then there exists a constant $C > 0$, independent of v and σ, such that for all $f \in L^2(d\sigma)$,*

$$\|(T_v^\sigma)^{-1/2}f\|_{L^2(d\sigma)} \le CD^{3/2}(\sigma)[v]_{A_2^d(d\sigma)}\|f\|_{L^2(d\sigma)}.$$

Notice also that if we specialize to the case $d\sigma = wdx$ and $v = w^{-1}$, and remember that $[w^{-1}]_{A_2^d(wdx)} = [w]^2_{RH_2^d}$, we obtain the following direct and inverse estimates for S_w^d in $L^2(\mathbb{R})$,

$$(5.5) \qquad \|S_w^d f\|_{L^2(\mathbb{R})} \le CD^2(w)\,[w]^2_{RH_2^d}\|f\|_{L^2(\mathbb{R})},$$

$$(5.6) \qquad \|f\|_{L^2(\mathbb{R})} \le CD^2(w)\,[w]_{RH_2^d}\|S_w^d f\|_{L^2(\mathbb{R})}.$$

In both cases the constants depend on $D(w)$. Note that in Section 3 we got linear dependence on the doubling constant for the direct estimate of S_w, whereas the argument just presented yields the larger quadratic power.

6. Sharp bounds for $(T_v^\sigma)^{1/2}$

In this section we will prove an estimate of the form

$$(6.1) \qquad \|(T_v^\sigma)^{1/2}f\|_{L^2(d\sigma)} \le CD^{3/2}(\sigma)[v]_{A_2(d\sigma)}^{1/2}\|f\|_{L^2(d\sigma)}.$$

Note that the dependence on the $A_2(d\sigma)$-characteristic of v is like a square root, and the dependence on the dyadic doubling constant of σ is like a power $3/2$.

Proof of Theorem 6. Suffices to prove the following estimate

$$(6.2) \qquad \sum_{I \in \mathcal{D}} \frac{1}{m_I^\sigma(v)}|\langle f, h_I^\sigma\rangle_\sigma|^2 \le CD^3(\sigma)[v]_{A_2^d(d\sigma)}\|f\|^2_{L^2(v^{-1}d\sigma)}.$$

Notice that the left hand side of (6.2) is $\|(T_\sigma^{1/2})^*(v^{-1/2}f)\|^2_{L^2(d\sigma)}$, and that $\|f\|^2_{L^2(v^{-1}d\sigma)} = \|v^{-1/2}f\|^2_{L^2(d\sigma)}$. So (6.2) is equivalent to the square root estimate (6.1) for $T_\sigma^{1/2}$. Estimate (6.2) is exactly what was proved in [33] for the case $d\sigma = dx$ and $v = w$, and we will follow their proof very closely.

The weighted Haar system $\{h_I^\sigma\}_{I \in \mathcal{D}}$ satisfies the following identity,

$$h_I = \delta_I^\sigma h_I^\sigma + \gamma_I^\sigma \chi_I,$$

where $\delta_I^\sigma = \sqrt{\frac{m_{I_+}(d\sigma)m_{I_-}(d\sigma)}{m_I(d\sigma)}}$, and $\gamma_I^\sigma = \frac{\langle d\sigma, h_I\rangle}{|I|m_I(d\sigma)}$.

The same holds for another doubling measure μ,

$$h_I = \delta_I^\mu h_I^\mu + \gamma_I^\mu \chi_I.$$

Combining these two we can obtain a similar relation between h_I^σ and h_I^μ when $d\mu = v d\sigma$, namely

$$(6.3) \qquad\qquad h_I^\sigma = \delta_I^{v,\sigma} h_I^\mu + \gamma_I^{v,\sigma} \chi_I,$$

where

$$\delta_I^{v,\sigma} = \sqrt{\frac{m_{I_+}^\sigma(v) m_{I_-}^\sigma(v)}{m_I^\sigma(v)}}, \quad \gamma_I^{v,\sigma} = \frac{\langle v, h_I^\sigma \rangle_\sigma}{\sigma(I) m_I^\sigma(v)}.$$

Switching in (6.2) to the weighted Haar system $\{h_I^\mu\}_{I \in \mathcal{D}}$, which is an orthonormal basis in $L^2(d\mu)$, where $d\mu = v d\sigma$, the left hand side becomes,

$$\sum_{I \in \mathcal{D}} \frac{1}{m_I^\sigma(v)} \left| \langle f, \delta_I^{v,\sigma} h_I^\mu + \gamma_I^{v,\sigma} \chi_I \rangle_\sigma \right|^2 = (I) + (II) + (III),$$

where

$$(I) = \sum_{I \in \mathcal{D}} \frac{1}{m_I^\sigma(v)} |\delta_I^{v,\sigma}|^2 |\langle f, h_I^\mu \rangle_\sigma|^2,$$

$$(II) = \sum_{I \in \mathcal{D}} \frac{1}{m_I^\sigma(v)} \delta_I^{v,\sigma} \gamma_I^{v,\sigma} \langle f, h_I^\mu \rangle_\sigma \langle f, \chi_I \rangle_\sigma,$$

$$(III) = \sum_{I \in \mathcal{D}} \frac{1}{m_I^\sigma(v)} |\gamma_I^{v,\sigma}|^2 |\langle f, \chi_I \rangle_\sigma|^2,$$

The following estimates will hold:

$$(6.4) \qquad\qquad (I) \;\leq\; CD(\sigma) \|f\|_{L^2(v^{-1}d\sigma)}^2,$$
$$(6.5) \qquad\qquad (III) \;\leq\; CD^3(\sigma)[v]_{A_2^d(d\sigma)} \|f\|_{L^2(v^{-1}d\sigma)}^2, .$$

With those estimates in hand, we can control the second term by Cauchy-Schwarz,

$$|(II)| \leq \sqrt{(I)}\sqrt{(III)} \leq CD^2(\sigma)[v]_{A_2^d(d\sigma)}^{1/2} \|f\|_{L^2(v^{-1}d\sigma)}^2.$$

All together these imply,

$$\sum_{I \in \mathcal{D}} \frac{1}{m_I^\sigma(v)} \left| \langle f, \delta_I^{v,\sigma} h_I^\mu + \gamma_I^{v,\sigma} \chi_I \rangle_\sigma \right|^2 \leq CD^3(\sigma)[v]_{A_2^d(d\sigma)} \|f\|_{L^2(v^{-1}d\sigma)}^2.$$

Which is what we wanted to prove. ∎

The first estimate (6.4) is not difficult, it is in estimating term (III), inequality (6.5), where we will have to prove a Carleson's Lemma. To estimate term (III) we will follow the scheme described in Section 3: first prove a Sawyer's estimate which boils down to finding exactly the same Bellman function we found before, second we will need a weight lemma to jump-start properly the Sawyer's estimate, this time such lemma is not trivial.

6.1. Estimate for term (I)

Proof of estimate (I). Note that

$$\frac{|\delta_I^{v,\sigma}|^2}{m_I^\sigma(v)} = \frac{m_{I_+}^\sigma(v)m_{I_-}^\sigma(v)}{m_I^\sigma(v)m_I^\sigma(v)},$$

and that

$$m_I^\sigma(v) = \frac{\sigma(I_+)}{\sigma(I)}m_{I_+}^\sigma(v) + \frac{\sigma(I_-)}{\sigma(I)}m_{I_-}^\sigma(v), \quad \frac{\sigma(I_+)}{\sigma(I)} + \frac{\sigma(I_-)}{\sigma(I)} = 1.$$

Denote by $s = \frac{\sigma(I_+)}{\sigma(I)}$, which will be uniformly bounded away from 1 and 0 because σ is dyadic doubling, more precisely, $\epsilon \leq s, 1 - s \leq 1 - \epsilon$, for $0 < \epsilon = D(\sigma)^{-1} \leq 1/2$. The geometric-arithmetic inequality implies that,

$$\sqrt{m_{I_+}^\sigma(v)m_{I_-}^\sigma(v)} \leq \frac{1}{2\sqrt{s(1-s)}}m_I^\sigma(v) \leq \sqrt{\frac{D(\sigma)}{2}}m_I^\sigma(v).$$

Therefore, for $d\mu = vd\sigma$,

$$(I) \leq \frac{D(\sigma)}{2}\sum_{I\in\mathcal{D}}|\langle f, h_I^\mu\rangle_\sigma|^2 = \frac{D(\sigma)}{2}\sum_{I\in\mathcal{D}}|\langle v^{-1}f, h_I^\mu\rangle_\mu|^2$$

$$= \frac{D(\sigma)}{2}\|v^{-1}f\|_{L^2(d\mu)} = \frac{D(\sigma)}{2}\|f\|_{L^2(v^{-1}d\sigma)}.$$

Which is what we wanted to prove.   ∎

6.2. Estimate for term (III)

We are trying to show the following Lemma, which can be thought as a Carleson's Embedding Lemma after noticing that $\langle f, \chi_I^\sigma\rangle = \sigma(I)m_I^\sigma f$.

Lemma 8. *If $v \in A_2^d(d\sigma)$, and σ is a doubling dyadic measure, then there exists a constant $C > 0$, independent of σ and v, such that for all $f \in L^2(v^{-1}d\sigma)$,*

$$\sum_{I\in\mathcal{D}}\frac{\sigma^2(I)}{m_I^\sigma(v)}|\gamma_I^{v,\sigma}|^2|m_I^\sigma(f)|^2 \leq CD^3(\sigma)[v]_{A_2^d(d\sigma)}\|f\|_{L^2(v^{-1}d\sigma)}^2.$$

To prove this Carleson's Embedding Lemma, it will suffice to prove a Sawyer's Estimate, and a Weight Lemma to jump-start it with the right constant. This was exactly the scheme followed in the proof of Lemma 2

Lemma 9 (σ-Sawyer's Estimate). *Given a dyadic doubling measure σ, $Q > 0$, and a sequence of positive numbers $\{\lambda_I\}_{I \in \mathcal{D}}$. Suppose that for all dyadic intervals J*

$$\frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} |m_I^\sigma(v)|^2 \lambda_I \leq Q m_J^\sigma(v),$$

then

$$\frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} |m_I^\sigma(f)|^2 \lambda_I \leq 4Q m_J^\sigma(f^2 v^{-1}).$$

Lemma 10 (Weight Lemma). *If $v \in A_2^d(d\sigma)$, and σ is a doubling dyadic measure, then*

$$\frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} \frac{|\langle v, h_I^\sigma \rangle_\sigma|^2}{m_I^\sigma(v)} \leq 18 D^3(\sigma)[v]_{A_2^d(d\sigma)} m_J^\sigma(v).$$

This Weight Lemma is known to hold in the case $d\sigma = dx$, a proof can be found in [37], and it is a refinement of Buckley's characterization of A_∞^d by summation conditions, in the case the weight is in the subset A_2^d, see [5].

We will prove the σ-Sawyer's Estimate (Lemma 9) in Section 6.3, and the Weight Lemma (Lemma 10) in Section 6.4.

Proof of Lemma 8. Apply σ-Sawyer's Estimate (Lemma 9) with $Q = 18 D^3(\sigma)[w]_{A_2^d(d\sigma)} > 0$, and

$$\lambda_I = \sigma^2(I) \frac{|\gamma_I^{v,\sigma}|^2}{m_I^\sigma(v)} = \frac{|\langle v, h_I^\sigma \rangle_\sigma|^2}{\left(m_I^\sigma(v)\right)^3}.$$

By Weight Lemma 10, the hypothesis of σ-Sawyer's Estimate are satisfied, namely,

$$\frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} |m_I^\sigma(v)|^2 \lambda_I = \frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} \frac{|\langle v, h_I^\sigma \rangle_\sigma|^2}{m_I^\sigma(v)}$$
$$\leq 18 D^3(\sigma)[v]_{A_2^d(d\sigma)} m_J^\sigma(v)$$
$$\leq Q m_J^\sigma(v).$$

Therefore the conclusion of σ-Sawyer's Estimate holds, namely, for each $J \in \mathcal{D}$,

$$\frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} |m_I^\sigma(f)|^2 \lambda_I = \sum_{I \in \mathcal{D}} \frac{\sigma^2(I)}{m_I^\sigma(v)} |\gamma_I^{v,\sigma}|^2 |m_I^\sigma(f)|^2$$
$$\leq 72 D^3(\sigma)[v]_{A_2^d(d\sigma)} \frac{1}{\sigma(J)} \int_J f^2(x) v^{-1}(x) \, d\sigma.$$

Cancel $\sigma(J)$ and let J grow to $[0, \infty)$ and to $(-\infty, 0]$. Finally add both estimates to obtain Lemma 8 with $C = 72$. ∎

6.3. σ-Sawyer's Estimate

Proof of Lemma 9. Let us write σ-Sawyer's estimate in term of Lebesgue averages, the hypothesis (or Carleson condition) now reads,

$$(6.6) \qquad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left| \frac{m_I(vd\sigma)}{m_I(d\sigma)} \right|^2 \lambda_I \leq Q m_J(vd\sigma),$$

and the conclusion will now read,

$$(6.7) \qquad \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left| \frac{m_I(fd\sigma)}{m_I(d\sigma)} \right|^2 \lambda_I \leq 4Q m_J(f^2 v^{-1} d\sigma),$$

where here $m_I(d\sigma) := \sigma(I)/|I|$, $m_I(fd\sigma) := \frac{1}{|I|} \int_I f \, d\sigma$.

Notice that if $d\sigma = wdx$ and $v = w$, then this is exactly Lemma 3 (our first Sawyer's Estimate), for which we have a Bellman function. In fact, we proved in Section 3.2 that there exists a real-valued function of 5 variables, $B(\mathbf{s}) = B(x, y, w, v, M)$ whose domain D is given by those $\mathbf{s} = (x, y, w, v, M) \in \mathbb{R}^5$ such that

$$x, y, w, v, M \geq 0, \quad M \leq v, \quad y^2 \leq xv;$$

whose range is given by $0 \leq B(\mathbf{s}) \leq x$, $\mathbf{s} \in D$, and such that the following *convexity* property holds: for all $\mathbf{s}, \mathbf{s}_\pm \in D$ such that $\mathbf{s} - \frac{\mathbf{s}_+ + \mathbf{s}_-}{2} = (0, 0, 0, 0, \alpha)$ then

$$B(\mathbf{s}) - \frac{B(\mathbf{s}_+) + B(\mathbf{s}_-)}{2} \geq \frac{1}{4} \left(\frac{y}{v} \right)^2 \alpha.$$

Fix a dyadic interval I. Let $\mathbf{s_I} = (x_I, y_I, w_I, v_I, M_I)$, where $x_I = m_I(f^2 v^{-1} d\sigma)$, $y_I = m_I(fd\sigma)$, $w_I = m_I(d\sigma)$, $v_I = m_I(vd\sigma)$, and

$$M_I = \frac{1}{Q} \frac{1}{|I|} \sum_{K \in \mathcal{D}(I)} \left(\frac{m_K(vd\sigma)}{m_K(d\sigma)} \right)^2 \lambda_K.$$

Clearly for each $J \in \mathcal{D}$, $\mathbf{s_J}$ belongs to the domain, these are all positive quantities, $M_J \leq v_J$ is the hypothesis of Lemma 9, and $y_J^2 \leq x_J v_J$ is nothing more than Cauchy-Schwarz inequality. Let now $\mathbf{s}_\pm = \mathbf{s_{J_\pm}} \in D$. By definition, $\mathbf{s_J} - \frac{\mathbf{s_{J_+}} + \mathbf{s_{J_-}}}{2} = (0, 0, 0, 0, \alpha_J)$, where $\alpha_J = \frac{1}{Q|J|} \left(\frac{m_J(vd\sigma)}{m_J(d\sigma)} \right)^2 \lambda_J$.

Hence, $B(\mathbf{s_J}) \leq m_J(f^2 v^{-1} d\sigma)$, and since the convexity property holds,

$$B(\mathbf{s_J}) - \frac{B(\mathbf{s_{J_+}}) + B(\mathbf{s_{J_-}})}{2} \geq \frac{1}{4} \left(\frac{y_J}{v_J} \right)^2 \alpha_J = \frac{1}{4Q|J|} \left(\frac{m_J(fd\sigma)}{m_J(d\sigma)} \right)^2 \lambda_J.$$

Iterating, we obtain that

$$m_J(f^2 v^{-1} d\sigma) \geq \frac{1}{4Q|J|} \sum_{I \in \mathcal{D}(J)} \left(\frac{m_I(f d\sigma)}{m_I(d\sigma)} \right)^2 \lambda_I.$$

And we are done proving Sawyer's estimate.   ∎

6.4. Weight Lemma

Janine Wittwer proved Lemma 10 in the case $d\sigma = dx$, her proof uses Buckley's characterization of weights by summation conditions [5], Gehring's self-improvement theorem [15], Hölder's inequality, see [37]. There is a proof also using Bellman functions (although we have not been able to find a full reference, but modifications of other proofs lead to it). We will adapt the later for dyadic doubling positive measures σ. We present a Bellman function that gives the linear bound $C_\epsilon[v]_{A_2(d\sigma)}$, where ϵ is the dyadic doubling constant of σ. The Bellman function that gives linear dependence in the case $d\sigma = dx$ has been kindly provided by F. Nazarov [23], a small variation of it works for the doubling measure case.

Proof of Lemma 10. Fix $0 < \epsilon \leq 1/2$. Define the domain

$$D_Q = \{(u,v) \in \mathbb{R}_+^2 : 1 \leq uv \leq Q\}.$$

Suppose we can find a function of two variables $B(u,v)$ defined on D_Q, such that

(i) $0 \leq B(u,v) \leq Q\,v$ for all $(u,v) \in D_Q$,

(ii) if for all triplets $(u,v), (u_\pm, v_\pm) \in D_Q$ such that $u = su_+ + (1-s)u_-$ and $v = sv_+ + (1-s)v_-$ where $\epsilon \leq s \leq 1 - \epsilon$, then the following convexity condition holds for all such s, with $C_\epsilon > 0$,

$$\Delta_s B(u,v) = B(u,v) - sB(u_+, v_+) - (1-s)B(u_-, v_-)$$

(6.8)
$$\geq C_\epsilon^{-1} \frac{|v_+ - v_-|^2}{v}.$$

Then the lemma will be proved with bound $C_\epsilon[w]_{A_2^d(d\sigma)}$.

Let $v_I = m_I^\sigma v$, $u_I = m_I^\sigma(v^{-1})$, and $v_+ = v_{I_r}$, $v_- = v_{I_l}$, similarly for $u_\pm$. Let $Q = [v]_{A_2^d(d\sigma)}$, then $(u_I, v_I), (u_\pm, v_\pm) \in D_Q$. For each $I \in \mathcal{D}$, let $s = s_I = \frac{\sigma(I_r)}{\sigma(I)}$ (note that $1 - s = \frac{\sigma(I_l)}{\sigma(I)}$), then the dyadic doubling condition on σ implies that there is an $0 < \epsilon \leq 1/2$ such that $\epsilon \leq s, 1 - s \leq 1 - \epsilon$,

such ϵ is nothing more than the reciprocal of $D(\sigma)$, the doubling constant of σ. Fix $J \in \mathcal{D}$, by the convexity and the range conditions (i) and (ii), we conclude that

$$\sigma(J)Q\, m_J^\sigma v \geq \sigma(J)B(u_J, v_J)$$

$$\geq \sigma(J)s_J B(u_{J_r}, v_{J_r}) + \sigma(J)(1 - s_J)B(u_{J_l}, v_{J_l}) + \ C_\epsilon^{-1} \frac{|m_{J_r}^\sigma v - m_{J_l}^\sigma v|^2}{m_J^\sigma v}\sigma(J)$$

$$= \sigma(J_r)B(u_{J_r}, v_{J_r}) + \sigma(J_l)B(u_{J_l}, v_{J_l}) + \ C_\epsilon^{-1} \frac{|m_{J_r}^\sigma v - m_{J_l}^\sigma v|^2}{m_J^\sigma v}\sigma(J).$$

The last equality because by definition of s_J, $\sigma(J)s_J = \sigma(J_r)$, $\sigma(J)(1-s_J) = \sigma(J_l)$. We can now iterate, to conclude that

$$\sigma(J)Q\, m_J^\sigma v \geq C_\epsilon^{-1} \sum_{I \in \mathcal{D}(J)} \frac{|m_{I_r}^\sigma v - m_{I_l}^\sigma v|^2}{m_I^\sigma v}\sigma(I).$$

The last thing to notice is that

$$|\langle v, h_I^\sigma \rangle_\sigma|^2 = \frac{\sigma(I_r)\sigma(I_l)}{\sigma(I)}|m_{I_r}^\sigma v - m_{I_l}^\sigma v|^2 \leq \sigma(I)|m_{I_r}^\sigma v - m_{I_l}^\sigma v|^2.$$

All together these imply that

$$\frac{1}{\sigma(J)} \sum_{I \in \mathcal{D}(J)} \frac{|\langle v, h_I^\sigma \rangle_\sigma|^2}{m_I^\sigma v} \leq C_\epsilon Q\, m_J^\sigma v,$$

which is exactly what we wanted to prove: a linear bound multiplied by a positive constant $C = C_\epsilon > 0$ depending only on ϵ, the reciprocal of $D(\sigma)$, the dyadic doubling constant of σ.

All these will work provided we can construct the function $B(u, v)$ with the desired properties. ∎

Here is such a function (a small variation over the corresponding function when $d\sigma = dx$ that was provided by Nazarov in [23]),

(6.9) $$B(u, v) = Qv - \frac{Q}{u} - \frac{\epsilon}{3}v \ln(uv).$$

Let $0 < \epsilon \leq 1/2$. Define the domain

$$D_{\epsilon,Q} = \left\{(u, v) \in \mathbb{R}_+^2 : 1 \leq uv \leq 3\epsilon^{-1}Q\right\}.$$

Lemma 11. *The function B given by (6.9) is defined on $D_{\epsilon,Q}$, $0 \le B(u,v) \le Qv$, for all $(u,v) \in D_Q$, and satisfies the following differential inequality in $D_{\epsilon,Q}$,*

$$-d^2 B(u,v) \ge \frac{2\epsilon}{9} \frac{|dv|^2}{v}.$$

Furthermore these imply the convexity condition (6.8) on D_Q with constant $C_\epsilon = \frac{18}{\epsilon^3} = 18D(w)^3$.

Proof. By definition (6.9) it is clear that

$$B(u,v) \le Qv.$$

The fact that $B(u,v) \ge 0$ is nothing more than a calculus exercise, which we now describe,

$$B(u,v) = Qv - \frac{Q}{u} - \frac{\epsilon}{3}v\ln(uv) = v\left(Q - \frac{Q}{uv} - \frac{\epsilon}{3}\ln uv\right).$$

Suffices to show that the function $f(x) = Q\left(1 - \frac{1}{x}\right) - \frac{\epsilon}{3}\ln x$ is positive for $1 \le x < 3\epsilon^{-1}Q$. One can check that in that range $f'(x) > 0$, therefore f is an increasing function, hence $f(x) \ge f(1) = 0$, and we are done.

A direct computation of the Hessian of B for all $(u,v) \in D_{\epsilon,Q}$, shows that,

$$-d^2 B - \frac{2\epsilon}{9}\frac{|dv|^2}{v} = \begin{pmatrix} du & dv \end{pmatrix} \begin{pmatrix} \frac{2Q}{u^3} - \frac{\epsilon}{3}\frac{v}{u^2} & \frac{\epsilon}{3}\frac{1}{u} \\ \frac{\epsilon}{3}\frac{1}{u} & \frac{\epsilon}{3}\frac{1}{v} - \frac{2\epsilon}{9}\frac{1}{v} \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \ge 0.$$

Hence

$$-d^2 B \ge \frac{2\epsilon}{9}\frac{|dv|^2}{v},$$

as we wanted to show.

As for the convexity condition (6.8) on D_Q, it is a calculus exercise to check that
(6.10)
$$\Delta_s B(u,v) = -\int_{-1+s}^{0} \left(s(1-s) + st\right)b''(t)dt - \int_0^s \left(s(1-s) - (1-s)t\right)b''(t)dt,$$

where $b(t) = B(u(t), v(t))$, and

$$u(t) = (s-t)u_+ + (t + (1-s))u_-,$$

$$v(t) = (s-t)v_+ + (t + (1-s))v_-.$$

Notice that $v(0) = v$, $v(-1+s) = v_+$, $v(s) = v_-$, and $dv = v_- - v_+$; similarly for $u(t)$. Furthermore, if $(u,v), (u_\pm, v_\pm) \in D_Q$, then $\big(u(t), v(t)\big) \in D_{\epsilon,Q}$ for

all $-1+s \leq t \leq s$. The convexity of the lower boundary of the domain gives the lower estimate $1 \leq u(t)v(t)$. It is not hard to check that $v_\pm \leq \epsilon^{-1}v$, similarly $u_\pm \leq \epsilon^{-1}u$ (remember $0 < \epsilon \leq 1/2$), therefore, $v(t) \leq \epsilon^{-1}v$ and $u(t) \leq \epsilon^{-1}u$, which in turn implies $u(t)v(t) \leq \epsilon^{-2}uv \leq \epsilon^{-2}Q$. however this is not sufficient to be in the domain, doing more carefully the estimate we get that $u(t)v(t) \leq 3\epsilon^{-1}Q$. Here are the details: $uv = s^2 u_+ v_+ + (1-s)^2 u_- v_- + s(1-s)(u_+ v_- + u_- v_+)$, thus,

$$(6.11) \qquad u_+ v_- + u_- v_+ = \frac{uv}{s(1-s)} - \frac{s}{1-s}u_+ v_+ - \frac{1-s}{s}u_- v_-.$$

Similarly, setting $\alpha = s - t$, $0 \leq \alpha \leq 1$, and using (6.11), we get

$$u(t)v(t) = \alpha^2 u_+ v_+ + (1-\alpha)^2 u_- v_- + \alpha(1-\alpha)(u_+ v_- + u_- v_+)$$
$$= \frac{\alpha(-t)}{1-s}u_+ v_+ + \frac{(1-\alpha)t}{s}u_- v_- + \frac{\alpha(1-\alpha)}{s(1-s)}uv.$$

Notice that the third summand is always positive and bounded above by $2\epsilon^{-1}Q$ (by hypothesis $uv < Q$ and $s(1-s) > \epsilon/2$). The sign of the first two summands is dictated by the sign of t, and one is always positive while the other is negative. Dropping the negative term, and observing that the positive term is always bounded by Q, we conclude that for all $-1+s \leq t \leq s$, $u(t)v(t) \leq (2\epsilon^{-1} + 1)Q \leq 3\epsilon^{-1}Q$, as claimed.

So we are entitled to integrate $b''(t)$ from $t = -1 + s$ to $t = s$.

Finally notice that since $u''(t) = v''(t) = 0$ then

$$-b''(t) = -d^2 B(u(t), v(t)) \geq \frac{2\epsilon}{9}\frac{|dv|^2}{v(t)} \geq \frac{2\epsilon^2}{9}\frac{|v_+ - v_-|^2}{v} > 0.$$

Therefore, if we now integrate according to (6.10) and use the lower bound just found for $-b''(t)$, we obtain

$$\Delta_s B(u, v) \geq \frac{2\epsilon^2}{9}\frac{|v_+ - v_-|^2}{v}\frac{s(1-s)}{2}$$
$$\geq \frac{\epsilon^3}{18}\frac{|v_+ - v_-|^2}{v}.$$

In the second inequality $\frac{s(1-s)}{2}$ is nothing more than the area under the triangle with base the interval $[-1+s, s]$ and height $s(1-s)$, which is bounded from below by $\epsilon/2$.   ∎

This argument provides an operator bound for $(T_v^\sigma)^{1/2}$ of the order $D^{3/2}(\sigma)[v]_{A_2(d\sigma)}^{1/2}$. The square root dependence on the $A_2(d\sigma)$ characteristic of v is optimal, one could try to verify whether the $3/2$ power dependence on the doubling constant of σ can be improved or not, but we will not pursue this issue further in this paper.

7. Sharpness of the bounds for T_w

F. Nazarov, in a personal communication [23], observed that one can approach the problem of the boundedness of T_w in $L^2(\mathbb{R})$ using a Bellman function argument directly on the adjoint problem.

Lemma 12. *Suppose there is a function* $B = B(x, y, w, v)$ *defined on the domain*

$$D = \left\{ \mathbf{s} = (x, y, w, v) \in \mathbb{R}_+^4 : y^2 \leq xv, \ \ w^2 \leq v \leq Q^2 w^2 \right\},$$

where $Q > 1$, *such that, for some* $M > 0$,

(*i*) $0 \leq B(\mathbf{s}) \leq Mx$ *for all* $\mathbf{s} = (x, y, w, v) \in D$.

(*ii*) *Convexity: for all* $\mathbf{s_0}, \mathbf{s_\pm} \in D$ *such that* $\mathbf{s_0} = \frac{\mathbf{s_+} + \mathbf{s_-}}{2}$, *then*

$$\Delta^2 B := B(\mathbf{s_0}) - \left(\frac{B(\mathbf{s_+}) + B(\mathbf{s_-})}{2} \right) \geq c \left(\frac{\Delta y}{w} \right)^2,$$

where $\Delta y = \frac{y_+ - y_-}{2}$.

Then

$$\|(T_w)^* f\|_{L^2(\mathbb{R})}^2 = \sum_{I \in \mathcal{D}} \left| \frac{\langle wf, h_I \rangle}{m_I w} \right|^2 \leq c^{-1} M \|f\|_{L^2(\mathbb{R})}^2.$$

Proof. We can deduce boundedness of T_w in $L^2(\mathbb{R})$ from this lemma by the same type of arguments we have presented in the previous sections. Consider the function whose existence is claimed in Lemma 12. For $I \in \mathcal{D}$, $f \in L^2(\mathbb{R})$, $w \in RH_2^d$ with RH_2^d-characteristic Q. Let $\mathbf{s_I} = (x_I, y_I, w_I, v_I)$, where $x_I = m_I(f^2)$, $y_I = m_I(fw)$, $w_I = m_I w$, and $v_I = m_I(w^2)$, then $\mathbf{s_I} \in D$. Denote $\mathbf{s_{I_+}} = \mathbf{s_{I_r}}$, $\mathbf{s_{I_-}} = \mathbf{s_{I_l}}$, then $\mathbf{s_{I_\pm}} \in D$ and $\mathbf{s_I} = \frac{\mathbf{s_{I_+}} + \mathbf{s_{I_-}}}{2}$. Fix $J \in \mathcal{D}$, iterating over the convexity condition (*ii*) we conclude that,

$$\frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \left| \frac{\langle wf, h_I \rangle}{m_I w} \right|^2 \leq c^{-1} M x_J = c^{-1} M \frac{1}{|J|} \int_J |f(x)|^2 \, dx.$$

Cancel $|J|$, let J grow to $[0, \infty)$ and to $(-\infty, 0]$, and add the estimates to obtain

$$\sum_{I \in \mathcal{D}} \left| \frac{\langle wf, h_I \rangle}{m_I w} \right|^2 \leq c^{-1} M \|f\|_{L^2(\mathbb{R})}^2.$$

∎

Note that if w is dyadic doubling, all we need is the convexity condition (ii) to hold for those triplets $\mathbf{s_0}$, $\mathbf{s_\pm}$ in D such that $\mathbf{s_0} = \frac{\mathbf{s_+}+\mathbf{s_-}}{2}$ and such that

$$(7.1) \qquad w \leq \epsilon^{-1}\min\{w_+, w_-\}, \quad \text{for } \epsilon^{-1} = D(w).$$

When $w \in RH_2^d$ and dyadic doubling, we can find B like in Lemma 12 with $c^{-1}M \sim \epsilon^{-2}Q^4 = D^2(w)[w]_{RH_2^d}^4$, because we proved in Section 3 that $\|T_w f\|_{L^2(\mathbb{R})}^2 \leq CD^2(w)[w]_{RH_2^d}^4\|f\|_{L^2(\mathbb{R})}^2$. If we can show that there is no Bellman function with $c^{-1}M$ growing slower than Q^4 then this will show that the quartic bound $[w]_{RH_2^d}^4$ is optimal. There are reasons to believe the dyadic doubling constant of w should not be in the estimate, since the boundedness of T_w requires only $w \in RH_2^d$ and not the dyadic doubling condition, see [17]. However, all the arguments presented in this paper require at some point or another the use of the doubling conditions.

To find a Bellman function one reduces the convexity condition (ii) to an infinitesimal condition.

Lemma 13. *If we can find a function B such that the range condition (i) holds and the following infinitesimal condition holds,*

$$(ii') \qquad -d^2B \geq c'\left(\frac{dy}{w}\right)^2,$$

on the larger domain

$$D_\epsilon = \left\{\mathbf{s} = (x, y, w, v) \in \mathbb{R}_+^4 : y^2 \leq xv, \ w^2 \leq v \leq \epsilon^{-1}Q^2w^2\right\}.$$

Then convexity condition (ii) will hold for those triplets $\mathbf{s_0}$, $\mathbf{s_\pm}$ in D, such that $\mathbf{s_0} = \frac{\mathbf{s_+}+\mathbf{s_-}}{2}$ and such that $w_0 \leq \epsilon^{-1}\min\{w_+, w_-\}$,

Proof of Lemma 13. The fact that the differential condition (ii') holds on the larger domain implies now that condition (ii) holds by the usual integration argument. Parametrize the segment joining $\mathbf{s_+}$ and $\mathbf{s_-}$ by

$$\mathbf{s_0}(t) = \frac{1-t}{2}\mathbf{s_+} + \frac{1+t}{2}\mathbf{s_-}, \quad \text{for } -1 \leq t \leq t,$$

note that $\mathbf{s}(1) = \mathbf{s_+}$, $\mathbf{s}(-1) = \mathbf{s_-}$, and $\mathbf{s}(0) = \mathbf{s_0}$, $\frac{dy}{dt} = \Delta y = \frac{y_+ - y_-}{2}$, similarly for $\frac{dw}{dt} = \Delta w$, and $w(t) = w_0 + t\Delta w$.

$$\Delta^2 B = -\int_{-1}^{1}(1-|t|)d^2B(\mathbf{s}(t))\,dt \geq c'\left(\frac{\Delta y}{w_0}\right)^2\int_{-1}^{1}\frac{1-|t|}{(1+t\frac{\Delta w}{w_0})^2}\,dt.$$

The right-hand-side integral can be computed explicitly and is identical to $f\left(\frac{\Delta w}{w_0}\right)$, where $f(x) = -\frac{\ln(1-x^2)}{x^2}$. The function $f(x) \geq 1$ for $0 < x < 1$, and by doubling $x = \frac{\Delta w}{w_0}$ is bounded away from 1, in fact, $\frac{\Delta w}{w_0} \leq 1 - \epsilon/2$. This argument gives a constant $c = c'$ in (ii).

The larger domain will guarantee that if

$$\mathbf{s_0}, \mathbf{s_\pm} \in D \quad \text{and} \quad w_0 \leq \epsilon^{-1} \min\{w_+, w_-\}$$

then

$$\mathbf{s}(t) = \frac{1}{2}[(1+t)\mathbf{s_+} + (1-t)\mathbf{s_-}] \in D_\epsilon \quad \text{for all } t \in [-1, 1].$$

Thus the integration argument just described can be carried away.

In the application the parameter $0 < \epsilon \leq 1/2$ will be the reciprocal of $D(w)$. We will sketch the argument so that the appearance of the larger domain and its dependence on ϵ become clear.

The non-convexity of the domain D is an issue only for the variables v and w. We will concentrate on them, consider in the (w, v)-plane the closed region D' in the first quadrant trapped between the parabolas $v = w^2$ and $v = Q^2 w^2$. Without loss of generality, assume $w_+ < w_-$, and assume that the points (w_+, v_+) and (w_0, v_0) are on the graph of the parabola $v = Q^2 w^2$, where $w_0 = \frac{w_+ + w_-}{2}$ and $v_0 = \frac{v_+ + v_-}{2}$, and by hypothesis the point (w_-, v_-) is inside D'. With these choices, the segment L_- joining (w_-, v_-) and (w_0, v_0) must be completely inside the region D', and the segment L_+ joining (w_+, v_+) and (w_0, v_0) must lie completely outside D'. Our job is to find $\delta > 0$ such that the segment L_+ lies completely inside D'_δ, where D'_δ is the region in the (w, v)-plane trapped between the parabolas $v = w^2$ and $v = \delta^{-1} Q^2 w^2$. Notice that at the origin, all three parabolas have horizontal tangent, so if we choose (w_+, v_+) arbitrarily close to the origin, independently of (w_0, v_0), then no δ will do the job. However, that scenario can be ruled out if we take advantage of the additional constraint (7.1) imposed on $w_0, w_\pm$. In fact, given $0 < w_+$, then $w_0 \leq \epsilon^{-1} w_+$. With that information at hand, we can now explicitly calculate and bound the slope m of the segment L_+,

$$m = \frac{Q^2 w_0^2 - Q^2 w_+^2}{w_0 - w_+} = Q^2(w_0 + w_+) \leq (\epsilon^{-1} + 1)Q^2 w_+ \leq 2\epsilon^{-1} Q^2 w_+.$$

We can also calculate the slope m_δ of the tangent line to the graph of $v = \delta^{-1} Q^2 w^2$, at the point $(w_+, \delta^{-1} Q^2 w_+^2)$, namely, $m_\delta = 2\delta^{-1} Q^2 w_+$. If we now choose $\delta = \epsilon$, we conclude that $m \leq m_\epsilon$, and this is sufficient to guarantee that the segment L_+ lies completely inside D'_ϵ (the reader is encouraged to draw her own picture illustrating what we just described in words). Now the integration argument can be carried out safely. ∎

7.1. Reducing the number of variables

A nice trick that we learned from F. Nazarov [23], together with obvious homogeneity considerations allows to restrict ourselves to functions of certain type, more precisely,

Lemma 14. *If B is a Bellman function satisfying the range and convexity properties (i) and (ii) on D, then there is a function $\widetilde{B}$ of the form*

$$(7.2) \qquad \widetilde{B}(x, y, w, v) = Mx - \frac{y^2}{v} \psi\left(\frac{w}{\sqrt{v}}\right)$$

where $\psi : [Q^{-1}, 1] \to \mathbb{R}$ is a non-negative function, that is also a Bellman function satisfying the same range and convexity properties (i) and (ii) on D with the same constant M. Furthermore, the function ψ is bounded by M, namely

$$(7.3) \qquad M \geq \max_{Q^{-1} \leq u \leq 1} \psi(u) \geq 0.$$

If we denote $-\frac{y^2}{v}\psi\left(\frac{w}{\sqrt{v}}\right)$ by $\Psi(y, w, v)$, then Ψ must satisfy the same convexity condition as B and $\widetilde{B}$ do, namely, for all $\mathbf{s} = (x, y, w, v), \mathbf{s}_\pm = (x_\pm, y_\pm, w_\pm, v_\pm) \in D$, such that $\mathbf{s} = \frac{\mathbf{s}_+ + \mathbf{s}_-}{2}$,

$$(7.4) \qquad \Delta^2 \Psi(y, w, v) \geq \left(\frac{\Delta y}{w}\right)^2,$$

where $\Delta y = \frac{y_+ - y_-}{2}$, and

$$\Delta^2 \Psi(y, w, v) := \Psi(y, w, v) - \left(\frac{\Psi(y_+, w_+, v_+) + \Psi(y_-, w_-, v_-)}{2}\right).$$

If we can locate a function $\psi : [Q^{-1}, 1] \to \mathbb{R}$, such that $0 \leq \psi \leq M$, and if $\Psi(y, w, v) := -\frac{y^2}{v}\psi\left(\frac{w}{\sqrt{v}}\right)$ satisfies condition (7.4) in the appropriate domain then the function $B(x, y, w, v) = Mx + \Psi(y, w, v)$ defined on D is a Bellman function satisfying conditions (i) and (ii) with $c = 1$.

If the function ψ is twice differentiable, then we can replace the convexity condition (7.4) by the following infinitesimal condition on the larger domain $[\sqrt{\epsilon}Q^{-1}, 1]$

$$(7.5) \qquad -d^2\Psi(y, w, v) = d^2\left[\frac{y^2}{v}\psi\left(\frac{w}{\sqrt{v}}\right)\right] \geq \frac{(dy)^2}{w^2}.$$

Proof of Lemma 14. Let B be the true Bellman function for the problem. Assume that M is a constant such that $B(x, y, w, v) \leq Mx$. Since multiplication of the function and the weight by a positive constant changes nothing in the problem, the function B must obey the following homogeneity rule,

$$B(\alpha^2 x, \alpha\beta y, \beta w, \beta^2 v) = \alpha^2 B(x, y, w, v), \qquad \text{for all} \quad \alpha, \beta > 0.$$

Notice that the domain remains invariant under the above scaling, that is, $(x, y, w, v) \in D$ if and only if $(\alpha^2 x, \alpha\beta y, \beta w, \beta^2 v) \in D$ for all $\alpha, \beta > 0$.

Now consider the function of 3 variables, defined for all triplets $(y, w, v) \in \mathbb{R}_+^3$ such that there exists $x > 0$ such that $(x, y, w, v) \in D$,

$$\Psi(y, w, v) := \sup_{\{x \,:\, (x,y,w,v) \in D\}} [B(x, y, w, v) - Mx]$$

Note that $\Psi(y, w, v) \leq 0$, and for any x such that $(x, y, w, v) \in D$, we have

$$Mx + \Psi(y, w, v) \geq Mx + B(x, y, w, v) - Mx = B(x, y, w, v) \geq 0.$$

Note also that for all $\alpha, \beta > 0$,

$$\Psi(\alpha\beta y, \beta w, \beta^2 v) = \alpha^2 \Psi(y, w, v)$$

(this follows almost immediately from the corresponding property of B) and, therefore, (using $\alpha = \frac{\sqrt{v}}{y}$, $\beta = \frac{1}{\sqrt{v}}$)

$$\Psi(y, w, v) = \frac{y^2}{v} \Psi\left(1, \frac{w}{\sqrt{v}}, 1\right) = -\frac{y^2}{v} \psi\left(\frac{w}{\sqrt{v}}\right)$$

for some non-negative function ψ. So, if we could show that Ψ satisfies the same finite difference inequality as B, then

$$\widetilde{B}(x, y, w, v) := Mx + \Psi(y, w, v) = Mx - \frac{y^2}{v} \psi\left(\frac{w}{\sqrt{v}}\right)$$

would be another Bellman type function yielding the same constant M as the original Bellman function B. Consider any three triples in the domain of definition of Ψ such that

$$(y, w, v) = \frac{1}{2}[(y_-, w_-, v_-) + (y_+, w_+, v_+)]$$

By definition of Ψ as a supremum, given $\delta > 0$ we can always choose $x_\pm$, so that $(x_\pm, y_\pm, w_\pm, v_\pm)$ are in the domain D and

$$\Psi(y_\pm, w_\pm, v_\pm) \leq (1 + \delta)\left[B(x_\pm, y_\pm, w_\pm, v_\pm) - Mx_\pm\right].$$

Since the function $(y, v) \mapsto \frac{y^2}{v}$ is convex, the point $x := \frac{1}{2}[x_- + x_+]$ satisfies $x \geq \frac{y^2}{v}$ and, $(x, y, w, v) \in D$, therefore, by the definitions of Ψ, x, and the dyadic convexity of B, for all $\delta > 0$,

$$
\begin{aligned}
\Psi(y, w, v) &\geq B(x, y, w, v) - Mx \\
&\geq \frac{(\Delta y)^2}{w^2} + \frac{1}{2}[B(x_-, y_-, w_-, v_-) - Mx_-] \\
&\quad + \frac{1}{2}[B(x_+, y_+, w_+, v_+) - Mx_+] \\
&\geq \frac{(\Delta y)^2}{w^2} + \frac{(1+\delta)^{-1}}{2}[\Psi(y_-, w_-, v_-) + \Psi(y_+, w_+, v_+)].
\end{aligned}
$$

Let $\delta \to 0$, then

$$
\Psi(y, w, v) \geq \frac{(\Delta y)^2}{w^2} + \frac{1}{2}[\Psi(y_-, w_-, v_-) + \Psi(y_+, w_+, v_+)].
$$

So, indeed, Ψ satisfies the same convexity condition as B and we are done. ∎

7.2. Non-existence of certain Bellman function

To prove that the estimate obtained so far is sharp (in terms of its dependence in Q), it is enough to show that there cannot exist a non-negative function $B(x, y, w, v)$ of the form (7.2) that is bounded from above by $o(Q^4)x$ and such that it satisfies the convexity property (ii) in the domain $w^2 \leq v \leq Q^2 w^2$, $y^2 \leq xv$. Thus, it is enough to show that condition (7.4) is incompatible with $0 \leq \psi \leq o(Q^4)$. We will show here that when ψ is twice differentiable on $[\sqrt{\epsilon}Q^{-1}, 1]$, then the infinitesimal convexity condition (7.5) on the larger domain is incompatible with $0 \leq \psi \leq o(Q^4)$.

Lemma 15. *There does not exist a function $\psi \in C^2([\sqrt{\epsilon}Q^{-1}, 1]$ such that $0 \leq \psi \leq o(Q^4)$ and (7.5) hold on $[\sqrt{\epsilon}Q^{-1}, 1]$.*

Proof . We will first explicitely calculate the Hessian matrix of Ψ, and reduce condition (7.5) to verifying the positive definiteness of a three by three matrix parametrized by $u \in [\sqrt{\epsilon}Q^{-1}, 1]$, with entries depending on ψ, ψ' and ψ''.

As F. Nazarov said: *"A nice thing is that we have here only one variable $u = \frac{w}{\sqrt{v}}$. A bad thing is that we have a 3×3 rather than 2×2 matrix. Fortunately, the matrix turns out to be fairly nice and easy to analyze though the computations are somewhat boring."*

Denoting $\frac{y^2}{v}\psi\left(\frac{w}{\sqrt{v}}\right)$ by $\Psi(y,w,v)$ (notice that in Ψ denoted $-\Psi$), we find

$$\Psi_y = \frac{2y}{v}\psi$$

$$\Psi_v = -\frac{y^2}{v^2}\psi - \frac{1}{2}\frac{y^2 w}{v^{5/2}}\psi'$$

$$\Psi_w = \frac{y^2}{v^{3/2}}\psi'$$

and, therefore

$$\Psi_{yy} = \tfrac{2}{v}\psi \qquad\qquad \Psi_{vv} = \tfrac{2y^2}{v^3}\psi + \tfrac{7}{4}\tfrac{y^2 w}{v^{7/2}}\psi' + \tfrac{1}{4}\tfrac{y^2 w^2}{v^4}\psi''$$

$$\Psi_{yv} = -\tfrac{2y}{v^2}\psi - \tfrac{yw}{v^{5/2}}\psi' \qquad \Psi_{vw} = -\tfrac{3}{2}\tfrac{y^2}{v^{5/2}}\psi' - \tfrac{1}{2}\tfrac{y^2 w}{v^3}\psi''$$

$$\Psi_{yw} = \tfrac{2y}{v^{3/2}}\psi' \qquad\qquad \Psi_{ww} = \tfrac{y^2}{v^2}\psi''$$

Condition (7.5) means that the matrix

$$\begin{pmatrix} \frac{2}{v}\psi - \frac{1}{w^2} & -\frac{2y}{v^2}\psi - \frac{yw}{v^{5/2}}\psi' & \frac{2y}{v^{3/2}}\psi' \\ -\frac{2y}{v^2}\psi - \frac{yw}{v^{5/2}}\psi' & \frac{2y^2}{v^3}\psi + \frac{7}{4}\frac{y^2 w}{v^{7/2}}\psi' + \frac{1}{4}\frac{y^2 w^2}{v^4}\psi'' & -\frac{3}{2}\frac{y^2}{v^{5/2}}\psi' - \frac{1}{2}\frac{y^2 w}{v^3}\psi'' \\ \frac{2y}{v^{3/2}}\psi' & -\frac{3}{2}\frac{y^2}{v^{5/2}}\psi' - \frac{1}{2}\frac{y^2 w}{v^3}\psi'' & \frac{y^2}{v^2}\psi'' \end{pmatrix}$$

must be non-negative definite in the domain.

Multiplying the first row and column by w, the second row and column by $\frac{v^{3/2}}{y}$ and the third row and column by $\frac{v}{y}$, we get the matrix in the variable u, where $u^2 = w^2/v$.

$$\begin{pmatrix} 2u^2\psi - 1 & -2u\psi - u^2\psi' & 2u\psi' \\ -2u\psi - u^2\psi' & 2\psi + \frac{7}{4}u\psi' + \frac{1}{4}u^2\psi'' & -\frac{3}{2}\psi' - \frac{1}{2}u\psi'' \\ 2u\psi' & -\frac{3}{2}\psi' - \frac{1}{2}u\psi'' & \psi'' \end{pmatrix}$$

Now add $u/2$ times the third row to the second row and then $u/2$ times the third column to the second column. We get the matrix

$$\begin{pmatrix} 2u^2\psi - 1 & -2u\psi & 2u\psi' \\ -2u\psi & 2\psi + \frac{1}{4}u\psi' & -\frac{3}{2}\psi' \\ 2u\psi' & -\frac{3}{2}\psi' & \psi'' \end{pmatrix}$$

Now add u times the second row to the first row and then u times the second column to the first column. We obtain the matrix

$$(7.6) \qquad A(u) = \begin{pmatrix} \frac{1}{4}u^3\psi' - 1 & \frac{1}{4}u^2\psi' & \frac{1}{2}u\psi' \\ \frac{1}{4}u^2\psi' & 2\psi + \frac{1}{4}u\psi' & -\frac{3}{2}\psi' \\ \frac{1}{2}u\psi' & -\frac{3}{2}\psi' & \psi'' \end{pmatrix}$$

At this point to verify that a function $\widetilde{B}$ given by (7.2) is a Bellman function for the problem all we need to do is verify that the given function $\psi(u) \in C^2([\sqrt{\epsilon}Q^{-1}, 1])$, $0 \leq \max_{Q^{-1} \leq u \leq 1} \psi(u) \leq M$, and the matrix $A(u)$ is positive definite for each u in $[\sqrt{\epsilon}Q^{-1}, 1]$. This is equivalent to verifying that $\det A(u) \geq 0$, and the 2×2 submatrix

$$A_1(u) = \begin{pmatrix} \frac{1}{4}u^3\psi' - 1 & \frac{1}{2}u\psi' \\ \frac{1}{2}u\psi' & \psi'' \end{pmatrix}$$

must be itself positive definite for all $\sqrt{\epsilon}Q^{-1} \leq u \leq 1$. The matrix A_1 is positive definite if and only if $\psi''(u) \geq 0$ and $\det A_1 = \psi'' \left(\frac{1}{4}u^3\psi' - 1\right) - \frac{1}{4}u^2(\psi')^2 \geq 0$.

To show that there is no twice differentiable function ψ such that $0 \leq \psi \leq o(Q^4)$, and the above matrix A is positive definite, we will proceed by contrapositive. If such function ψ existed, then the 2×2 submatrix A_1 of A must be itself positive definite for all $\sqrt{\epsilon}Q^{-1} \leq u \leq 1$. Looking at A_1 we see that we must have $\psi'' > 0$, moreover, $\psi'(u) \geq 4u^{-3}$, and (from the determinant condition)

(7.7)
$$u\psi''(u) > \psi'(u).$$

Note that since $\psi''(u) > 0$ implies that ψ' is increasing, a lower bound for ψ' must be attained at $u = \sqrt{\epsilon}Q^{-1}$, therefore $\psi'(u) \geq \psi'(\sqrt{\epsilon}Q^{-1}) \geq 4\epsilon^{-3/2}Q^3$. Furthermore the differential inequality (7.7) can be rewritten as $\left(\frac{\psi'(u)}{u}\right)' > 0$, which implies that the function $\frac{\psi'(u)}{u}$ is increasing and it attains its lower bound at $u = \sqrt{\epsilon}Q^{-1}$, hence $\psi'(u) \geq 4\epsilon^{-2}Q^4u$ on $[\sqrt{\epsilon}Q^{-1}, 1]$. From here we can see that ψ itself must be of order at least Q^4 for large Q and we are done. More precisely, integrating the inequality from $\sqrt{\epsilon}Q^{-1}$ to 1, and using the fundamental theorem of calculus, we get

$$\psi(u) \geq \psi(u) - \psi(\sqrt{\epsilon}Q^{-1}) \geq 2\epsilon^{-2}Q^4(u^2 - \epsilon Q^{-2}).$$

That is, $\frac{\psi(u)}{\epsilon^{-2}Q^4} \geq 2u^2 - 2\epsilon Q^{-2}$, and letting now $Q \to \infty$ we reach a contradiction that $2u^2 \leq 0$ since we assumed $\psi(u) = o(Q^4)$. ∎

This trick is used to show that, though the function $4\left(x - \frac{y^2}{1+M}\right)$ is not the true Bellman function for the Carleson embedding theorem, the constant it yields is the best possible [26]. The method of reducing the number of variables for the Bellman function is Burkholder's idea from his paper [6], as it is clearly stated in [26].

8. Concluding Remarks

One can deduce $L^p(w)$ results for the square function S^d by "sharp extrapolation" techniques as described in [12]. For $1 \le p < 2$ the operator norm is of the order $[w]_{A_p}^{1/(p-1)}$ and this is optimal [12, Sec 4.1]. However for $p > 2$ the linear results obtained by extrapolation in [12] are not optimal. Sharper results have been obtained by Lerner [18, 19]. The optimal rate of growth for S^d on $L^p(w)$, $p > 2$ with respect to $[w]_{A_p}$ characteristic is not known yet. We expect similar bound for S_σ on $L^p(vd\sigma)$ extrapolated form the $L^2(vd\sigma)$ norms obtained in this paper. This will be discussed in a future paper.

Theorems 4, 5, 6 and 7, are optimal in the sense that we will not find a better bound that will work for ALL positive doubling measures σ, and weights $v \in A_2^d(d\sigma)$, because it will mean that it will work for the particular case $d\sigma = dx$, $v = w$, for which we have shown no better bound exists.

Lemma 9 (σ-Sawyer's Estimate) was proved using a Bellman function argument. The Bellman function that worked for Lemma 3 (a Sawyer's Estimate in the case $d\sigma = wdx$ and $v = w^{-1}$) worked as well for Lemma 9 in a slightly larger domain. Similarly in the proof of the Weight Lemma 10 the Bellman function needed was the same that worked for the lemma in the case $d\sigma = dx$, which was known to exist because the lemma was known to hold in that case. In both cases, once we had a Bellman function for one case, we could use it for the other, provided we could define it on a slightly larger domain (deformation of the domain or the function dictated by the doubling constant of σ).

The question now is, how far can the following principle be pushed: *Given any theorem for the Lebesgue measure dx that allows for a Bellman function proof "in principle"* (that is regardless of the proof having been obtained by Bellman functions or completely different methods, the fact that the theorem is true guarantees the existence of a Bellman function with certain properties: even if we do not have an explicit formula for such function, we know it exists) *then the proof should be generalizable, in a way similar to how we have done it here for some particular examples, to the case of the doubling measure $d\sigma$.*

References

[1] Astala, K., Iwaniec, T. and Saksman, E.: Beltrami operators in the plane. *Duke Math. J.* **107** (2001), no. 1, 27–56.

[2] Bañuelos, R. and Janakiraman, P.: L^p-bounds for the Beurling-Ahlfors transform. *Trans. Amer. Math. Soc.* **360** (2008), no. 7, 3603–3612.

[3] Beznosova, O.: Linear bound for the dyadic paraproduct on weighted Lebesgue space $L^2(w)$. *J. Funct. Anal.* **255** (2008), no. 4, 994–1007.

[4] Buckley, S.: Estimates for operator norms and reverse Jensen's inequalities. *Trans. Amer. Math. Soc.* **340** (1993), no. 1, 253–272.

[5] Buckley, S.: Summation conditions on weights. *Michigan Math. J.* **40** (1993), no. 1, 153–170.

[6] Burkholder, D. L.: Explorations in martingale theory and its applications. In *Ecole d'Eté de Probabilité de Saint-Flour XIX-1989*, 1–66. Lecture Notes in Math. **1464**. Springer, Berlin, 1991.

[7] Coifman, R. R. and Fefferman, C.: Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.* **51** (1974), 241–250.

[8] Coifman, R. R., Jones, P. W. and Semmes, S.: Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves. *J. Amer. Math. Soc.* **2** (1989), no. 3, 553–564.

[9] Dindos, M., Petermichl, S. and Pipher, J.: The L^p Dirichlet problem for second order elliptic operators and a p-adapted square function. *J. Funct. Anal.* **249** (2007), no. 2, 372–392.

[10] Dragičević, O.: *Riesz transforms and the Bellman function technique.* PhD Thesis. Michigan State University, 2003.

[11] Dragičević, O. and Volberg, A.: Sharp estimate of the Ahlfors–Beurling operator via averaging martingale transforms. *Michigan Math. J.* **51** (2003), no. 2, 415–435.

[12] Dragičević, O., Grafakos, L., Pereyra, M. C. and Petermichl, S.: Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Mat.* **49** (2005), no. 1, 73–91.

[13] Fefferman, R., Kenig, C. and Pipher, J.: The theory of weights and the Dirichlet problem for elliptic equations. *Ann. of Math. (2)* **134** (1991), no. 1, 65–124.

[14] Fefferman, R. and Pipher, J.: Multiparameter operators and sharp weighted inequalities. *Amer. J. Math.* **119** (1997), no. 2, 337–369.

[15] Gehring, F. W.: The L^p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.* **130** (1973), 265–277.

[16] Hukovic, S., Treil, S. and Volberg, A.: The Bellman function and sharp weighted inequalities for square functions. In *Complex analysis, operators and related topics*, 97–113. Oper. Theory Adv. Appl. **113**. Birkhäuser, Basel, 2000.

[17] Katz, N. H. and Pereyra, M. C.: Haar multipliers, paraproducts and weighted inequalities. In *Analysis of Divergence (Orono, ME, 1997)*, 145–170. Appl. Numer. Harmon. Anal. Birkhäuser, Boston, MA, 1999.

[18] Lerner, A. K.: On some sharp weighted norm inequalities. *J. Funct. Anal.* **232** (2006), no. 2, 477–494.

[19] LERNER, A. K.: On some weighted norm inequalities for Littlewood-Paley operators. To appear in *Illinois J. Math.*

[20] MOEN, K.: *Linear and multilinear fractional operators: weighted inequalities, sharp bounds and other properties.* PhD Thesis. University of Kansas, 2009.

[21] MUCKENHOUPT, B.: Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.* **165** (1972), 207–226.

[22] MURPHY, G.: C^*-algebras and operator theory. Academic Press, Boston, MA, 1990.

[23] NAZAROV, F.: Personal communications, 2004.

[24] NAZAROV, F. AND TREIL, S.: The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. *Algebra i Analiz* **8** (1996), no. 5, 32–162; translation in *St. Petersburg Math. J.* **8** (1997), no. 5, 721–824.

[25] NAZAROV, F., TREIL, S. AND VOLBERG, A.: The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.* **12** (1999), no. 4, 909–928.

[26] NAZAROV, F., TREIL, S. AND VOLBERG, A.: Bellman function in stochastic control and harmonic analysis. In *Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000)*, 393–423. Oper. Theory Adv. Appl. **129**. Birkhäuser, Basel, 2001.

[27] PANEK, D.: *On sharp extrapolation theorems.* PhD Thesis. University of New Mexico, 2008.

[28] PEREYRA, M. C.: On the resolvent of dyadic paraproducts. *Rev. Mat. Iberoamericana* **10** (1994), no. 3, 627–664.

[29] PEREYRA, M. C.: Lecture notes on dyadic harmonic analysis. In *Second Summer School in Analysis and Mathematical Physics (Cuernavaca, 2000)*, 1–60. Contemp. Math. **289**. Amer. Math. Soc., Providence, RI, 2001.

[30] PÉREZ, C.: *Weighted norm inequalities for potential maximal operators.* PhD Thesis. Washington University, 1989.

[31] PETERMICHL, S.: The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p-characteristic. *Amer. J. Math.* **129** (2007) no. 5, 1355–1375.

[32] PETERMICHL, S.: The sharp weighted bound for the Riesz transforms. *Proc. Amer. Math. Soc.* **136** (2008), no. 4, 1237–1249.

[33] PETERMICHL, S. AND POTT, S.: An estimate for weighted Hilbert transform via square functions. *Trans. Amer. Math. Soc.* **354** (2002), no. 4, 1699–1703.

[34] PETERMICHL, S. AND VOLBERG, A.: Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular. *Duke Math. J.* **112** (2002), no. 2, 281–305.

[35] Petermichl, S. and Wittwer, J.: A sharp estimate for the weighted Hilbert transform via Bellman functions. *Michigan Math. J.* **50** (2002), no. 1, 71–87.

[36] Volberg, A. and Nazarov, F.: Heat extension of the Beurling operator and estimates for its norm. *Algebra i Analiz* **15** (2003), no. 4, 142–158; translation in *St. Petersburg Math. J.* **15** (2004), no. 4, 563–573.

[37] Wittwer, J.: A sharp estimate on the norm of the martingale transform. *Math. Res. Lett.* **7** (2000), no. 1, 1–12.

[38] Wittwer, J.: A sharp estimate on the norm of the continuous square function. *Proc. Amer. Math. Soc.* **130** (2002), no. 8, 2335–2342.

María Cristina Pereyra
Department of Mathematics and Statistics
MSC03 2150
1 University of New Mexico
Albuquerque, NM 87131-0001
`crisp@math.unm.edu`