

THE BELLMAN FUNCTIONS AND TWO-WEIGHT INEQUALITIES FOR HAAR MULTIPLIERS

F. NAZAROV, S. TREIL, AND A. VOLBERG

0. INTRODUCTION

Let T_0 be an operator defined on some class of “sufficiently good” functions $f : \mathbb{R} \rightarrow \mathbb{R}$. Let u, v be two (almost) everywhere positive locally integrable functions (from now on, we shall call such functions “weights”). The question we are going to discuss is

When is the operator T_0 bounded from $L^2(u)$ to $L^2(v)$?

or, in other words,

When does the inequality

$$\int_{\mathbb{R}} |T_0 f|^2 v \leq C \int_{\mathbb{R}} |f|^2 u$$

hold with some constant C independent of f ?

(Unless otherwise specified, all integrals are taken with respect to the standard Lebesgue measure on \mathbb{R} .)

Denoting $w := u^{-1}$, we can reformulate the above question as follows:

When is the operator $T := M_{\sqrt{v}} T_0 M_{\sqrt{w}}$ bounded in L^2 ?

(Here M_φ stands for the operator of multiplication by φ .)

Such weighted norm inequalities arise naturally in many areas of analysis, operator theory (including the perturbation of self-adjoint operators), and probability theory.

The one-weight case is now pretty well understood for many interesting operators T_0 . For the Hilbert transform $Hf(t) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(s)}{t-s} ds$ (defined as the principal value $\lim_{\varepsilon \rightarrow 0} \frac{1}{\pi} \int_{\mathbb{R} \setminus (t-\varepsilon, t+\varepsilon)} \frac{f(s)}{t-s} ds$) the answer is given by the famous theorems of Helson–Szegö and Hunt–Muckenhoupt–Wheeden.

The Helson–Szegö theorem states that the Hilbert transform is bounded in $L^2(v)$ if and only if the weight v can be represented as $v = \exp\{\varphi + H\psi\}$, where $\varphi, \psi \in L^\infty$, $\|\psi\|_{L^\infty} < \pi/2$.

The Hunt–Muckenhoupt–Wheeden theorem states that the Hilbert transform is bounded in $L^2(v)$ if and only if the weight v satisfies the so-called Muckenhoupt

Received by the editors December 31, 1997.

1991 *Mathematics Subject Classification*. Primary 42B20, 42A50, 47B35.

This work was partially supported by NSF grant DMS 9622936, the joint Israeli–USA grant BSF 00030, and MSRI programs of the Fall 1995 and the Fall 1997.

A_2 condition

$$(A_2) \quad \sup_I \left(\frac{1}{|I|} \int_I v \right) \cdot \left(\frac{1}{|I|} \int_I v^{-1} \right) < \infty,$$

where the supremum is taken over all intervals $I \subset \mathbb{R}$.

It is worth mentioning that no direct proof of the equivalence of the Helson–Szegő condition and the Muckenhoupt A_2 condition is known.

The same Muckenhoupt A_2 condition is necessary and sufficient for boundedness of the maximal function operator

$$\mathcal{M}f(x) := \sup_{I \ni x} \frac{1}{|I|} \int_I |f|$$

(here the supremum is taken over all intervals I containing x) and sufficient for boundedness of all classical Calderón–Zygmund operators in $L^2(v)$.

Actually, the maximal operator M plays a very important part in the theory. For example, one of the proofs of the Hunt–Muckenhoupt–Wheeden theorem uses the above-mentioned fact that the A_2 condition characterizes the boundedness of \mathcal{M} in $L^2(w)$. Then the inequality $\|Hf\|_{L^2(w)} \leq C \|\mathcal{M}f\|_{L^2(w)}$ is used, which holds, for example, for $A_\infty = \bigcup_p A_p$ weights (see [CF] for this result) or for C_q ($q > p$) weights of Muckenhoupt and Sawyer (see [Sa]). Let us remind the reader that A_∞ is the class of weights satisfying the inverse Jensen inequality (many amazing properties of such weights can be found in [St]):

$$(A_\infty) \quad \sup_I \left(\frac{1}{|I|} \int_I w \right) \cdot \left(\exp \left(-\frac{1}{|I|} \int_I \log w \right) \right) < \infty,$$

and C_q is defined as follows: There are positive constants C, ε such that

$$(C_q) \quad \int_E w \leq C \left(\frac{|E|}{|I|} \right)^\varepsilon \int |M\chi_I|^q w$$

whenever E is a subset of an interval I .

Things look much more complicated in the two-weight case. Though there exists a beautiful two-weight analogue of the Helson–Szegő theorem due to Cotlar and Sadosky (see [CS1]), the consensus now is that there can be no explicit (Muckenhoupt type) necessary and sufficient condition for the boundedness of the Hilbert transform.

It was a big surprise when Eric Sawyer [S1] found a necessary and sufficient condition for boundedness of the maximal function operator \mathcal{M} from $L^2(u)$ to $L^2(v)$. His theorem states that it is enough to test the boundedness on a very special class of “simple” test functions, namely on the functions $u^{-1}\chi_I$ (here and below χ_I stands for the characteristic function of an interval $I \subset \mathbb{R}$). Later (see [S2]) Sawyer presented a necessary and sufficient condition of similar type for boundedness from $L^2(u)$ to $L^2(v)$ of certain integral operators T_0 with positive kernels. The main difference from the case of the maximal function operator was that now one should also check boundedness of the adjoint operator T_0^* on an appropriate class of simple test-functions (see more about that in Section 2).

Everywhere below we shall call conditions of such kind *Sawyer type* conditions. The main aim of this paper is to present a uniform approach that allows us to recover many of Sawyer’s results (at least, in L^2) and to obtain at least one new theorem of Sawyer’s type: the necessary and sufficient condition for boundedness of Haar multipliers. As far as we know, it is the first theorem of Sawyer’s type for

singular integral operators: all previous results concerned operators with positive kernels only.

The paper is organized as follows: the Sawyer type theorem on Haar multipliers is formulated and proved in Section 4. The reader who wants to go “straight to the point” should keep in mind that some results from previous sections will be used there without explicit reference. In Section 1 the weighted version of the dyadic Carleson imbedding theorem is discussed. Section 2 deals with the dyadic version of the classical Sawyer’s theorem on operators with positive kernels. Section 3 concerns the Sawyer type theorem on the (dyadic) square function operator.

We will restrict ourselves to the L^2 -case exclusively. Our methods allow us to treat the L^p -case as well, but since the main result of our paper — Sawyer type theorem for Haar multipliers — fails in any L^p with $p \neq 2$, we decided not to include the corresponding pieces. The reader may find it a (quite challenging, but definitely doable) exercise to find the “Bellman function proofs” of the L^p -versions of the Weighted Carleson Imbedding Theorem and of the classical Sawyer’s theorem on operators with positive kernels and to figure out what happens in the case of the square function operator.

The reader may wonder what happens for the Hilbert transform H ? What if one just adopts all our proofs for the Haar multipliers to the case of $M_{\sqrt{v}}HM_{\sqrt{w}}$? It turns out that the proofs will work fine and will give a certain list of sufficient conditions for $M_{\sqrt{v}}HM_{\sqrt{w}}$ to be bounded. The reader can find these conditions in [NTV1] or in the manuscript “The Bellman functions and two weight inequalities for Haar multipliers” on the homepage <http://www.math.msu.edu/~volberg>. However these conditions are most probably not necessary.

1. WEIGHTED CARLESON IMBEDDING THEOREM

We shall start with a statement of the weighted version of the dyadic Carleson imbedding theorem. It is worth mentioning at this point that Sawyer’s theorem on the maximal function operator is an easy consequence of the Weighted Carleson Imbedding Theorem (see [SW] for details). What may be interesting (and, perhaps, new) to the reader in this section is not the result itself, but our way of proving it.

Everywhere below we will use the notation $\langle f \rangle_I$ for the average $\frac{1}{|I|} \int_I f$ of the function $f : \mathbb{R} \rightarrow \mathbb{R}$ over the interval $I \subset \mathbb{R}$.

Theorem (Weighted Carleson Imbedding Theorem). *Let \mathcal{D} be a dyadic lattice. Let w be any weight. Let, at last, $\{\alpha_J\}_{J \in \mathcal{D}}$ be a sequence of non-negative numbers. The imbedding*

$$\sum_{J \in \mathcal{D}} \alpha_J \langle f \sqrt{w} \rangle_J^2 \leq C \|f\|_{L^2}^2$$

holds for all $f \in L^2$ if and only if for all $I \in \mathcal{D}$,

$$(SC) \quad \frac{1}{|I|} \sum_{J \subset I} \alpha_J \langle w \rangle_J^2 \leq C \langle w \rangle_I.$$

Proof. The necessity of the Sawyer–Carleson condition (SC) for the imbedding is almost obvious: just consider $f = \sqrt{w} \chi_J$ and take the sum only over the intervals $J \subset I$. The non-trivial part is the sufficiency.

Without loss of generality, one can assume that $f \geq 0$ and that the constant C in the Sawyer–Carleson condition equals 1. It is enough to show that for any

dyadic interval I ,

$$\sum_{J \in \mathcal{D}, J \subset I} \alpha_J \langle f\sqrt{w} \rangle_J^2 \leq C \int_I f^2.$$

We do not know anything about the sequence α_J except the condition (CS). So, let us treat this sequence as an arbitrary one and consider the worst case scenario.

Namely, fix a dyadic interval I and fix the averages of all unknowns involved:

$$\langle w \rangle_I = \mathbf{w}, \quad \langle f\sqrt{w} \rangle_I = \mathbf{f}, \quad \langle f^2 \rangle_I = \mathbf{F}, \quad \frac{1}{|I|} \sum_{J \subset I} \alpha_J \langle w \rangle_J^2 = M.$$

Define the function \mathcal{B} of 4 real variables by

$$\mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M) := \sup \frac{1}{|I|} \sum_{J \in I} \alpha_J \langle f\sqrt{w} \rangle_J^2,$$

where the supremum is taken over all possible choices of $f \geq 0$, w , $\{\alpha_J\}_{J \in \mathcal{D}}$ with fixed averages and such that for all dyadic intervals K ,

$$\frac{1}{|K|} \sum_{J \subset K} \alpha_J \langle w \rangle_J^2 \leq \langle w \rangle_K.$$

Note that the function \mathcal{B} does not depend on the choice of the interval I (this is essentially just the observation that a linear change of variable does not alter the average value of a function over an interval). The function \mathcal{B} is called the *Bellman function* associated with the Weighted Carleson Imbedding Theorem.

1.1. Properties of the Bellman function.

- (1) The domain of \mathcal{B} consists of all non-negative quadruples $(\mathbf{F}, \mathbf{f}, \mathbf{w}, M)$ such that

$$\mathbf{f}^2 \leq \mathbf{F}\mathbf{w}, \quad M \leq \mathbf{w}.$$

- (2) For any $\mathbf{F}, \mathbf{f}, \mathbf{w}, M$ in the domain, $0 \leq \mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M) \leq C\mathbf{F}$.
- (3) For any three quadruples $(\mathbf{F}, \mathbf{f}, \mathbf{w}, M)$, $(\mathbf{F}_\pm, \mathbf{f}_\pm, \mathbf{w}_\pm, M_\pm)$ in the domain satisfying

$$\mathbf{F} = \frac{\mathbf{F}_- + \mathbf{F}_+}{2}, \quad \mathbf{f} = \frac{\mathbf{f}_- + \mathbf{f}_+}{2}, \quad \mathbf{w} = \frac{\mathbf{w}_- + \mathbf{w}_+}{2},$$

$$M = m + \frac{M_- + M_+}{2}, \quad m \geq 0,$$

the following inequality holds:

$$(*) \quad \mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M) - \frac{1}{2} \{ \mathcal{B}(\mathbf{F}_+, \mathbf{f}_+, \dots) + \mathcal{B}(\mathbf{F}_-, \mathbf{f}_-, \dots) \} \geq \frac{\mathbf{f}^2}{\mathbf{w}^2} m.$$

Let us discuss these properties a little bit. The inequality $\mathbf{f}^2 \leq \mathbf{F}\mathbf{w}$ is just the Cauchy-Schwartz inequality:

$$\left(\frac{1}{|I|} \int_I f\sqrt{w} \right)^2 \leq \left(\frac{1}{|I|} \int_I w \right) \cdot \left(\frac{1}{|I|} \int_I f^2 \right) = \langle w \rangle_I \langle f^2 \rangle_I,$$

the inequality $\mathcal{B} \geq 0$ follows from the definition of \mathcal{B} , and the inequality $\mathcal{B} \leq C\mathbf{F}$ must hold if the theorem is true.

The property (*), which we will call *the main inequality*, can be obtained by analyzing how the averages are distributed between the right half I_+ and the left half I_- of the interval I . Actually, (*) means that the supremum taken over all

functions with fixed averages on interval I is larger than the supremum taken over all functions with averages fixed on I_+ and I_- in such a way that their arithmetic average equals the average over I . (The former set of functions just include the latter one.) Inequality (*) also shows how much larger the first supremum is; see [NT].

We will not prove this property here because, in fact, we do not need to *prove* it to obtain the theorem! To *observe* it is enough for our purposes.

1.2. The Bellman function proves the theorem. The trick is that to prove the theorem, it suffices to find *any* function satisfying the above properties (1)–(3)!

Indeed pick f, w and $\{\alpha_J\}_{J \in \mathcal{D}}$. For every dyadic interval $J \in \mathcal{D}$, let $\mathbf{F}_J, \mathbf{f}_J, \mathbf{w}_J, M_J$ be the corresponding averages:

$$\mathbf{F}_J = \langle f^2 \rangle_J, \quad \mathbf{f}_J = \langle f\sqrt{w} \rangle_J, \quad \mathbf{w}_J = \langle w \rangle_J, \quad M_J = \frac{1}{|J|} \sum_{K \subset J} \alpha_K \mathbf{w}_K^2.$$

Note that

$$M_J = \frac{1}{2}(M_{J_+} + M_{J_-}) + |J|^{-1} \alpha_J \mathbf{w}_J^2.$$

For the interval J , the main inequality (*) implies

$$\alpha_J \mathbf{f}_J^2 \leq |J| \cdot \mathcal{B}(\mathbf{F}_J, \mathbf{f}_J, \dots) - |J_+| \cdot \mathcal{B}(\mathbf{F}_{J_+}, \mathbf{f}_{J_+}, \dots) - |J_-| \cdot \mathcal{B}(\mathbf{F}_{J_-}, \mathbf{f}_{J_-}, \dots).$$

Now, if we take the sum of the above inequalities for all $J \subset I, |J| > 2^{-k}|I|$, we shall get

$$\begin{aligned} \sum_{\substack{J \subset I \\ |J| > 2^{-k}|I|}} \alpha_J \mathbf{f}_J^2 &\leq |I| \mathcal{B}(\mathbf{F}_I, \mathbf{f}_I, \mathbf{w}_I, M_I) - \sum_{\substack{J \subset I \\ |J| = 2^{-k}|I|}} |J| \mathcal{B}(\mathbf{F}_J, \mathbf{f}_J, \mathbf{w}_J, M_J) \\ &\leq |I| \mathcal{B}(\mathbf{F}_I, \mathbf{f}_I, \mathbf{w}_I, M_I) \leq C|I| \mathbf{F}_I. \end{aligned}$$

Passing to the limit as $k \rightarrow \infty$ and recalling what \mathbf{f}_J and \mathbf{F}_I mean, we obtain

$$\sum_{J \subset I} \alpha_J \langle f\sqrt{w} \rangle_J^2 \leq C|I| \langle f^2 \rangle_I,$$

which is exactly what we need.

1.3. Finding a Bellman function. Now our task is to find a function \mathcal{B} satisfying properties (1)–(3). We will look for a function satisfying the infinitesimal version of the main inequality (*), namely, for a function satisfying (1), (2) and

$$(3') \quad \text{The function } -\mathcal{B} \text{ is convex, i.e., } d^2\mathcal{B} \leq 0.$$

$$(3'') \quad \frac{\partial \mathcal{B}}{\partial M} \geq \frac{\mathbf{f}^2}{\mathbf{w}^2}.$$

Since the domain of \mathcal{B} is convex, properties (3') and (3'') imply the main inequality (*):

$$\begin{aligned} m \frac{\mathbf{f}^2}{\mathbf{w}^2} &\leq \mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M) - \mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M - m) \\ &\leq \mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M) - \frac{1}{2} \left\{ \mathcal{B}(\mathbf{F}_+, \mathbf{f}_+, \mathbf{w}_+, M_+) + \mathcal{B}(\mathbf{F}_-, \mathbf{f}_-, \mathbf{w}_-, M_-) \right\}. \end{aligned}$$

The first inequality follows from (3''), and the second one is just the convexity condition (3').

Now we shall just present a possible function \mathcal{B} , leaving it to the reader (as an easy exercise) to check the properties (2), (3') and (3'') in the domain given by (1):

$$\mathcal{B}(\mathbf{F}, \mathbf{f}, \mathbf{w}, M) = 4\left(\mathbf{F} - \frac{\mathbf{f}^2}{\mathbf{w} + M}\right).$$

□

2. SAWYER TYPE THEOREM FOR OPERATORS WITH POSITIVE KERNELS (DYADIC VERSION)

Let us first remind the reader of the classical Sawyer's theorem on operators with "good" positive kernels.

Theorem. For $t, s \in \mathbb{R}$, denote by $I(t, s)$ the interval of length $2|t - s|$ centered at the point $\frac{t+s}{2}$. Let $K(t, s)$ be a positive kernel such that for every $t, s \in \mathbb{R}$ and for every $t', s' \in I(t, s)$,

$$K(t', s') \geq cK(t, s)$$

with some $c > 0$ independent of t and s . Let T_0 be the integral operator with kernel $K(t, s)$, i.e.,

$$T_0 f(t) = \int_{\mathbb{R}} K(t, s) f(s) ds.$$

Let, at last, v, w be two weights.

The operator $T = M_{\sqrt{v}} T_0 M_{\sqrt{w}}$ is bounded in L^2 if and only if for every interval $I \subset \mathbb{R}$,

1. $\int_I |T_0(w \chi_I)|^2 v \leq C \int_I w;$
2. $\int_I |T_0^*(v \chi_I)|^2 w \leq C \int_I v,$

with some constant $C > 0$ independent of I .

This version of Sawyer's theorem is due to Verbitsky. In the original version proved by Sawyer the integrals on the left hand side in Conditions 1 and 2 were taken over the whole real line \mathbb{R} , i.e., the first condition was just the boundedness of the operator T on the test-functions $\sqrt{w} \chi_I$, while the second condition was the boundedness of the adjoint operator T^* on the test-functions $\sqrt{v} \chi_I$.

In this section we shall present a proof of the dyadic version of Sawyer's theorem, namely, we shall fix some dyadic lattice \mathcal{D} and consider the operator

$$T_0 f = \sum_{J \in \mathcal{D}} \alpha_J \langle f \rangle_J \chi_J$$

where $\{\alpha_J\}_{J \in \mathcal{D}}$ is any family of non-negative numbers. The result is exactly the same as that of the classical Sawyer's theorem except now it is enough to check Conditions 1 and 2 for dyadic intervals $I \in \mathcal{D}$ only. It is worth mentioning at this point that the classical Sawyer theorem can be easily derived from our dyadic one, but, as far as we know, not vice versa.

The boundedness of the operator $T := M_{\sqrt{v}} T_0 M_{\sqrt{w}}$ is equivalent to the bilinear inequality

$$\sum_{J \in \mathcal{D}} \alpha_J \langle f \sqrt{w} \rangle_J \langle g \sqrt{v} \rangle_J |J| \leq C \|f\|_{L^2} \|g\|_{L^2}.$$

We shall prove the following

Theorem (Bilinear Imbedding Theorem). **(A)** *Under the conditions of Sawyer’s theorem, the above bilinear inequality holds for all non-negative functions $f, g \in L^2$.*

(B) *What is more, each of Conditions 1 and 2 is responsible for the boundedness of a certain half of the sum. Namely, let*

$$\mathcal{D}_1 = \left\{ J \in \mathcal{D} : \frac{\langle f\sqrt{w} \rangle_J^2}{\langle f^2 \rangle_J \langle w \rangle_J} \geq \frac{\langle g\sqrt{v} \rangle_J^2}{\langle g^2 \rangle_J \langle v \rangle_J} \right\}$$

(note that the set \mathcal{D}_1 of dyadic intervals does depend on the functions f and g as well as on the weights w and v). Then Condition 1 is necessary and sufficient for the estimate

$$\sum_{J \in \mathcal{D}_1} \alpha_J \langle f\sqrt{w} \rangle_J \langle g\sqrt{v} \rangle_J |J| \leq C \|f\|_{L^2} \|g\|_{L^2}.$$

Note that the ratio $\frac{\langle f\sqrt{w} \rangle_J^2}{\langle f^2 \rangle_J \langle w \rangle_J}$ never exceeds 1 and can be equal to 1 only if the function f is a multiple of \sqrt{w} on the interval J . So, roughly speaking, Condition 1 is responsible for the estimate for the sum over those intervals $J \in \mathcal{D}$ on which f is closer to \sqrt{w} than g is to \sqrt{v} . This seems to make quite good sense (at least, philosophically). As far as we know, this elegant “separation” of the conditions has somehow been completely overlooked in the literature and may be considered the second new result of this paper.

Proof. The necessity of Condition 1 is almost obvious. Indeed, for a fixed interval $I \in \mathcal{D}$, consider the test function $f := \sqrt{w} \chi_I$. Let $g \in L^2$ be any function supported by the interval I . Since no such function can be closer to \sqrt{w} than $\sqrt{w} \chi_I$ on any interval $J \subset \mathbb{R}$ in the sense of the inequality in the definition of \mathcal{D}_1 , we have $\mathcal{D}_1 = \mathcal{D}$. Therefore,

$$(Tf, g)_{L^2} = \sum_{J \in \mathcal{D}, J \subset I} \langle f\sqrt{w} \rangle_J \langle g\sqrt{v} \rangle_J \alpha_J \leq C \|f\|_{L^2} \|g\|_{L^2}.$$

Since g was an arbitrary L^2 -function supported by I , we can conclude from here that $\int_I |Tf|^2 \leq C \|f\|_{L^2}^2$, which is exactly Condition 1. So, again, the only non-trivial part is sufficiency.

2.1. Young versus Hölder. We are going to prove the following lemma first. Let, as before, $\{\alpha_J\}_{J \in \mathcal{D}}$ be a sequence of non-negative numbers. For every interval $I \in \mathcal{D}$, define the truncated operator $T_0^{(I)} f := \sum_{J \subset I} \alpha_J \langle f \rangle_J \chi_J$. We obviously have $T_0^{(I)}(f \chi_I) = T_0^{(I)} f \leq T_0 f$ for any non-negative function f .

Lemma. *Assume that the weights v and w satisfy the condition*

$$1'. \quad M_I := \frac{1}{|I|} \int_I [T_0^{(I)} w]^2 v \leq C \langle w \rangle_I$$

for every $I \in \mathcal{D}$. (This is just a slightly relaxed Sawyer’s condition 1.) Then for any two non-negative functions $f, g \in L^2$,

$$\sum_{J \in \mathcal{D}_1} \alpha_J \langle f\sqrt{w} \rangle_J \langle g\sqrt{v} \rangle_J |J| \leq C (\|f\|_{L^2}^2 + \|g\|_{L^2}^2)$$

where \mathcal{D}'_1 is the set of all dyadic intervals $J \in \mathcal{D}$ on which

$$\frac{\langle f\sqrt{w} \rangle_J^2}{\langle w \rangle_J} \geq \frac{\langle g\sqrt{v} \rangle_J^2}{\langle v \rangle_J}.$$

This lemma is already good enough to imply part (A) of the Bilinear Imbedding Theorem. Indeed, using both Conditions 1 and 2 (or even their relaxed versions), we can estimate the sum over all dyadic intervals:

$$\sum_{J \in \mathcal{D}} \alpha_J \langle f\sqrt{w} \rangle_J \langle g\sqrt{v} \rangle_J |J| \leq C(\|f\|_{L^2}^2 + \|g\|_{L^2}^2).$$

Let t be any positive number. Replacing f by tf and g by $t^{-1}g$, we obtain

$$\sum_{J \in \mathcal{D}} \alpha_J \langle f\sqrt{w} \rangle_J \langle g\sqrt{v} \rangle_J |J| \leq C(t^2\|f\|_{L^2}^2 + t^{-2}\|g\|_{L^2}^2).$$

Taking the infimum of the right hand part over $t \in (0, \infty)$, we get

$$\sum_{J \in \mathcal{D}} \alpha_J \langle f\sqrt{w} \rangle_J \langle g\sqrt{v} \rangle_J |J| \leq 2C\|f\|_{L^2}\|g\|_{L^2}.$$

On the other hand, the separation of conditions provided by this lemma is far less satisfactory than that provided by part (B) of the Bilinear Imbedding Theorem: firstly, the ratios $\frac{\langle f\sqrt{w} \rangle_J}{\langle w \rangle_J}$ do not seem to have any “physical sense”, and secondly, now it is not at all clear why (or, even, whether) Condition 1 is necessary for the estimate for the sum over \mathcal{D}'_1 . Why then are we willing to switch to this Young type lemma? The reason is that it is much easier to find an appropriate Bellman function for it than for part (B) of the Bilinear Imbedding Theorem.

2.2. Hidden variable of the Bellman function. We shall try to prove the lemma in the same way as for the Weighted Carleson Imbedding Theorem from the previous section. First of all, it would be nice to discern the variables of the future Bellman function. Some of them are obvious. These are the averages

$$\begin{aligned} \mathbf{F}_J &= \langle f^2 \rangle_J, & \mathbf{f}_J &= \langle f\sqrt{w} \rangle_J, & \mathbf{w}_J &= \langle w \rangle_J, \\ \mathbf{G}_J &= \langle g^2 \rangle_J, & \mathbf{g}_J &= \langle g\sqrt{v} \rangle_J, & \mathbf{v}_J &= \langle v \rangle_J, \end{aligned}$$

and the variable M_J from the Relaxed Sawyer Condition 1'. Nothing else is ready right away, so it seems to be a good time to pass to the next step and to establish the dynamics of our variables, i.e., to see how the values corresponding to the interval J are related to those corresponding to its left and right halves J_- and J_+ .

First six variables (denoted by **bold** letters) have trivial *martingale* dynamics, i.e., $\mathbf{F}_J = \frac{1}{2}(\mathbf{F}_{J_-} + \mathbf{F}_{J_+})$, $\mathbf{f}_J = \frac{1}{2}(\mathbf{f}_{J_-} + \mathbf{f}_{J_+})$, etc., but we get a serious problem with M_J . Like the variable M_J from the proof of the Weighted Carleson Imbedding Theorem, it has *supermartingale* dynamics, that means, $M_J \geq \frac{1}{2}(M_{J_-} + M_{J_+})$. The problem is that now it is impossible to express the difference $M_J - \frac{1}{2}(M_{J_-} + M_{J_+})$ in terms of our seven variables. Let us take a closer look at this difference. On each half J_\pm of the interval J , we can write $T_0^{(J_\pm)}w = T_0^{(J)}w - \alpha_J \langle w \rangle_J$. Using the

inequality $A^2 - (A - a)^2 \geq aA$ ($0 \leq a \leq A$), we obtain

$$\begin{aligned} M_J - \frac{1}{2}(M_{J_+} + M_{J_-}) &= \frac{1}{|J|} \int_J \left[(T_0^{(J)} w)^2 - (T_0^{(J)} w - \alpha_J \langle w \rangle_J)^2 \right] v \geq \alpha_J \langle w \rangle_J \frac{1}{|J|} \int_J (T_0^{(J)} w) v. \end{aligned}$$

So, if we introduce a new variable

$$N_J := \frac{1}{|J|} \int_J (T_0^{(J)} w) v,$$

we can write

$$M_J - \frac{1}{2}(M_{J_+} + M_{J_-}) \geq \alpha_J \langle w \rangle_J N_J.$$

Note that in fact this inequality is almost an identity: since we also have $A^2 - (A - a)^2 \leq 2aA$ for all $0 \leq a \leq A$, we can conclude that

$$M_J - \frac{1}{2}(M_{J_+} + M_{J_-}) \leq 2\alpha_J \langle w \rangle_J N_J.$$

Actually we shall need the estimate from below only, but it is encouraging to know that all we lose during this step is just a constant factor.

Now we are done with the dynamics of M_J , but how about the dynamics of N_J ? Fortunately, we do not get any problem with that one: a straightforward computation shows that our new quantity N_J also has supermartingale dynamics and

$$N_J - \frac{1}{2}(N_{J_+} + N_{J_-}) = \alpha_J \mathbf{w}_J \mathbf{v}_J.$$

So, we do not need any more variables and we may start

2.3. Guessing a Bellman function for the problem. We are looking for a function \mathcal{B} of 8 real variables $\mathbf{F}, \mathbf{f}, \mathbf{w}, \mathbf{G}, \mathbf{g}, \mathbf{v}, M, N$. The domain of \mathcal{B} should be the set of all non-negative octets $X = (\mathbf{F}, \mathbf{f}, \dots, N)$ satisfying

$$\mathbf{f}^2 \leq \mathbf{F}\mathbf{w}, \quad \mathbf{g}^2 \leq \mathbf{G}\mathbf{v}, \quad M \leq \mathbf{w}.$$

The first 2 inequalities are just Cauchy–Schwartz estimates:

$$\mathbf{f}^2 = \left(\frac{1}{|J|} \int_J f \sqrt{w} \right)^2 \leq \left(\frac{1}{|J|} \int_J f^2 \right) \left(\frac{1}{|J|} \int_J w \right) = \mathbf{F}\mathbf{w},$$

and similarly for \mathbf{g} . The inequality $M \leq \mathbf{w}$ is the Relaxed Sawyer Condition 1' with the constant $C = 1$ (the general case can be easily reduced to this special one by division of one of the weights v or w by C). We could also add the inequality $N^2 \leq M\mathbf{v}$ to this list (one more Cauchy–Schwartz), but, as it turns out, we shall not need it anywhere below.

We also want our function \mathcal{B} to satisfy the estimates

$$0 \leq \mathcal{B}(X) \leq C(\mathbf{F} + \mathbf{G}).$$

Finally, we want our function to satisfy the *main inequality*,

$$(**) \quad \mathcal{B}(X) - \frac{1}{2}(\mathcal{B}(X_+) + \mathcal{B}(X_-)) \geq \begin{cases} \gamma \alpha \mathbf{f}\mathbf{g}, & \text{if } \frac{\mathbf{f}^2}{\mathbf{w}} \geq \frac{\mathbf{g}^2}{\mathbf{v}}; \\ 0, & \text{if } \frac{\mathbf{f}^2}{\mathbf{w}} < \frac{\mathbf{g}^2}{\mathbf{v}}, \end{cases}$$

with some absolute constant $\gamma > 0$. This inequality should hold for all octets $X = (\mathbf{F}, \mathbf{f}, \dots)$, $X_{\pm} = (\mathbf{F}_{\pm}, \mathbf{f}_{\pm}, \dots)$ from the domain and numbers $\alpha \geq 0$ satisfying the relations

$$\mathbf{F} = \frac{1}{2}(\mathbf{F}_- + \mathbf{F}_+), \quad \mathbf{f} = \frac{1}{2}(\mathbf{f}_- + \mathbf{f}_+), \quad \dots, \quad \mathbf{v} = \frac{1}{2}(\mathbf{v}_- + \mathbf{v}_+),$$

and

$$M - \frac{1}{2}(M_+ + M_-) \geq \alpha \mathbf{w}N, \quad N - \frac{1}{2}(N_+ + N_-) = \alpha \mathbf{w} \mathbf{v}.$$

Once such a function is found, the lemma is proved. Indeed, take any two weights w and v satisfying the Relaxed Sawyer Condition 1' and any two functions $f, g \in L^2$. For an interval J , let X_J be the corresponding octet $(\mathbf{F}_J, \mathbf{f}_J, \dots)$. Then the main inequality yields

$$\gamma |J| \alpha_J \langle f \sqrt{w} \rangle_J \langle g \sqrt{v} \rangle_J \leq |J| \mathcal{B}(X_J) - |J_+| \mathcal{B}(X_{J_+}) - |J_-| \mathcal{B}(X_{J_-}).$$

Adding these inequalities up over all dyadic intervals $J \subset I$ of length $|J| > 2^{-k}|I|$, we get

$$(2.1) \quad \sum_{\substack{J \subset I \\ |J| > 2^{-k}|I|}} \alpha_J \langle f \sqrt{w} \rangle_J \langle g \sqrt{v} \rangle_J |J| \leq |I| \cdot \mathcal{B}(X_I) - \sum_{\substack{J \subset I \\ |J| = 2^{-k}|I|}} |J| \cdot \mathcal{B}(X_J) \leq |I| \cdot \mathcal{B}(X_I).$$

Passing to the limit as $k \rightarrow \infty$, we obtain

$$(2.2) \quad \sum_{J \subset I} \alpha_J \langle f \sqrt{w} \rangle_J \langle g \sqrt{v} \rangle_J |J| \leq |I| \cdot \mathcal{B}(X_I) \leq |I| \cdot C \cdot (\mathbf{F}_I + \mathbf{G}_I) = |I| \cdot C \cdot (\langle f^2 \rangle_I + \langle g^2 \rangle_I) = C \left(\int_I f^2 + \int_I g^2 \right).$$

Now, to prove the lemma, it remains only to pass to the limit as I expands to the whole real line \mathbb{R} .

2.4. Construction of a Bellman function. The Bellman function we are going to present will consist of two parts. The first part is guessed after the Bellman function for the Weighted Carleson Imbedding Theorem:

$$\mathcal{B}_1 := \mathbf{F} - \frac{\mathbf{f}^2}{\mathbf{w} + M}.$$

Since $\mathbf{f}^2 \leq \mathbf{F} \mathbf{w}$, we have $0 \leq \mathcal{B}_1 \leq \mathbf{F}$. It is easy to check that the function \mathcal{B}_1 is concave ($d^2 \mathcal{B}_1 \leq 0$). At last, since $M \leq \mathbf{w}$, the inequality

$$\frac{\partial \mathcal{B}_1}{\partial M} \geq \frac{1}{4} \cdot \frac{\mathbf{f}^2}{\mathbf{w}^2}$$

holds everywhere in the domain.

Therefore

$$\mathcal{B}_1(X) - \frac{1}{2}(\mathcal{B}_1(X_+) + \mathcal{B}_1(X_-)) \geq \frac{\partial \mathcal{B}_1}{\partial M} \cdot \left(M - \frac{1}{2}(M_+ + M_-) \right) \geq \frac{1}{4} \cdot \frac{\mathbf{f}^2}{\mathbf{w}} \alpha N.$$

Hence, if $N \geq \frac{8\mathbf{w}}{4\mathbf{f}}$, we are done (with $\gamma = \frac{1}{16}$).

The second part \mathcal{B}_2 of the Bellman function $\mathcal{B} := \mathcal{B}_1 + \mathcal{B}_2$ takes care of the case when N is small.

Consider a one-parameter family of functions

$$\mathcal{B}_2^{(t)}(X) := \mathbf{F} + \mathbf{G} - \left(\frac{\mathbf{f}^2}{\mathbf{w} + tN} + \frac{\mathbf{g}^2}{\mathbf{v} + t^{-1}N} \right),$$

and put $\mathcal{B}_2(X) := \inf_{t>0} \mathcal{B}_2^{(t)}(X)$.

Fix an octet X , and let $t_* = t_*(X)$ be the value of t where the infimum of $\mathcal{B}_2^{(t)}(X)$ is attained. Then

$$\begin{aligned} \mathcal{B}_2(X) - \frac{1}{2}(\mathcal{B}_2(X_+) + \mathcal{B}_2(X_-)) &\geq \mathcal{B}_2^{(t_*)}(X) - \frac{1}{2}(\mathcal{B}_2^{(t_*)}(X_-) + \mathcal{B}_2^{(t_*)}(X_+)) \\ &\geq \frac{\partial \mathcal{B}_2^{(t_*)}}{\partial N}(X) \cdot \left[N - \frac{1}{2}(N_+ + N_-) \right] \geq \left(\frac{\mathbf{f}^2 t_*}{(\mathbf{w} + t_* N)^2} + \frac{\mathbf{g}^2 t_*^{-1}}{(\mathbf{v} + t_*^{-1} N)^2} \right) \alpha \mathbf{w} \mathbf{v} \\ &= \frac{2\mathbf{f}\mathbf{g}}{(\mathbf{w} + t_* N)(\mathbf{v} + t_*^{-1} N)} \cdot \alpha \mathbf{w} \mathbf{v}. \end{aligned}$$

Thus, if we show that $\mathbf{w} + t_* N \leq 2\mathbf{w}$ and $\mathbf{v} + t_*^{-1} N \leq 2\mathbf{v}$, then we are done (with $\gamma = \frac{1}{2}$). Recall that we only need to prove these inequalities under the assumption $N < \frac{\mathbf{g}\mathbf{w}}{4\mathbf{f}}$ because otherwise \mathcal{B}_1 already gives us all that we need.

By definition, t_* is the value of t at which the supremum

$$\sup_{t>0} \left(\frac{\mathbf{f}^2}{\mathbf{w} + tN} + \frac{\mathbf{g}^2}{\mathbf{v} + t^{-1}N} \right)$$

is attained. If we show that this supremum is at least

$$\frac{\mathbf{f}^2}{\mathbf{w}} + \frac{1}{2} \cdot \frac{\mathbf{g}^2}{\mathbf{v}} \geq \frac{1}{2} \cdot \frac{\mathbf{f}^2}{\mathbf{w}} + \frac{\mathbf{g}^2}{\mathbf{v}}$$

(it is here where we use the inequality $\frac{\mathbf{f}^2}{\mathbf{w}} \geq \frac{\mathbf{g}^2}{\mathbf{v}}$), then we shall indeed be able to conclude that $\mathbf{w} + t_* N \leq 2\mathbf{w}$ and $\mathbf{v} + t_*^{-1} N \leq 2\mathbf{v}$, finishing the proof.

Since

$$\frac{\mathbf{f}^2}{\mathbf{w} + tN} \geq \frac{\mathbf{f}^2}{\mathbf{w}} - tN \frac{\mathbf{f}^2}{\mathbf{w}^2}, \quad \frac{\mathbf{g}^2}{\mathbf{v} + t^{-1}N} \geq \frac{\mathbf{g}^2}{\mathbf{v}} - t^{-1}N \frac{\mathbf{g}^2}{\mathbf{v}^2},$$

it is enough to show that

$$\inf_{t>0} \left\{ tN \frac{\mathbf{f}^2}{\mathbf{w}^2} + t^{-1}N \frac{\mathbf{g}^2}{\mathbf{v}^2} \right\} \leq \frac{1}{2} \cdot \frac{\mathbf{g}^2}{\mathbf{v}}.$$

But this infimum equals $2 \frac{\mathbf{f}\mathbf{g}}{\mathbf{w}\mathbf{v}} N$ which is clearly less than $\frac{1}{2} \cdot \frac{\mathbf{g}^2}{\mathbf{v}}$ under our assumption $N < \frac{\mathbf{g}\mathbf{w}}{4\mathbf{f}}$. The lemma is proved. \square

Proof of the Bilinear Imbedding Theorem. By now we have constructed a Bellman function \mathcal{B} of an octet $X = (\mathbf{F}, \mathbf{f}, \mathbf{w}, \dots)$ such that

$$0 \leq \mathcal{B}(X) \leq 2(\mathbf{F} + \mathbf{G})$$

and

$$\mathcal{B}(X) - \frac{1}{2}(\mathcal{B}(X_+) + \mathcal{B}(X_-)) \geq \begin{cases} \frac{1}{16} \alpha \mathbf{f}\mathbf{g}, & \text{if } \frac{\mathbf{f}^2}{\mathbf{w}} \geq \frac{\mathbf{g}^2}{\mathbf{v}}; \\ 0, & \text{if } \frac{\mathbf{f}^2}{\mathbf{w}} < \frac{\mathbf{g}^2}{\mathbf{v}}. \end{cases}$$

To prove part (B) of the Bilinear Imbedding Theorem, it would suffice to find a function $\tilde{\mathcal{B}}$ satisfying

$$0 \leq \tilde{\mathcal{B}}(X) \leq 14\sqrt{\mathbf{F}\mathbf{G}}$$

and

$$\tilde{\mathcal{B}}(X) - \frac{1}{2}(\tilde{\mathcal{B}}(X_+) + \tilde{\mathcal{B}}(X_-)) \geq \begin{cases} \frac{1}{16}\alpha\mathbf{f}\mathbf{g}, & \text{if } \frac{\mathbf{f}^2}{\mathbf{F}\mathbf{w}} \geq \frac{\mathbf{g}^2}{\mathbf{G}\mathbf{v}}; \\ 0, & \text{if } \frac{\mathbf{f}^2}{\mathbf{F}\mathbf{w}} < \frac{\mathbf{g}^2}{\mathbf{G}\mathbf{v}}. \end{cases}$$

To this end, consider the family of functions

$$\tilde{\mathcal{B}}^{(\tau)}(\mathbf{F}, \mathbf{f}, \dots) := \tau\mathbf{F} + 14\tau^{-1}\mathbf{G} + \mathcal{B}(\tau\mathbf{F}, \sqrt{\tau}\mathbf{f}, \mathbf{w}, \tau^{-1}\mathbf{G}, \sqrt{\tau^{-1}}\mathbf{g}, \mathbf{v}, M, N).$$

(This family has a clear physical sense: essentially we are just looking at what happens to our Bellman function \mathcal{B} if we multiply the function f by $\sqrt{\tau}$, divide the function g by the same $\sqrt{\tau}$ and leave the weights v and w without change.) Observe that we always have

$$\tau\mathbf{F} + 14\tau^{-1}\mathbf{G} \leq \tilde{\mathcal{B}}^{(\tau)}(X) \leq 3\tau\mathbf{F} + 16\tau^{-1}\mathbf{G}.$$

The sought after function $\tilde{\mathcal{B}}$ can now be defined by

$$\tilde{\mathcal{B}}(X) := \inf_{\tau>0} \tilde{\mathcal{B}}^{(\tau)}(X).$$

To check the properties, note first of all that for $\tau = 2\sqrt{\mathbf{G}/\mathbf{F}}$, we have

$$\tilde{\mathcal{B}}^{(\tau)}(X) \leq 6\sqrt{\mathbf{F}\mathbf{G}} + 8\sqrt{\mathbf{F}\mathbf{G}} = 14\sqrt{\mathbf{F}\mathbf{G}}.$$

Thereby, $\tilde{\mathcal{B}}(X) \leq 14\sqrt{\mathbf{F}\mathbf{G}}$ as well.

Let $\tau_* = \tau_*(X)$ be the value of τ at which the infimum is attained. Since

$$14\tau_*^{-1}\mathbf{G} \leq \tilde{\mathcal{B}}^{(\tau_*)}(X) \leq 14\sqrt{\mathbf{F}\mathbf{G}},$$

we should have $\tau_* \geq \sqrt{\mathbf{G}/\mathbf{F}}$. Note that the difference

$$\tilde{\mathcal{B}}(X) - \frac{1}{2}(\tilde{\mathcal{B}}(X_-) + \tilde{\mathcal{B}}(X_+)) \geq \tilde{\mathcal{B}}^{(\tau_*)}(X) - \frac{1}{2}(\tilde{\mathcal{B}}^{(\tau_*)}(X_-) + \tilde{\mathcal{B}}^{(\tau_*)}(X_+))$$

is always non-negative and can be estimated from below by $\frac{1}{16}\alpha\mathbf{f}\mathbf{g}$ provided that

$$\frac{\tau_*\mathbf{f}^2}{\mathbf{w}} \geq \frac{\tau_*^{-1}\mathbf{g}^2}{\mathbf{v}}.$$

But for $\tau_* \geq \sqrt{\mathbf{G}/\mathbf{F}}$, this inequality is implied by the inequality

$$\frac{\mathbf{f}^2}{\mathbf{F}\mathbf{w}} \geq \frac{\mathbf{g}^2}{\mathbf{G}\mathbf{v}},$$

which means that we are done.

3. TWO-WEIGHT INEQUALITIES FOR THE (GENERALIZED) SQUARE FUNCTION

Recall that for a dyadic interval $J \in \mathcal{D}$, the corresponding Haar function h_J is defined by

$$h_J(t) := \begin{cases} 0, & t \notin J; \\ -1, & t \in J_-; \\ +1, & t \in J_+. \end{cases}$$

For a function $f \in L^1_{\text{loc}}$ and for an interval $I \in \mathcal{D}$, define

$$\Delta_I f := \frac{1}{|I|} \int_{\mathbb{R}} f h_I = \frac{1}{2}(\langle f \rangle_{I_+} - \langle f \rangle_{I_-}).$$

In this section we are going to prove a Sawyer type theorem for the (generalized) square function operator

$$T_0 f := \left(\sum_{J \in \mathcal{D}} a_J [\Delta_J f]^2 \chi_J \right)^{\frac{1}{2}}$$

where $\{a_J\}_{J \in \mathcal{D}}$ is an arbitrary family of positive numbers. The “standard” dyadic square function $\mathcal{S}f$ corresponds to a_J all equal to 1.

Theorem (Sawyer type theorem for the square function operator). *The operator $T = M_{\sqrt{v}} T_0 M_{\sqrt{w}}$ is bounded in L^2 if and only if it is uniformly bounded on all test-functions $f = \sqrt{w} \chi_I$, i.e., if and only if there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}} [T_0(w \chi_I)]^2 \leq C \int_I w \quad \text{for all } I \in \mathcal{D}.$$

Proof. Necessity is obvious. To prove the sufficiency part, let us note first of all that

$$T_0(w \chi_{I_{\pm}})(t) \geq \frac{\sqrt{a_{I_{\pm}}}}{2} \langle w \rangle_{I_{\pm}}$$

for every $t \in I$ (just consider the term corresponding to the interval I in the definition of $T_0(w \chi_{I_{\pm}})$). Therefore, the Sawyer condition implies the inequalities

$$\frac{a_I}{4} \langle w \rangle_{I_{\pm}}^2 \int_I v \leq \int_{I_{\pm}} w,$$

i.e.,

$$\langle w \rangle_{I_{\pm}} \langle v \rangle_I \leq 4a_I^{-1}.$$

Taking into account that

$$\langle w \rangle_I \leq \max\{\langle w \rangle_{I_-}, \langle w \rangle_{I_+}\},$$

we obtain the Muckenhoupt type inequality

$$(3.1) \quad \langle w \rangle_I \langle v \rangle_I \leq 4a_I^{-1}.$$

Now, for every interval $I \subset \mathcal{D}$, let us consider the truncated square function operator

$$T_0^{(I)} f := \left(\sum_{J \in \mathcal{D}, J \subset I} a_J [\Delta_J f]^2 \chi_J \right)^{\frac{1}{2}}.$$

Since $T_0^{(I)} f \leq T_0 f$ pointwise on \mathbb{R} , the assumption of the theorem implies

$$\int_I [T_0^{(I)}(w \chi_I)]^2 v \leq C \int_I w,$$

i.e.,

$$(3.2) \quad \frac{1}{|I|} \sum_{J \in \mathcal{D}, J \subset I} a_J [\Delta_J w]^2 \langle v \rangle_J |J| \leq C \langle w \rangle_I.$$

We are going to show now that Conditions (3.1) and (3.2) are sufficient for the boundedness of the operator T in L^2 .

3.1. Disbalanced Haar functions. We would like to modify the Haar functions a little bit so that they become orthogonal in the weighted space $L^2(w)$ instead of the non-weighted L^2 . More precisely, we would like to find constants A_J ($J \in \mathcal{D}$) such that the disbalanced Haar functions

$$H_J := h_J - A_J \chi_J$$

are orthogonal in $L^2(w)$, or, which is the same, the functions $\sqrt{w} H_J$ are orthogonal in the non-weighted L^2 . Their orthogonality is equivalent to the relations

$$0 = (1, H_J)_{L^2(w)} = \int_{\mathbb{R}} H_J w = |J| [\Delta_J w - A_J \langle w \rangle_J] \quad (J \in \mathcal{D}),$$

from which we conclude that we must take

$$A_J = \frac{\Delta_J w}{\langle w \rangle_J}.$$

Now, let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an arbitrary L^2 -function. Our aim is to show that

$$\int_{\mathbb{R}} [T_0 f]^2 \leq C \int_{\mathbb{R}} f^2,$$

or, which is the same,

$$\sum_{J \in \mathcal{D}} a_J [\Delta_J (f \sqrt{w})]^2 \langle v \rangle_J |J| \leq C \int_{\mathbb{R}} f^2.$$

Note that

$$\begin{aligned} \Delta_J (f \sqrt{w}) &= \frac{1}{|J|} (f \sqrt{w}, h_J)_{L^2} = \frac{1}{|J|} (f \sqrt{w}, H_J + A_J \chi_J)_{L^2} \\ &= \frac{1}{|J|} (f, \sqrt{w} H_J)_{L^2} + A_J \langle f \sqrt{w} \rangle_J. \end{aligned}$$

Since $(a + b)^2 \leq 2(a^2 + b^2)$ for any $a, b \in \mathbb{R}$, it would suffice to show that

$$\sum_1 := \sum_{J \in \mathcal{D}} a_J \left[\frac{1}{|J|} (f, \sqrt{w} H_J)_{L^2} \right]^2 \langle v \rangle_J |J| \leq C \int_{\mathbb{R}} f^2$$

and

$$\sum_2 := \sum_{J \in \mathcal{D}} a_J A_J^2 \langle f \sqrt{w} \rangle_J^2 \langle v \rangle_J |J| \leq C \int_{\mathbb{R}} f^2.$$

3.2. Estimate for \sum_1 . Note that the functions $\sqrt{w} H_J$ are orthogonal in L^2 and

$$\|\sqrt{w} H_J\|_{L^2}^2 = |J| \langle w \rangle_J \left[1 - \left(\frac{\Delta_J w}{\langle w \rangle_J} \right)^2 \right] \leq |J| \langle w \rangle_J.$$

Recall that for any orthogonal system of functions $\{\varphi_J\}_{J \in \mathcal{D}}$, the Bessel inequality

$$\sum_{J \in \mathcal{D}} \frac{1}{\|\varphi_J\|_{L^2}^2} (f, \varphi_J)_{L^2}^2 \leq \int_{\mathbb{R}} f^2$$

holds. Therefore,

$$\sum_{J \in \mathcal{D}} \frac{1}{|J| \langle w \rangle_J} (f, \sqrt{w} H_J)_{L^2}^2 \leq \int_{\mathbb{R}} f^2.$$

Now it remains only to note that, due to the Muckenhoupt condition (1), the factors $\frac{1}{|J|}a_J\langle v \rangle_J$ in front of $(f, \sqrt{w}H_J)_{L^2}$ in the definition of \sum_1 do not exceed $\frac{C}{|J|\langle w \rangle_J}$.

3.3. Estimate for \sum_2 . The estimate for \sum_2 is exactly the same as the estimate of the Weighted Carleson Imbedding Theorem with

$$\alpha_J := a_J A_J^2 \langle v \rangle_J |J|.$$

Therefore, to prove it, it is enough to check the Sawyer–Carleson condition

$$\frac{1}{|I|} \sum_{J \subset I} \alpha_J \langle w \rangle_J^2 \leq C \langle w \rangle_I \quad (I \in \mathcal{D}).$$

Recalling that $A_J = \frac{\Delta_J w}{\langle w \rangle_J}$, we see that this condition coincides with (3.2). The Sawyer type theorem for the dyadic square function operator is thus completely proved.

4. TWO-WEIGHT INEQUALITIES FOR HAAR MULTIPLIERS

Again let \mathcal{D} be some dyadic lattice and let $h_J, J \in \mathcal{D}$, be the corresponding family of Haar functions.

Let $\mathbf{a} = \{a_J\}_{J \in \mathcal{D}}$ be any sequence of real numbers. Define the operator $T_0(\mathbf{a})$ by

$$T_0(\mathbf{a})f := \sum_{J \in \mathcal{D}} \frac{a_J}{|J|} (f, h_J)_{L^2} h_J = \sum_{J \in \mathcal{D}} a_J [\Delta_J f] h_J.$$

As before, we would like to know whether the “weighted operator” $T(\mathbf{a}) := M_{\sqrt{v}} T_0(\mathbf{a}) M_{\sqrt{w}}$ is bounded in L^2 if and only if it is bounded on the test-functions $\sqrt{w} \chi_I$ and the adjoint operator $T(\mathbf{a})^* = M_{\sqrt{w}} T_0(\mathbf{a})^* M_{\sqrt{v}}$ is bounded on the test-functions $\sqrt{v} \chi_I$ ($I \in \mathcal{D}$).

At the present moment we are unable to answer this question for an *individual* operator $T_0(\mathbf{a})$. We are quite sure that in general the answer is negative, but there may exist some interesting partial cases for which the Sawyer type theorem still holds. For instance, it surely holds when all $a_J = 1$ (the identity operator) though this case can hardly be called “interesting”.

What we are going to do below is to prove a Sawyer type theorem for a “family of operators”. Namely, fix some sequence $\mathbf{a} = \{a_J\}_{J \in \mathcal{D}}$ of positive numbers and consider all the operators $T_0(\sigma \mathbf{a})$ and the corresponding weighted operators $T(\sigma \mathbf{a}) = M_{\sqrt{v}} T_0(\sigma \mathbf{a}) M_{\sqrt{w}}$ where $\sigma = \{\sigma_J\}_{J \in \mathcal{D}}$ is an arbitrary sequence of signs, i.e., $\sigma_J = +1$ or -1 for every J . (As usual, $\sigma \mathbf{a}$ stands for the sequence of products $\{\sigma_J a_J\}_{J \in \mathcal{D}}$.)

Theorem (Sawyer type theorem for Haar multipliers). *The operators $T(\sigma \mathbf{a})$ are uniformly bounded in L^2 , i.e., $\sup_{\sigma} \|T(\sigma \mathbf{a})\|_{L^2 \rightarrow L^2} < +\infty$ if and only if they are uniformly bounded on the test-functions $\sqrt{w} \chi_I$ and the adjoint operators $T(\sigma \mathbf{a})^*$ are uniformly bounded on the test-functions $\sqrt{v} \chi_I$, that is, if and only if there exists a constant $C > 0$ such that*

$$(*) \quad \int_{\mathbb{R}} [T_0(\sigma \mathbf{a})(w \chi_I)]^2 v \leq C \int_I w \quad \text{and} \quad \int_{\mathbb{R}} [T_0(\sigma \mathbf{a})(v \chi_I)]^2 w \leq C \int_I v$$

for all dyadic intervals $I \in \mathcal{D}$ and sign sequences σ .

Proof. Necessity is obvious. To prove the sufficiency part, note first of all, that Sawyer’s condition (*) implies that for each dyadic interval $I \in \mathcal{D}$, the following three estimates hold:

- (1) $\langle w \rangle_I \langle v \rangle_I \leq 4Ca_I^{-2}$ (Muckenhoupt type condition);
- (2_w) $\frac{1}{|I|} \sum_{J \subset I} a_J^2 [\Delta_J w]^2 \langle v \rangle_J |J| \leq C \langle w \rangle_I$;
- (2_v) $\frac{1}{|I|} \sum_{J \subset I} a_J^2 [\Delta_J v]^2 \langle w \rangle_J |J| \leq C \langle v \rangle_I$.

Indeed, fixing some $I \in \mathcal{D}$ and taking the average of $\int_{\mathbb{R}} [T_0(\sigma \mathbf{a})(w \chi_I)]^2 v$ over all possible sign sequences σ , we get

$$\frac{1}{|I|} \sum_{J \in \mathcal{D}} a_J^2 [\Delta_J (w \chi_I)]^2 \langle v \rangle_J |J| = \text{Avg}_{\sigma} \frac{1}{|I|} \int_{\mathbb{R}} [T_0(\sigma \mathbf{a})(w \chi_I)]^2 v \leq C \langle w \rangle_I.$$

Observing that $\Delta_J (w \chi_I) = \Delta_J w$ for all $J \subset I$ and that every square is non-negative, we immediately arrive at (2_w).

To prove (1), let us apply the above inequality with I_- and I_+ instead of I . Taking into account that $|\Delta_I (w \chi_{I_{\pm}})| = \frac{1}{2} \langle w \rangle_{I_{\pm}}$, we find

$$\frac{a_I^2}{4} \langle w \rangle_{I_{\pm}}^2 \langle v \rangle_I \leq C \langle w \rangle_{I_{\pm}},$$

i.e.,

$$\langle w \rangle_{I_{\pm}} \langle v \rangle_I \leq 4Ca_I^{-2}.$$

Recalling that $\langle w \rangle_I \leq \max\{\langle w \rangle_{I_-}, \langle w \rangle_{I_+}\}$, we get (1).

The proof of (2_v) is similar to that of (2_w).

To continue we shall use the disbalanced Haar functions $H_J = h_J - A_J \chi_J$ constructed in the previous section. Only now we shall need them for both weights w and v . So, let us put

$$A_J^{(w)} := \frac{\Delta_J w}{\langle w \rangle_J}, \quad H_J^{(w)} := h_J - A_J^{(w)} \chi_J; \quad A_J^{(v)} := \frac{\Delta_J v}{\langle v \rangle_J}, \quad H_J^{(v)} := h_J - A_J^{(v)} \chi_J.$$

The uniform boundedness of the operators $T(\sigma \mathbf{a})$ is equivalent to the bilinear inequality

$$\Sigma(f, g) := \sum_{J \in \mathcal{D}} a_J |\Delta_J (f \sqrt{w})| \cdot |\Delta_J (g \sqrt{v})| \cdot |J| \leq C \|f\|_{L^2} \|g\|_{L^2}.$$

Note again that

$$\Delta_J (f \sqrt{w}) = \frac{1}{|J|} (f, \sqrt{w} H_J^{(w)})_{L^2} + A_J^{(w)} \langle f \sqrt{w} \rangle_J$$

and

$$\Delta_J (g \sqrt{v}) = \frac{1}{|J|} (g, \sqrt{v} H_J^{(v)})_{L^2} + A_J^{(v)} \langle g \sqrt{v} \rangle_J.$$

Plugging these expressions into the formula for $\Sigma(f, g)$, we get 4 sums to estimate. Our next aim will be to show that all these sums except

$$\tilde{\Sigma}(f, g) = \sum_{J \in \mathcal{D}} a_J |A_J^{(w)}| \cdot |A_J^{(v)}| \cdot |\langle f \sqrt{w} \rangle_J| \cdot |\langle g \sqrt{v} \rangle_J| \cdot |J|$$

can be estimated using conditions (1), (2_w) and (2_v). More precisely, we shall prove that under these conditions,

$$|\Sigma(f, g) - \tilde{\Sigma}(f, g)| \leq C\|f\|_{L^2}\|g\|_{L^2}.$$

Indeed, this difference does not exceed $\Sigma_1(f, g) + \Sigma_{2,w}(f, g) + \Sigma_{2,v}(f, g)$ where

$$\Sigma_1(f, g) := \sum_{J \in \mathcal{D}} a_J \left| \frac{1}{|J|} (f, \sqrt{w} H_J^{(w)})_{L^2} \right| \cdot \left| \frac{1}{|J|} (g, \sqrt{v} H_J^{(v)})_{L^2} \right| \cdot |J|;$$

$$\Sigma_{2,w}(f, g) := \sum_{J \in \mathcal{D}} a_J |A_J^{(w)}| \cdot |\langle f \sqrt{w} \rangle_J| \cdot \left| \frac{1}{|J|} (g, \sqrt{v} H_J^{(v)})_{L^2} \right| \cdot |J|;$$

and

$$\Sigma_{2,v}(f, g) := \sum_{J \in \mathcal{D}} a_J |A_J^{(v)}| \cdot |\langle g \sqrt{v} \rangle_J| \cdot \left| \frac{1}{|J|} (f, \sqrt{w} H_J^{(w)})_{L^2} \right| \cdot |J|.$$

4.1. **Estimate for Σ_1 .** As we have seen in Section 3,

$$\sum_{J \in \mathcal{D}} \frac{1}{|J| \langle w \rangle_J} (f, \sqrt{w} H_J^{(w)})_{L^2}^2 \leq \int_{\mathbb{R}} f^2.$$

Analogously,

$$\sum_{J \in \mathcal{D}} \frac{1}{|J| \langle v \rangle_J} (g, \sqrt{v} H_J^{(v)})_{L^2}^2 \leq \int_{\mathbb{R}} g^2.$$

The Cauchy–Schwartz inequality now yields

$$\sum_{J \in \mathcal{D}} \frac{1}{|J| \sqrt{\langle w \rangle_J \langle v \rangle_J}} \left| (f, \sqrt{w} H_J^{(w)})_{L^2} \right| \cdot \left| (g, \sqrt{v} H_J^{(v)})_{L^2} \right| \leq \|f\|_{L^2} \|g\|_{L^2}.$$

To estimate Σ_1 , it remains only to note that, according to (1),

$$a_J \leq \frac{C}{\sqrt{\langle w \rangle_J \langle v \rangle_J}}.$$

4.2. **Estimates for $\Sigma_{2,w}$ and $\Sigma_{2,v}$.** Since

$$\sum_{J \in \mathcal{D}} \frac{1}{|J| \langle v \rangle_J} (g, \sqrt{v} H_J^{(v)})_{L^2}^2 \leq \int_{\mathbb{R}} g^2,$$

to get the desired estimate for $\Sigma_{2,w}$, it is enough to show that

$$\sum_{J \in \mathcal{D}} a_J^2 |A_J^{(w)}|^2 \langle f \sqrt{w} \rangle_J^2 \langle v \rangle_J |J| \leq C \int_{\mathbb{R}} f^2$$

for every $f \in L^2$. According to the Weighted Carleson Imbedding Theorem, it suffices to check that for every $I \in \mathcal{D}$,

$$\frac{1}{|I|} \sum_{J \subset I} a_J^2 |A_J^{(w)}|^2 \langle w \rangle_J \langle v \rangle_J |J| \leq C \langle w \rangle_I.$$

But the last inequality coincides with (2_w), and we are done. Similarly, the estimate for $\Sigma_{2,v}$ is implied by (2_v).

4.3. Auxiliary operator \tilde{T} and the end of the proof. Now let \mathcal{F} be any family of non-negative functions. The above argument shows that under Conditions (1), (2_w) and (2_v), the following assertions are equivalent:

- $\sup\left\{\frac{\|T(\sigma\mathbf{a})f\|_{L^2}}{\|f\|_{L^2}} : \sigma_J = \pm 1, f \in \mathcal{F}\right\} < +\infty;$
- $\sup\left\{\frac{\Sigma(f, g)}{\|f\|_{L^2}\|g\|_{L^2}} : f \in \mathcal{F}, g \in L^2\right\} < +\infty;$
- $\sup\left\{\frac{\tilde{\Sigma}(f, g)}{\|f\|_{L^2}\|g\|_{L^2}} : f \in \mathcal{F}, g \in L^2\right\} < +\infty;$
- $\sup\left\{\frac{\|\tilde{T}f\|_{L^2}}{\|f\|_{L^2}} : f \in \mathcal{F}\right\} < +\infty,$

where $\tilde{T} = M_{\sqrt{v}}\tilde{T}_0M_{\sqrt{w}}$ and the (linear) operator \tilde{T}_0 is defined by

$$\tilde{T}_0f := \sum_{J \in \mathcal{D}} a_J |A_J^{(w)}| \cdot |A_J^{(v)}| \cdot \langle f \rangle_J \chi_J.$$

Applying this observation to the family of all test-functions $\sqrt{w}\chi_I$, $I \in \mathcal{D}$, and recalling that the operators $T(\sigma\mathbf{a})$ are uniformly bounded on such test-functions, we conclude that the operator \tilde{T} is also bounded on such test-functions. Similarly, the adjoint operator \tilde{T}^* is bounded on the family of test-functions $\sqrt{v}\chi_I$, $I \in \mathcal{D}$. According to our dyadic version of Sawyer's theorem on operators with positive kernels, this means that the operator \tilde{T} is bounded in L^2 . Therefore for any two non-negative functions $f, g \in L^2$,

$$\tilde{\Sigma}(f, g) = (\tilde{T}f, g)_{L^2} \leq C\|f\|_{L^2}\|g\|_{L^2}.$$

The trivial inequality $\tilde{\Sigma}(f, g) \leq \tilde{\Sigma}(|f|, |g|)$ allows us to extend this estimate to all (not necessarily non-negative) functions f and g , finishing the proof of the bilinear inequality $\Sigma(f, g) \leq C\|f\|_{L^2}\|g\|_{L^2}$ and thereby of the theorem.

5. CONCLUDING REMARKS

Notice that conditions (1), (2_v), (2_w), allowed us to reduce the question of boundedness of Haar multipliers to the boundedness of operator \tilde{T}_0 with positive kernel. Conditions (1), (2_v), (2_w) are necessary for the boundedness of Haar multipliers, and, thus, the reduction is an entirely equivalent one. On the other hand for \tilde{T}_0 (as for dyadic operators with positive kernel) we have two necessary and sufficient conditions for its boundedness. They are listed in Section 2. Thus we have a list of five conditions in terms of v, w necessary and sufficient for the two-weight boundedness of Haar multipliers.

We do not know whether these five conditions are independent. We strongly believe they are. However if either weight w or weight v satisfy the A_∞ condition, then only (1) is enough for two-weight boundedness! The same is true if the weights are mutually A_∞ . These facts are well-known to experts. The reader can find the proofs in [NTV1].

Coming back to our list of five dyadic conditions, one can notice that they can be rewritten in "continuous" form. Then they form the list of conditions sufficient for

the Hilbert transform H to be bounded between two weighted spaces with different weights. Again the reader can find this in [NTV1]. Unfortunately some of these conditions are most probably not necessary. This is why we decided not to include them in this article.

REFERENCES

- [B] St. Buckley, Summation conditions on weights, *Mich. Math. J.*, **40** (1993), 153-170. MR **94d**:42021
- [Bu] D.L Burkholder, Explorations in martingale theory and its applications. *Ecole d'Été de Probabilité de Saint-Flour XIX-1989*, 1-66, *Lecture Notes in Mathematics*, **1464**, Springer, Berlin, 1991. MR **92m**:60037
- [CF] R.R. Coifman, Ch. Fefferman, Weighted norm inequalities for maximal functions and singular integrals, *Studia Math.*, **51** (1974), 241-250. MR **50**:10670
- [CJS] R.R. Coifman, P.W. Jones, and St. Semmes, Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves, *J. of Amer. Math. Soc.*, **2** (1989), No. 3, 553-564. MR **90k**:42017
- [CS1] M. Cotlar, C. Sadosky, On the Helson-Szegö theorem and a related class of modified Toeplitz kernels, in *Harmonic Analysis in Euclidean spaces*, ed. by G.Weiss and S. Wainger, *Proc. Symp. Pure Math.* **35**, Amer. Math. Soc., Providence, R.I., 1979, 383-407. MR **81j**:42022
- [CS2] M. Cotlar, C. Sadosky, On some L^p version of the Helson-Szegö theorem, *Conference on Harmonic Analysis in honor of Antony Zygmund* (Chicago, 1981), **vol.1**, ed. by W. Beckner et al., *Wadsworth Math. Ser.*, Wadsworth, Belmont, CA, 1983, 306-317. MR **85i**:42015
- [ChWW] A. Chang, J.M. Wilson, Th. Wolff, Some weighted norm inequalities concerning the Schrödinger operators. *Comment. Math. Helvetici*, **60** (1985), 217-246. MR **87d**:42027
- [F] C. Fefferman, The uncertainty principle. *Bull. of Amer. Math. Soc.*, **9** (1983), No. 2, 127-206. MR **85f**:35001
- [FKP] R.A. Fefferman, C.E. Kenig, J. Pipher, The theory of weights and the Dirichlet problem for elliptic equations, *Ann. of Math.* **134** (1991), 65-124. MR **93h**:31010
- [G] John B. Garnett, *Bounded analytic functions*. *Pure and Applied Mathematics*, 96. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981. xvi+467 MR **83g**:30037
- [KV] N.J. Kalton, J.E. Verbitsky, Nonlinear equations and weighted norm inequalities, *Trans. Amer. Math. Soc.*, to appear. CMP 98:02
- [N] F. Nazarov, A counterexample to a problem of Sarason on boundedness of the product of two Toeplitz operators. Preprint, 1996, 1-5.
- [NT] F. L. Nazarov and S. R. Treil, 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. MR **99d**:42026
- [NT1] F.Nazarov, S.Treil, The weighted norm inequalities for Hilbert transform are now trivial, *C.R. Acad. Sci. Paris, Série J*, **323**, (1996), 717-722. CMP 97:03
- [NTV] F.Nazarov, S.Treil, A.Volberg, Cauchy integral and Calderón-Zygmund operators on non-homogeneous spaces. *IMRN (Int. Math. Res. Notes.)*, **1997**, No. 15, 703-726. MR **99e**:42028
- [NTV1] F.Nazarov, S.Treil, A.Volberg, The Bellman functions and two weight inequalities for Haar multipliers, *MSRI Preprint 1997-103*, p. 1-31.
- [Ne] C.J. Neugebauer, Inserting A_p -weights, *Proc. of the Amer. Math. Soc.* **87** (1983), 644-648. MR **84d**:42026
- [S] C. Sadosky, Lifting of kernels shift-invariant in scattering systems, *Holomorphic spaces*, *MSRI publications*, **32**, 1997. MR **99e**:47034
- [Sa] E.T. Sawyer, Norm inequalities relating singular integrals and the maximal functions, *Studia Math.* **75** (1983), No. 3, 253-263. MR **85c**:42018
- [S1] E.T. Sawyer, A characterization of a two weight norm inequality for maximal operators, *Studia Math.* **75** (1982), 1-11. MR **84i**:42032
- [S2] E.T. Sawyer, A characterization of two weight norm inequality for fractional and Poisson integrals. *Trans. Amer. Math. Soc.* **308** (1988), 533-545. MR **89d**:26009

- [St] E.M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Math. Series, 43, Monographs in Harmonic Analysis, Princeton Univ. Press, Princeton, NJ, 1993. MR **95c**:42002
- [SW] E.T. Sawyer, R.L.Wheeden, Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces, Amer. J. Math. **114** (1992), 813-874. MR **94i**:42024
- [T] X. Tolsa, Boundedness of the Cauchy integral operator. Preprint, 1997.
- [TV1] S. Treil and A. Volberg, Wavelets and the angle between past and future, J. Funct. Anal. **143** (1997), no. 2, 269–308. CMP 97:06
- [TV2] S.R. Treil, A.L. Volberg, Weighted embeddings and weighted norm inequalities for the Hilbert transform and the maximal operator. St. Petersburg Math. J. **7** (1996), 207-226. MR **97c**:42017
- [TVZ] S.R. Treil, A.L. Volberg, D. Zheng, Hilbert transform, Toeplitz operators and Hankel operators, and invariant A_∞ weights. Revista Mat. Iberoamericana, **13** (1997), No. 2, 319-360. CMP 98:11
- [VW] J.E. Verbitsky, R.L. Wheeden, Weighted norm inequalities for integral operators. Preprint, 1996. 1-25.
- [V] A. Volberg, Matrix A_p weights via S -functions. J. Amer. Math. Soc. **10** (1997), 445-466. MR **98a**:42013
- [Zh] Dechao Zheng, The distribution function inequality and products of Toeplitz operators and Hankel operators, J. Funct. Anal. **138** (1996), no. 2, 477–501. MR **97e**:47040

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MICHIGAN 48824-1027

E-mail address: fedja@math.msu.edu

E-mail address: treil@math.msu.edu

E-mail address: volberg@math.msu.edu