THE \(A_2\) THEOREM: REMARKS AND COMPLEMENTS

TUOMAS P. HYTÖNEN

Abstract. I give a mini-survey of several approaches to the \(A_2\) theorem, biased towards the “corona” rather than the “Bellman” side of the coin. There are two new results (a streamlined form of Lerner’s local oscillation formula, and the sharpness of the linear-in-complexity weak \((1,1)\) bound for dyadic shifts) and two new proofs of known results (the \(A_p-A_\infty\) testing conditions, and the two-weight \(T1\) theorem for positive dyadic operators).

1. Introduction

\(A_2\) theorem for a Calderón–Zygmund operator \(T\) is the sharp weighted bound

\[ \|Tf\|_{L^2(w)} \leq c_{T,A_2} \|f\|_{L^2(w)} \]  

in terms of the Muckenhoupt “norm” (or “characteristic”, or “constant”)

\[ [w]_{A_2} := \sup_Q \langle w \rangle_Q \langle 1/w \rangle_Q := \sup_Q \frac{1}{|Q|} \int_Q w \cdot \frac{1}{Q} \frac{1}{|Q|} := \sup_Q \int_Q \frac{w}{|Q|} \cdot \int_Q \frac{1}{|Q|}. \]

After being established for several particular operators first \cite{10, 25, 26, 29}, I proved (1.1) for an arbitrary Calderón–Zygmund operator \(T\) in July 2010 \cite{6}. This was considered quite a difficult result back then. It used, in particular: a weighted \(T1\) theorem of Pérez, Treil and Volberg \cite{24}, permitting the reduction of (1.1) to the “testing condition”

\[ \|T(w^{-1}1_Q)\|_{L^2(w)} \leq c_{T,A_2} w^{-1}(Q)^{1/2}, \]

\[ \|T^*(w1_Q)\|_{L^2(w^{-1})} \leq c_{T,A_2} w(Q)^{1/2}; \]

a reduction of the operator \(T\) to dyadic model operators called shifts \(S_k\) via a probabilistic argument inspired by the work of Nazarov, Treil and Volberg on non-doubling harmonic analysis \cite{22}; and a subtle multi-step (“corona”) decomposition, elaborating on earlier work of Lacey, Petermichl and Reguera \cite{14}. (Personally, I think that “corona” is a misused word in this context, since the original Corona Problem is rather distant, but I adopt this common terminology for this article.)

The last two years have greatly expanded our understanding of the \(A_2\) theorem, and several known approaches are illustrated in the following diagram. The nodes “\(A_2\) for \(T\)” and “testing \(T\)” (for different choices of \(T\)) represent intermediate results, as defined in (1.1) and (1.2), whereas the arrows indicate different routes of passing from one intermediate result to the next. Whenever an arrow crosses a dashed line, it means that a corresponding auxiliary result is needed at that point. The further...
right one applies the weighted $T1$ theorem, the more difficult it becomes. To a smaller extent, this is also true for the corona decomposition.

The dotted lines indicate steps that are possible but unnecessary, since there is also a short direct proof of $A2$ for $S_0^+$; but I will say more about this path below.

In the diagram, CZO denotes an arbitrary Calderón–Zygmund operator, and $S_k$ an arbitrary dyadic shift of order $k \in \mathbb{N}$, as defined in [6] [14]. In the conditions of $A2$ for $S_k$ and “testing $S_k$”, as defined in (1.1) and (1.2), it is understood that $c_{S_k}$ should grow at most polynomially in $k$. For the purposes of the present discussion, it is not necessary to recall the general definition of a dyadic shift, since we only explicitly deal with the following particular case: The symbol $S_k^+$ (of which $S_0^+$ is a special case) denotes a positive dyadic shift of order $k$ of the specific form

$$S_k^+ f = \sum_{K \in \mathcal{K}} 1_K \int_{K^{(k)}} f \, dx,$$

where $K^{(k)}$ is the $k$ generations older dyadic ancestor of $K$ (so that $K^{(0)} := K$), and $\mathcal{K}$ is an arbitrary sparse collection of dyadic cubes, i.e., there are pairwise disjoint subsets $E(K) \subset K$ with $|E(K)| \geq c|K|$ for a fixed constant $c > 0$.

My original proof of the $A2$ theorem [6] proceeded via the “top right” route

corona $\rightarrow$ testing $S_k$ $\rightarrow$ testing CZO $\rightarrow$ $A2$ for CZO,

where the last step was borrowed from Pérez, Treil and Volberg [24]. This difficult step was avoided by the somewhat easier route

corona $\rightarrow$ testing $S_k$ $\rightarrow$ $A2$ for $S_k$ $\rightarrow$ $A2$ for CZO, (1.3)
taken by Hytönen–Pérez–Treil–Volberg [12]. The estimates along this route were further elaborated by Hytönen, Lacey, Martikainen et al. [8] to show that even the maximal truncated singular integrals

$$T\# f(x) := \sup_{\epsilon > 0} |T, f(x)|, \quad T, f(x) := \int_{|x-y| > \epsilon} K(x, y) f(y) \, dy,$$

can be reached, proving “$A2$ for $T\#$.” It was at this point that the $A2$ technology was at the peak of its difficulty: In addition to the methods shown in the diagram, ideas coming from the proof of Carleson’s theorem on pointwise convergence of Fourier
The goal of this paper is to further simplify this lower-left route (1.8) to the theorem and the mixed estimate (1.7). A detailed technical outline of this route is given in Section 2. After this, the new contributions are as follows:

\[
\text{corona} \rightarrow \text{testing } S_0^+ \rightarrow A_p \text{ for } S_0^+ \rightarrow A_p \text{ for } S_k^+ \rightarrow A_p \text{ for CZO}, \quad (1.8)
\]

Although the full \(A_2\) theorem was reduced to positive operators, a theme further elaborated in 2012. Before going into these most recent developments, it should be mentioned that the corona and testing condition parts can also be replaced by alternative Bellman function arguments (like those by Nazarov and Volberg [23]), but I would say that they remain roughly on the same level of difficulty.

However, both corona, testing, and Bellman functions were completely avoided by Lerner’s discovery [17] of “\(A_2\) for \(S_0^+\)” since the starting point, \(A_2\) for the simplest operator \(S_0^+\), can be directly verified by an elegant few-lines argument due to Cruz-Uribe, Martell and Pérez [1]. The final shortcut “\(A_2\) for \(S_k^+\)” which even avoided the random dyadic representation, was independently found by Hytönen–Lacey–Pérez [9] and Lerner [18].

Altogether, it now seems that the lower route to the \(A_2\) theorem, \(A_2 \text{ for } S_0^+ \rightarrow A_2 \text{ for } S_k^+ \rightarrow A_2 \text{ for CZO}, \) is the easiest one available as of today. On the other hand, it also seems that for a number of closely related results, it is necessary to take some additional steps.

Until recently, this was the case for the \(A_p\) theorem

\[
\|Tf\|_{L^p(w)} \leq c_{T,p} [w]_{A_p} + [w]_{A_p}^{1/(p-1)} \|f\|_{L^p(w)}, \quad 1 < p < \infty,
\]

which was originally deduced from the \(A_2\) theorem after an additional extrapolation argument from [3]. It can also be obtained directly from some paths of the above diagram by changing “\(A_2\) for \(T\)” to “\(A_p\) for \(T\)” and modifying the “testing \(T\)” conditions accordingly: This was achieved via the route [13] by Hytönen, Lacey, Martikainen et al. [8], and via [14] by Hytönen and Lacey [7]. However, recently Moen [20] found a short direct proof of “\(A_p\) for \(S_0^+\)” making the easy direct route (1.8) also available for the full \(A_p\) theorem (1.5).

Still, it seems that for the mixed \(A_p-A_{\infty}\) improvement of (1.6),

\[
\|Tf\|_{L^p(w)} \leq c_{T,p} [w]_{A_p}^{1/p} ([w]_{A_{\infty}}^{1/p'} + [\sigma]_{A_{\infty}}^{1/p}) \|f\|_{L^p(w)}, \quad \sigma := w^{1-p'}, \quad (1.7)
\]

an approach via the testing conditions and a weighted \(T1\) theorem is necessary. The bound (1.7) was first obtained by Hytönen–Pérez [11] for \(p = 2\) via (1.3), and then in general by Hytönen–Lacey [7] via (1.4). As pointed out by Lerner [17] Sec. 2.2], this can be somewhat simplified to the lower-left route

\[
\text{corona} \rightarrow \text{testing } S_0^+ \rightarrow A_p \text{ for } S_0^+ \rightarrow A_p \text{ for } S_k^+ \rightarrow A_p \text{ for CZO}, \quad (1.8)
\]

although it still needs many of the same ideas as (1.4) in the easier case of \(k = 0\).

The goal of this paper is to further simplify this lower-left route (1.8) to the \(A_p\) theorem and the mixed estimate (1.7). A detailed technical outline of this route is given in Section 2.
Sec. 3 A streamlined form of Lerner’s local oscillation formula, where a maximal function term from the original formulation is seen to be redundant.

Sec. 4 An example showing the sharpness of the known weak \((1,1)\) bound for the operators \(S_k^+\), a key lemma to prove that “\(A_2\) for \(S_0^+\)” implies “\(A_2\) for \(S_k^+\)”.

Sec. 5 A direct verification of the \(A_p\)-\(A_\infty\) testing conditions for \(S_0^+\). This means that the corona decomposition (and Bellman function) is once again avoided.

Sec. 6 A (slight) variant of the proof of the two-weight T1 theorem (for the simplest positive operators \(S_0^+\)) that is still necessary to follow this route.

The new insight might give hints towards some related open questions, of which I mention the following: For all cubes \(Q\), let \(X_Q, Y_Q\) by Banach function spaces on \(Q\), with duals \(X_Q', Y_Q'\) with respect to the duality \(\int_Q fg \, dx\), and let \(M_X f(x) \coloneqq \sup_{Q \ni x} \|f\|_{X_Q}\) and \(M_Y\) be defined similarly. In this set-up, Lerner [17] has shown that the following two-weight condition is sufficient for the boundedness of \(T(\cdot, \sigma) : L^p(w) \to L^p(w)\) for an arbitrary Calderón–Zygmund operator \(T\):

\[
(\sup_Q \|w^{1/p}\|_{X_Q} \|\sigma^{1/p'}\|_{Y_Q}) \|M_{X'}\|_{\mathcal{B}(L^p)} \|M_{Y'}\|_{\mathcal{B}(L^p')} < \infty.
\]

Note that the \(L^p\)-boundedness of \(M_X\), roughly means that “\(X_Q\) has a weaker norm than \(L^p(Q)\)” , hence “\(X_Q\) has a stronger norm than \(L^p(Q)\)”, so that \(\|w^{1/p}\|_{X_Q}\) is a bigger (“bumped up”) quantity than \(\langle w \rangle_{X_Q}^{1/p}\). Although partial progress was achieved by Cruz-Uribe, Reznikov and Volberg [2], it is open if the following one-sided bump condition is sufficient:

\[
(\sup_Q \|w^{1/p}\|_{X_Q} \|\sigma^{1/p'}\|_{Y_Q}) \|M_{Y'}\|_{\mathcal{B}(L^p')} + (\sup_Q \|w^{1/p}\|_{X_Q} \|\sigma^{1/p'}\|_{Y_Q}) \|M_{X'}\|_{\mathcal{B}(L^p)} < \infty.
\]

This last quantity, for \(w, \sigma \in A_\infty\) and \(X_Q = L^{p+\epsilon}(Q)\), \(Y_Q = L^{p'+\epsilon}(Q)\), is dominated by the product of \(A_p\) and \(A_\infty\) norms in [1,7]. Thus the one-sided bump conjecture would recover the \(A_p\) theorem, while the two-sided bump theorem does not.

2. Detailed outline of the lower-left route

2.A. Lerner’s formula. The key ingredient of the recent proofs of the \(A_2\) theorem is Lerner’s local oscillation formula from [17,19]. It involves the following concepts:

- The median of a measurable function \(f\) on a set \(Q\) is any real number \(m_f(Q)\) such that

\[
|Q \cap \{f > m_f(Q)\}| \leq \frac{1}{2}|Q|, \quad |Q \cap \{f < m_f(Q)\}| \leq \frac{1}{2}|Q|.
\]

- The decreasing rearrangement of \(f\) is the nonnegative function

\[
f^*(t) \coloneqq \inf \{ \alpha \geq 0 : \|\{ |f| > \alpha\} \| \leq t \} = \inf_{E, |E| \leq t} \|f 1_E^*\|_{\infty},
\]

where both infima are actually reached by \(\alpha = f^*(t)\) and \(E = \{ |f| > f^*(t)\}\).

- The oscillation of \(f\) on \(Q\), off a \(\lambda\)-fraction, is

\[
\omega_{\lambda}(f; Q) \coloneqq \inf_c \|1_Q(f-c)^*(\lambda|Q)\|.
\]

The key properties of these objects a summarized in the following simple lemma:
2.1. Lemma. We have the estimates
\[ |m_f(Q)| \leq (1_Q f)^*(\nu|Q|) \quad \forall \nu \in (0, \frac{1}{2}), \quad f^*(t) \leq \frac{1}{t} \|f\|_{L_1^\infty} \quad \forall t \in (0, \infty), \]
\[ (1_Q(f - m_f(Q)))^*(\nu|Q|) \leq 2\omega_{\nu}(f; Q) \quad \forall \nu \in (0, \frac{1}{2}). \]

2.2. Remark. These estimates are in general invalid for \( \nu = \frac{1}{2} \) and some medians. Indeed, consider \( Q = [0, 1) \subset \mathbb{R} \) and \( f = 1_{[0, \frac{1}{2})} \). Then any \( c \in [0, 1] \) is a median of \( f \) on \( Q \), but \( (1_Q f)^*(\frac{1}{2}|Q|) = 0 \), so the first bound is only true for the special median \( m_f(Q) = 0 \). Likewise, one can check with either \( c = 0 \) or \( c = 1 \) that \( \omega_{1/2}(f; Q) = 0 \), but \( (1_Q(f - m_f(Q)))^*(\frac{1}{2}|Q|) > 0 \) for all other medians \( m_f(Q) \in (0, 1) \).

In Section 2.1 below, I prove Lerner’s formula in the following form:

2.3. Theorem. For any measurable function \( f \) on a cube \( Q^0 \subset \mathbb{R}^d \), we have
\[ |f(x) - m_f(Q^0)| \leq 2 \sum_{L \in \mathcal{L}} \omega_\lambda(f; L) 1_L(x), \quad \lambda = 2^{-d-2}, \]
where \( \mathcal{L} \subset \mathcal{D}(Q^0) \) is sparse: there are pairwise disjoint major subsets \( E(L) \subset L \) with \( |E(L)| \geq \gamma |L| \). In fact, we can take \( \gamma = \frac{1}{2} \).

I now discuss the application of this formula in the proof of the \( A_2 \) theorem.

2.4. The reduction “\( A_2 \) for \( S^+_k \to A_2 \) for CZO”. With minor modifications, everything here extends to the maximal truncated singular integral \( T_\# \), and even, for a smaller class of Calderón–Zygmund operators, a stronger nonlinearity given by the so-called \( q \)-variation of singular integrals \( V^q_\# T \); see Hytönen–Lacey–Pérez [9]. But for the sake of simplicity I only present this discussion for the linear operator \( T \).

We consider a (say, bounded) compactly supported \( f \) and pick some \( Q^0 \supset \text{supp } f \). Lerner’s formula (applied to \( T f \)) guarantees that
\[ |T f(x)| \leq |m_{T f}(Q_0)| + 2 \sum_{L \in \mathcal{L}} \omega_\lambda(T f; L) 1_L(x). \]
Using the first and third estimates from Lemma 2.1, we see that
\[ |m_{T f}(Q_0)| \leq (1_{Q_0} T f)^*(\nu|Q_0|) \leq \frac{1}{\nu|Q_0|} \|T f\|_{L_1^\infty} \lesssim \frac{1}{|Q_0|} \|f\|_{L^1} = \int_{Q_0} |f| \]
by the boundedness \( T : L^1 \to L_1^\infty \), and the fact that \( \text{supp } f \subset Q_0 \) in the last two steps. We also have the following estimate that essentially goes back to Jawerth and Torchinsky [13]:

2.4. Lemma. Let the kernel \( K \) of \( T \) satisfy
\[ |K(x, y) - K(x', y)| \lesssim \Omega \left( \frac{|x - x'|}{|x - y|} \right) \frac{1}{|x - y|} \quad \forall |x - y| > 2|x - x'| > 0, \]
where the modulus of continuity \( \Omega : [0, \infty) \to [0, \infty) \) is increasing and subadditive with \( \Omega(0) = 0 \). Then \( \omega_\lambda(T f; Q) \lesssim \sum_{k=0}^\infty \Omega(2^{-k}) \int_{2^k Q} |f|. \)

To replace the concentric expansion \( 2^k Q \) by the dyadic ancestor \( Q^{(k)} \), we use the following geometric lemma, well known for \( k = 0 \), and proven in [9] as stated here.
2.5. Lemma. For any cube \( Q \subset \mathbb{R}^d \), there exists a shifted dyadic cube
\[
R \in \mathcal{D}^\alpha := \{2^{-k}((0,1)^d + m + (-1)^k \alpha) : k \in \mathbb{Z}, m \in \mathbb{Z}^d \},
\]
for some \( \alpha \in \{0, \frac{1}{3}, \frac{2}{3} \}^d \), such that \( Q \subseteq R \), \( 2^k Q \subseteq R^{(k)} \), \( \ell(R) \leq 6 \ell(Q) \).

Let us denote the index \( \alpha \) and cube \( R \) produced by this lemma by \( \alpha(Q,k) \) and \( R(Q,k) \). Thus we have
\[
|T f(x) - 1_Q f(x)| \lesssim \int_{Q_0} |f| \cdot 1_Q(x) + \sum_{L \in \mathcal{L}} \omega_L(T f; L) 1_L(x)
\]
\[
\lesssim \int_{Q_0} |f| \cdot 1_Q(x) + \sum_{L \in \mathcal{L}} \sum_{k=0}^\infty \Omega(2^{-k}) \int_{2^k L} |f| \cdot 1_L(x)
\]
\[
\lesssim \sum_{\alpha} \sum_{k=0}^\infty \Omega(2^{-k}) \sum_{R \in \mathcal{R}^\alpha_k} \int_{R(k)} |f| \cdot 1_R(x) =: \sum_{\alpha} \sum_{k=0}^\infty \Omega(2^{-k}) S^{+}_{\alpha,k}(f)(x)
\]
where
\[
\mathcal{R}^\alpha_k := \{ R(L,k) : L \in \mathcal{D}, \alpha(L,k) = \alpha \} \subset \mathcal{D}^\alpha
\]
and \( 1_{Q_0}(x) \int_{Q_0} |f| \) was absorbed into the sum \( \sum_{R \in \mathcal{R}^\alpha_k} \) with \( \alpha = k = 0 \).

The collections \( \mathcal{R}^\alpha_k \) are sparse; indeed, the sets \( E(L), L \in \mathcal{L} \), are pairwise disjoint, and
\[
|E(L)| \geq 2^{-1}|L| \geq 2^{-1}6^{-d}|R(L,k)|.
\]

Since \( Q_0 \supset \text{supp } f \) was arbitrary, from (2.6) we see that to estimate \( T f \) in a Banach function space, it clearly suffices to estimate the \( S^{+}_{\alpha,k}(f) \) in the same space, and this proves the claimed reduction.

2.6. The reduction “\( A_2 \rightarrow A_2 \) for \( S^{+}_{\alpha,k} \)”. This reduction is due to Lerner. Since we deal with positive operators, we can also restrict to nonnegative functions. Dualizing \( S^{+}_{\alpha,k} f \) with a function \( g \) of bounded support (choosing \( Q_0 \supset \text{supp } f, \text{supp } g \)), we have
\[
(S^{+}_{\alpha,k} f, g) = (f, (S^{+}_{\alpha,k})^t g),
\]
and we apply Lerner’s formula to \((S^{+}_{\alpha,k})^t g \) on a large enough cube \( Q_\alpha \in \mathcal{D}^\alpha \). This gives
\[
(S^{+}_{\alpha,k})^t g(x) \leq m((S^{+}_{\alpha,k})^t g(Q_\alpha) + \sum_{L \in \mathcal{Z}_{\alpha,k}} \omega_L((S^{+}_{\alpha,k})^t g, L) 1_L(x),
\]
and one can check that
\[
\omega_L((S^{+}_{\alpha,k})^t g; L) \lesssim (1 + k) \int_L g,
\]
where the essential ingredient is the weak \((1,1)\) bound for dyadic shifts, the only remaining component of the original proof of the \( A_2 \) conjecture [6 Prop. 5.1]. This holds for a general dyadic shift of complexity \( k \), and it is not in any substantial way easier for the particular shifts \((S^{+}_{\alpha,k})^* \), a special case treated in [18 Lemma 3.2].

2.7. Proposition. Any dyadic shift \( S_k \) of order \( k \) satisfies
\[
\|S_k f\|_{L^1,\infty} \lesssim (1 + k) \|f\|_{L^1}.
\]
The median term is handled similarly, and the result is that
\[(S^+_{\alpha,k})^t g(x) \lesssim (1 + k) \sum_{L \in \mathbb{Z}, k} \int_L g \cdot 1_L(x) =: (1 + k)(S^+_{0,(\alpha,k)})g(x),\]
where \(S^+_{0,(\alpha,k)}\) is a positive dyadic shift of complexity zero, a self-dual operator. Hence
\[\langle S^+_{\alpha,k}f, g \rangle = \langle f, (S^+_{\alpha,k})^t g \rangle \leq (1 + k)\langle S^+_{0,(\alpha,k)}f, g \rangle.\]

If we combine this with the previous reduction, we arrive at
\[|\langle Tf, g \rangle| \lesssim \sum_{k=0}^\infty \Omega(2^{-k})(1 + k)\langle S^+_{0,(\alpha,k)}(|f|), |g| \rangle,\]
and thus, for any Banach function space \(X\),
\[\|Tf\|_X \lesssim \left( \sum_{k=0}^\infty \Omega(2^{-k})(1 + k) \right) \sup_{S^+_0} \|S^+_0|f|\|_X,\]
where the supremum is over all positive dyadic shifts of complexity zero. As long as the above series converges (which is clear for the Hölder moduli \(\Omega(t) = t^p\), but also for some weaker moduli of continuity), if suffices to estimate the norm of \(S^+_0|f|\).

Note that \(\sum_{k=0}^\infty \Omega(2^{-k}) < \infty\) is the classical Dini condition \(\int_0^1 \Omega(t) \, dt/t < \infty\), where as we require a logarithmic strengthening \(\int_0^1 \Omega(t)(1 + \log(1/t)) \, dt/t < \infty\) caused by Proposition 2.7. While it is open if this logarithm is actually necessary in the final result, any attempt to remove it would have to circumvent the application of Proposition 2.7. Namely, this proposition is actually sharp, already for the special shifts \(S^+_{\alpha,k}\), as shown by example in Section 4.

2.D. The reduction “testing \(S^+_0 \to A_2\) for \(S^+_0^n\).” This is a direct application of the following elegant two-weight result of Lacey, Sawyer and Uriarte-Tuero [16]. A non-dyadic variant of this result goes back to Sawyer [27], and the stated dyadic version in the case \(p = q = 2\) to Nazarov–Treil–Volberg [21]. There is also a more recent simplification of the proof of the full theorem by Treil [28].

2.8. Theorem. Let \(\lambda_Q \geq 0\) be some coefficients, and
\[Sf := \sum_{Q \in \mathcal{P}} \lambda_Q(f)Q1_Q, \quad Sqf := \sum_{Q' \in \mathcal{P}, Q' \leq Q} \lambda_{Q'}(f)Q'1_{Q'} \quad (2.9)\]
be the associated positive dyadic shift of complexity zero, and its subshifts. For any two weights \(w\) and \(\sigma\), we have
\[\sup_{f} \frac{\|S(f)\|_{L^p(w)}}{\|f\|_{L^p(\sigma)}} \approx \sup_{Q \in \mathcal{P}} \frac{\|S_Q(\sigma)\|_{L^p(w)}}{\sigma(Q)^{1/p}} + \sup_{Q' \in \mathcal{P}} \frac{\|S_Q(w)\|_{L^p(\sigma)}}{w(Q)^{1/p'}}.\]

Observe that the shifts \(S^+_0\) correspond to the special case where \(\lambda_Q = 1_{\mathcal{L}}(Q)\) for some sparse family \(\mathcal{L} \subset \mathcal{P}\). Note also that with a special choice of the other weight, we have
\[\sup_{f} \frac{\|S(f)\|_{L^p(w)}}{\|f\|_{L^p(\sigma)}} = \sup_{f} \frac{\|Sf\|_{L^{p'}(w)}}{\|f\|_{L^{p'}(\sigma)}}, \quad \sigma = w^{1-p'}\]

Thus for example the mixed \(A_p-A_{\infty}\) bound
\[\|S^+_0 f\|_{L^p(w)} \lesssim [w]_{A_p}^{1/p'} ([w]_{A_{\infty}}^{1/p'} + [w^{1-p'}]_{A_{\infty}}^{1/p'}) \|f\|_{L^p(w)}\]
is a special case of the two-weight bound
\[ \|S_0^+(f \sigma)\|_{L^p(w)} \lesssim \|w, \sigma\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'} \|w^{1-p'}\|_{A_\infty}^{1/p'} \|f\|_{L^p(w)}, \]
where
\[ [w, \sigma]_{A_p} := \sup_{Q} \langle w \rangle^p_Q \sigma^{p-1}_Q, \]
which in turn follows, by Theorem 2.8 and symmetry, from the testing condition
\[ \|S_0^+(1_Q \sigma)\|_{L^p(w)} \lesssim [w, \sigma]_{A_p} [\sigma]_{A_\infty} \sigma(Q). \]
In Section 5 I give a new direct proof of the bound (2.11), without using either a corona decomposition or a Bellman function technique.

3. Proof of Lerner’s formula

For any family of pairwise disjoint subcubes $Q_j$ of $Q^0$, we can write the median Calderón–Zygmund decomposition
\[ 1_{Q^0}(f - m_f(Q^0)) = 1_{Q^0 \setminus \bigcup Q_j}(f - m_f(Q^0)) + \sum_j 1_{Q_j}(m_f(Q_j) - m_f(Q^0)) \]
\[ + \sum_j 1_{Q_j}(f - m_f(Q_j)). \quad (3.1) \]

We apply this with the following specific choice of the stopping cubes $Q_j$: they are the maximal dyadic subcubes of $Q^0$ with the property that
\[ \max_{Q' \in \text{ch}(Q_j)} |m_f(Q') - m_f(Q^0)| > (1_{Q^0}(f - m_f(Q^0)))^*(\lambda|Q^0|), \]
where $\text{ch}(Q) := \{ Q' \in \mathcal{D}(Q) : \ell(Q') = \frac{1}{2} \ell(Q) \}$ is the collection of dyadic children of $Q$. From the maximality it follows that $Q_j$ in place of $Q' \in \text{ch}(Q_j)$ satisfies the opposite estimate, and hence the second term on the right of (3.1) is dominated by
\[ 1_{\bigcup Q_j}(1_{Q^0}(f - m_f(Q^0)))^*(\lambda|Q^0|) \leq 1_{\bigcup Q_j^1} \cdot 2\omega_\lambda(f; Q^0). \]

On the other hand, if $x \in Q^0 \setminus \bigcup Q_j$, then the estimate opposite to (3.2) holds for all dyadic $Q' \ni x$. A lemma of Fujii [5] Lemma 2.2 (“a Lebesgue differentiation theorem for the median”) guarantees that $m_f(Q') \to f(x)$ as $Q' \to x$ for almost every $x$, and hence also the first term on the right of (3.1) is dominated by
\[ 1_{Q^0 \setminus \bigcup Q_j}(1_{Q^0}(f - m_f(Q^0)))^*(\lambda|Q^0|) \leq 1_{Q^0 \setminus \bigcup Q_j^1} \cdot 2\omega_\lambda(f; Q^0). \]

Altogether, we find that
\[ |1_{Q^0}(f - m_f(Q^0))| \leq 1_{Q^0} \cdot 2\omega_\lambda(f; Q^0) + \sum_j |1_{Q_j}(f - m_f(Q_j))|, \]
where the terms in the sum are of the same form as the left side, with $Q^0$ replaced by $Q_j$, and we are in a position to iterate. This gives
\[ |1_{Q^0}(f - m_f(Q^0))| \]
\[ \leq 1_{Q^0} \cdot 2\omega_\lambda(f; Q^0) + \sum_j 1_{Q_j} \cdot 2\omega_\lambda(f; Q_j^1) + \sum_{i} |1_{Q_i^j}(f - m_f(Q_i^j))| \]
\[ \leq \ldots \leq \sum_{k=0}^{m} \sum_j 1_{Q_j^k} \cdot 2\omega_\lambda(f; Q_j^k) + \sum_{i} |1_{Q_i^{m+1}}(f - m_f(Q_i^{m+1}))|, \quad (3.3) \]
where the cubes \( Q_{i}^{m+1} \) are dyadic subcubes of some \( Q_{j}^{m} \), chosen by a similar stopping criterion as the \( Q_{j}^{1} \) from \( Q_{0}^{0} \) in (3.2).

We claim that

\[
|E(Q_{j}^{m})| := \left| Q_{j}^{m} \setminus \bigcup_{i} Q_{i}^{m+1} \right| \geq \frac{1}{2} |Q_{j}^{m}|. \tag{3.4}
\]

This would show in particular that

\[
|\Omega^{m+1}| := \left| \bigcup_{i} Q_{i}^{m+1} \right| \leq 2^{-1} |\Omega^{m}| \leq \ldots \leq 2^{-m-1} |Q_{0}|,
\]

and hence the last term in (3.3) is supported on a set \( \Omega^{m+1} \subset \Omega^{m} \subset \ldots \subset Q_{0}^{0} \) of measure at most \( 2^{-m-1} |Q_{0}| \). As \( m \to \infty \), the support of this last term tends to a null set, and hence we obtain that

\[
|1_{Q_{0}}(f - m_{f}(Q_{0}))| \leq \sum_{k=0}^{\infty} \sum_{j} 1_{Q_{j}^{k}} : 2\omega_{k}(f; Q_{j}^{k})
\]

pointwise almost everywhere. This is the claimed formula with \( \mathcal{L} := \{Q_{j}^{k}\}_{k,j} \), and it only remains to check the sparseness condition (3.4). By symmetry, it suffices to consider \( m = 0 \).

We abbreviate \( f_{0} := f - m_{f}(Q_{0}) \). Then the stopping condition gives for some \( Q' \in \text{ch}(Q_{j}^{1}) \) and any \( \nu \in (0, \frac{1}{2}) \) the estimate

\[
\alpha := (1_{Q'_{0}}f_{0})^{*}(\lambda|Q_{0}|) \leq |m_{f_{0}}(Q')| \leq (1_{Q'_{0}}f_{0})^{*}(\nu|Q'|) \leq (1_{Q'_{0}}f_{0})^{*}(\nu 2^{-d}|Q_{j}^{1}|).
\]

Thus \( |Q_{j}^{1} \cap \{|f_{0}| > \alpha\}| \geq \nu 2^{-d}|Q_{j}^{1}| \), and hence

\[
\nu 2^{-d} \sum_{j} |Q_{j}^{1}| \leq \sum_{j} |Q_{j}^{1} \cap \{|f_{0}| > \alpha\}| \leq |Q_{0} \cap \{|f_{0}| > \alpha\}| \leq \lambda|Q_{0}| = 2^{-d-2}|Q_{0}|.
\]

Letting \( \nu \to \frac{1}{2} \), we get \( \sum_{j} |Q_{j}^{1}| \leq \frac{1}{2} |Q_{0}| \), which is the same as (3.4) for \( m = 0 \).

4. Sharpness of the weak \((1,1)\) estimate

I show by example (on \( \mathbb{R}^{1} \)) that the known bound

\[
\|(S_{k}^{+})^{*}f\|_{L^{1}} \lesssim (1 + k)\|f\|_{L^{1}}
\]

is sharp in terms of dependence on \( k \).

For \( L \in \mathcal{D} \), let \( L_{(j)} \) be the dyadic interval with \( (L_{(j)})^{(j)} = L \) and \( \inf L_{(j)} = \inf L \).

\[
\mathcal{L} := \{ L \in \mathcal{D} : L^{(k)} = [0,1) \}, \quad \mathcal{K} := \bigcup_{L \in \mathcal{L}} \{ L_{(j)} : j = 0, \ldots, k \}.
\]

Clearly this is a sparse family, with

\[
|E(K)| := \left| K \setminus \bigcup_{K' \in \mathcal{K}, K' \subseteq K} K' \right| \geq \frac{1}{2} |K| \quad \forall K \in \mathcal{K}.
\]

Consider \( f := 2^{k} \sum_{L \in \mathcal{L}} 1_{L^{(k)}} \) so that \( \|f\|_{L^{1}} = 2^{k} \sum_{L \in \mathcal{L}} |L^{(k)}| = \sum_{L \in \mathcal{L}} |L| = 1 \). For \( K = L_{(j)} \in \mathcal{K} \), we then have \( \int_{K} f = |L| \). Note that the intervals \( (L_{(j)})^{(k)} = \)
\( L^{(k-j)}, L \in \mathcal{L} \) cover every point of \([0,1]\) exactly \(2^{k-j}\) times. Hence

\[
\sum_{K \in \mathcal{X}} \frac{1_{K^{(k)}}}{|K^{(k)}|} \int_K f = \sum_{j=0}^{k} \sum_{L \in \mathcal{L}} \frac{1_{(L_{(j))}^{(k)}}}{|L^{(k)}|} \int_{L_{(j)}} f = \sum_{j=0}^{k} \sum_{L \in \mathcal{L}} \frac{1_{L^{(k-j)}}}{|L^{(k-j)}|} |L| = \sum_{j=0}^{k} 2^{j-k} \sum_{L \in \mathcal{L}} 1_{L^{(k-j)}} = \sum_{j=0}^{k} 2^{j-k} \times 2^{k-j} 1_{[0,1)} = (k+1)1_{[0,1)}.
\]

Then clearly \( \|(S_k^+)^* f\|_{L^1} = \|(k+1)1_{[0,1)}\|_{L^1} = k + 1 = (k+1)\|f\|_{L^1} \).

5. A direct verification of the testing conditions

We want to prove that

\[
\|S_Q \sigma\|^p_{L^p(w)} = \int_Q \left( \sum_{L \in \mathcal{L}} \langle \sigma \rangle_{L} 1_L \right)^p w \lesssim [w, \sigma]_{A_p} [\sigma]_{A_\infty} \sigma(Q),
\]

where \( L \) is a sparse family of cubes. Henceforth, we will suppress the summation condition \( \forall L \in \mathcal{L} \) with the understanding that all summation variables \( L, L', L_1, L_2, \ldots \) are always taken from the collection \( \mathcal{L} \). Recall that the two-weight \( A_p \) constant is defined by (2.10), and the \( A_\infty \) constant by

\[
[w, \sigma]_{A_\infty} := \sup_Q \frac{1}{\sigma(Q)} \int_Q M(1_Q \sigma).
\]

5.A. The \( A_2 \) case. This case is particularly simple:

\[
\int_Q \left( \sum_{L \in \mathcal{L}} \langle \sigma \rangle_{L} 1_L \right)^2 w = 2 \sum_{L \in \mathcal{L}} \sum_{L' \in \mathcal{L}} \langle \sigma \rangle_{L} \langle \sigma \rangle_{L'} |1_L| w \leq 2[w, \sigma]_{A_2} \sum_{L \in \mathcal{L}} \langle \sigma \rangle_{L} \sum_{L' \in \mathcal{L}} |L'| \lesssim [w, \sigma]_{A_2} \sum_{L \in \mathcal{L}} |\sigma_{L}| \sum_{L' \in \mathcal{L}} |L'| \lesssim [w, \sigma]_{A_2} \sum_{L \in \mathcal{L}} \inf_L M(\sigma_{L}) \cdot |E(L)| \lesssim [w, \sigma]_{A_2} \int_{E(L)} M(\sigma_{L}) \leq [w, \sigma]_{A_2} \int_Q M(\sigma_{L}) \leq [w, \sigma]_{A_2} [\sigma]_{A_\infty} \sigma(Q).
\]

5.B. The general \( A_p \) case. To “multiply out” the expression

\[
\left( \sum_{L \in \mathcal{L}} \langle \sigma \rangle_{L} 1_L \right)^p
\]

for a possibly non-integer value of \( p \in (1, \infty) \), we need the following observation:

5.1. Lemma. For all \( k \in \mathbb{N} \) and \( \alpha \in [0,1] \), and all nonnegative sequences of numbers \( a_i \), we have

\[
\left( \sum_i a_i \right)^{k+\alpha} \leq (k+1) \sum_{i_1, \ldots, i_k} a_{i_1} \cdots a_{i_k} \left( \sum_{j \leq \min\{i_1, \ldots, i_k\}} a_j \right)^{\alpha} \leq (k+1)! \sum_{i_1 \geq i_2 \geq \cdots \geq i_k \geq j} a_{i_1} \cdots a_{i_k} \cdot a_j^{\alpha}.
\]
Proof. Note that the second estimate is obvious, we only prove the first one. Let $A_i := \sum_{j \leq a_i}$, and for $i := (i_1, \ldots, i_k)$, write $a_i := a_{i_1} \cdots a_{i_k}$ and $A_i := A_{\min\{i_1, \ldots, i_k\}}$. Then the claim reads as $A_{k+\alpha}^k \leq (k+1) \sum_i a_i A_i^\alpha$. We consider the fraction

$$f(\alpha) := A_{k+\alpha}^k / \sum_i a_i A_i^\alpha,$$

and prove that it is bounded by $(k+1)$. Its derivative satisfies

$$f'(\alpha) = \frac{A_{k+\alpha}^k}{\left(\sum_i a_i A_i^\alpha\right)^2} \sum_i (\log A_i - \log A_i^\alpha) a_i A_i^\alpha \geq 0,$$

since $A_\infty \geq A_i$, and all quantities are nonnegative. Thus $f(\alpha) \leq f(1)$, and it is clear that this is at most $k+1$, since $A_{k+1}^k = \sum_{i_1, \ldots, i_{k+1}} a_{i_1} \cdots a_{i_{k+1}}$, and there are $k+1$ possible choices for which of the indices $i_1, \ldots, i_{k+1}$ is the smallest. □

For $1 \leq k < p \leq k + 1$, Lemma 5.2 gives

$$\int Q \left( \sum_{L \subseteq Q} \langle \sigma \rangle_{L_1}^p w \right) \lesssim \int Q \sum_{L_1 \geq \ldots \geq L_k \geq L_{k+1}} \langle \sigma \rangle_{L_1} \cdots \langle \sigma \rangle_{L_k} \langle \sigma \rangle_{L_{k+1}}^p L_{k+1} w$$

$$= \sum_{L_1 \geq \ldots \geq L_k \geq L_{k+1}} \langle \sigma \rangle_{L_1} \cdots \langle \sigma \rangle_{L_k} \langle \sigma \rangle_{L_{k+1}}^p L_{k+1} |L_{k+1}|$$

To proceed more smoothly, we record two further lemmas:

5.2. Lemma. For $\gamma \in [0, 1)$, we have $\sum_{L \subseteq P} \langle \sigma \rangle_{L_1}^\gamma |L| \lesssim \langle \sigma \rangle_{p}^\gamma |P|$.

Proof.

$$\sum_{L \subseteq P} \langle \sigma \rangle_{L_1}^\gamma |L| \lesssim \sum_{L \subseteq P} \langle \sigma \rangle_{L_1}^\gamma |E(L)| \lesssim \sum_{L \subseteq P} \inf_{L \subseteq P} M(w_{1P})^\gamma |E(L)| \lesssim \int_P M(w_{1P})^\gamma$$

$$\lesssim \|M(w_{1P})\|_{L^\infty}^\gamma |P|^{1-\gamma} \leq \|w_{1P}\|_{L^\infty}^\gamma |P|^{1-\gamma} = \langle \sigma \rangle_{p}^\gamma |P|. \quad \square$$

5.3. Lemma. For all $0 \leq \alpha \leq \beta(p-1) < \alpha + p - 1$, we have

$$\sum_{L \subseteq P} \langle \sigma \rangle_{L}^\alpha \langle \sigma \rangle_{p}^\beta |L| \lesssim |w, \sigma|_{Q}^{\alpha/(p-1)} \langle \sigma \rangle_{p}^{\beta-\alpha/(p-1)} |P|.$$

Proof.

$$\sum_{L \subseteq P} \langle \sigma \rangle_{L}^\alpha \langle \sigma \rangle_{p}^\beta |L| = \sum_{L \subseteq P} \left( \langle \sigma \rangle_{L}^{\alpha/(p-1)} \langle \sigma \rangle_{L}^{\beta/(p-1)} - \langle \sigma \rangle_{L}^{\alpha/(p-1)} \langle \sigma \rangle_{L}^{\beta-\alpha/(p-1)} |L| \right)$$

$$\leq |w, \sigma|_{Q}^{\alpha/(p-1)} \sum_{L \subseteq P} \langle \sigma \rangle_{L}^{\beta-\alpha/(p-1)} |L| \lesssim |w, \sigma|_{Q}^{\alpha/(p-1)} \langle \sigma \rangle_{p}^{\beta-\alpha/(p-1)} |P|,$$

where we used the assumption that $\beta - \alpha/(p-1) \in [0, 1)$ and Lemma 5.2. □

With $\alpha = p - k$, $\beta = 1$, Lemma 5.3 gives

$$\sum_{Q \supseteq L_1 \supseteq \cdots \supseteq L_k \supseteq L_{k+1}} \langle \sigma \rangle_{L_1} \cdots \langle \sigma \rangle_{L_k} L_{k+1} |L_{k+1}|$$

$$\lesssim |w, \sigma|_{Q}^{(p-k)/(p-1)} \sum_{Q \supseteq L_1 \supseteq \cdots \supseteq L_k} \langle \sigma \rangle_{L_1} \cdots \langle \sigma \rangle_{L_k} |L_{k+1}^{(k-1)/(p-1)}|L_k|$$
Then, using Lemma \ref{lem3} subsequently with $\alpha = 1$ and $\beta = j/(p-1)$, where $j = k - 1, \ldots, 1$, we obtain
\[
\sum_{Q \supseteq L_1 \supseteq \cdots \supseteq L_k} \langle \sigma \rangle_{L_1} \cdots \langle \sigma \rangle_{L_k} (w)^{(k-1)/(p-1)} |L_k|
\]
\[
\lesssim [w, \sigma]^{1/(p-1)}_{A_p} \sum_{Q \supseteq L_1 \supseteq \cdots \supseteq L_k-1} \langle \sigma \rangle_{L_1} \cdots \langle \sigma \rangle_{L_k-1} (w)^{(k-2)/(p-1)} |L_{k-1}|
\]
\[
\lesssim \cdots \lesssim [w, \sigma]^{(k-1)/(p-1)}_{A_p} \sum_{Q \supseteq L_1} \langle \sigma \rangle_{L_1} |L_1|,
\]
and here
\[
\sum_{L_1 \subseteq Q} \langle \sigma \rangle_{L_1} |L_1| \lesssim \sum_{L_1 \subseteq Q} \inf_{L_1} M(\sigma_{1Q}) \cdot |E(L_1)| \leq \int_Q M(\sigma_{1Q}) \leq [\sigma]_{A_{\infty}} \sigma(Q).
\]
Thus, altogether, we have checked that
\[
\int_Q \left( \sum_{L \subseteq Q} \langle \sigma \rangle_{L} \right)^p w \lesssim [w, \sigma]^{(p-k)/(p-1)}_{A_p} [w, \sigma]^{(k-1)/(p-1)}_{A_p} [\sigma]_{A_{\infty}} \sigma(Q),
\]
and the total power of $[w, \sigma]_{A_p}$ is one, as claimed.

6. The two-weight \textit{T1 theorem for positive operators}

To be in line with the \textit{T1} literature, I now write $T$ and $T_Q$ instead of $S$ and $S_Q$ as defined in \cite{22}. I restate and prove the main estimate of Theorem 2.8 in the following form:

6.1. \textbf{Theorem.} For $1 < p \leq q < \infty$ and any two weights $\sigma, \omega$, we have
\[
\|T(\cdot, \sigma)\|_{L^p(\sigma) \to L^q(\omega)} \leq 20(p' q^{\mathcal{T}} + pq^{\mathcal{T}^*}),
\]
\[
\mathcal{T} := \sup_{Q \in \mathcal{D}} \|T_Q(\sigma)\|_{L^q(\omega)}, \quad \mathcal{T}^* := \sup_{Q \in \mathcal{D}} \|T_Q(\omega)\|_{L^{p'}(\sigma)}.
\]

The proof below follows the main lines of Treil’s argument \cite{22}, with one key difference: rather than splitting the summation over $Q \in \mathcal{D}$ in the expansion of $\langle T(f \sigma), g \omega \rangle$ into parts according to an \textit{ad hoc} criterion such as $\sigma(Q) \langle f \rangle_Q \geq \omega(Q) \langle g \rangle_Q$, I simply apply the “parallel corona” decomposition from the recent work of Lacey, Sawyer, Shen and Uriarte-Tuero \cite{15} on the two-weight boundedness of the Hilbert transform. Thus, the proof below can also been seen as a toy introduction to some of the innovative techniques of \cite{15}.

\textbf{Proof.} We analyse the pairing
\[
\langle T(f \sigma), g \omega \rangle = \sum_{Q \in \mathcal{D}} \lambda_Q \langle f \rangle_Q \langle g \rangle_Q (Q) = \sum_{Q \in \mathcal{D}} \lambda_Q \langle f \rangle_Q \sigma(Q) (Q) \langle g \rangle_Q (Q).
\]  

(6.2)

It suffices to make a uniform estimate over all subseries with $Q \subseteq Q_0$ for some large dyadic cube $Q_0$, and we may assume that both $f, g \geq 0$ are supported in $Q_0$. Then we define the collections of principal cubes $\mathcal{F}$ for $(f, \sigma)$ and $\mathcal{G}$ for $(g, \omega)$. Namely,
\[
\mathcal{F} := \bigcup_{k=0}^{\infty} \mathcal{F}_k, \quad \text{where} \quad \mathcal{F}_0 := \{Q_0\},
\]
\[
\mathcal{F}_{k+1} := \bigcup_{F \in \mathcal{F}_k} \text{ch}_{\mathcal{F}}(F), \quad \text{ch}_{\mathcal{F}}(F) := \{Q \subseteq F \text{ maximal s.t. } \langle f \rangle_Q^\circ > 2 \langle f \rangle_F^\circ\},
\]
and analogously for \( \mathcal{G} \). Observe that

\[
\sum_{F' \in \text{ch}_\mathcal{G}(F)} \sigma(F') \leq \sum_{F' \in \text{ch}_\mathcal{G}(F)} \frac{\int_{E'} f \sigma}{2(f')^2} \leq \frac{\int_{E} f \sigma}{2(f)^2} = \sigma(F),
\]

and hence

\[
\sigma(E_\mathcal{G}(F)) := \sigma \left( F \setminus \bigcup_{F' \in \text{ch}_\mathcal{G}(F)} F' \right) \geq \frac{1}{2} \sigma(F),
\]

where the sets \( E_\mathcal{G}(F) \) are pairwise disjoint.

We further define the stopping parents

\[
\pi_\mathcal{G}(Q) := \min \{ F \supseteq Q : F \in \mathcal{G} \}, \quad \pi(Q) := (\pi_\mathcal{G}(Q), \pi_\mathcal{G}(Q)).
\]

Then we rearrange the series in (6.2) as

\[
\sum_Q \sum_{F' \in \mathcal{G}, \pi(Q) = (F,G)} \sum_{Q'} = \sum_F \sum_{G \subseteq F} \sum_{\pi(Q) = (F,G)} + \sum_F \sum_{G \subseteq F} \sum_Q, \quad (6.3)
\]

where we observed that if the inner sum over \( Q : \pi(Q) = (F,G) \) is not empty, then there is some \( Q \subseteq F \cap G \), hence \( F \cap G \neq \emptyset \), and thus \( G \subseteq F \) or \( F \subseteq G \). By symmetry, we concentrate on the first case only.

Consider a \( Q \) with \( \pi(Q) = (F,G) \) for some \( G \subseteq F \). Then

\[
\int_Q g \omega = \int_{Q \cap \text{ch}_\mathcal{G}(F)} g \omega + \sum_{F' \in \text{ch}_\mathcal{G}(F)} \int_{Q \cap F'} g \omega.
\]

If \( Q \cap F' \neq \emptyset \), then either \( F' \subseteq Q \) or \( Q \subseteq F' \). But the latter is not possible, since it would imply that \( \pi_\mathcal{G}(Q) \subseteq F' \subseteq F \), contradicting \( \pi(Q) = (F,G) \). Thus, for the nonzero terms in the last summation in (6.3), we must have \( F' \subseteq Q \subseteq G \) for some \( G \in \mathcal{G} \) with \( G \in \mathcal{G} \). Since \( Q \subseteq G \subseteq F \) and \( \pi_\mathcal{G} Q = F \), also \( \pi_\mathcal{G} G = F \). Thus, we may actually restrict the summation to

\[
\text{ch}^*_\mathcal{G}(F) := \{ F' \in \text{ch}_\mathcal{G}(F) : \pi_\mathcal{G} \pi_\mathcal{G}(F') = F \}.
\]

So in fact

\[
\int_Q g \omega = \int_{Q \cap \text{ch}_\mathcal{G}(F)} g \omega + \sum_{F' \in \text{ch}_\mathcal{G}(F)} \int_{F'} g \omega
\]

\[
= \int_Q \left( g1_{\text{ch}_\mathcal{G}(F)} + \sum_{F' \in \text{ch}_\mathcal{G}(F)} \langle g \rangle_{F'} 1_{F'} \right) \omega =: \int_Q g_F \omega.
\]

Thus we find that

\[
\sum_{G \subseteq F} \sum_{\pi(Q) = (F,G)} \lambda_Q \langle f \rangle_G \langle \sigma \rangle_Q \int_Q g \omega \leq 2(f)^2 \sum_{G \subseteq F} \sum_{\pi(Q) = (F,G)} \lambda_Q \langle \sigma \rangle_Q \int_Q g_F \omega
\]

\[
\leq 2(f)^2 \sum_{Q \subseteq F} \lambda_Q \langle \sigma \rangle_Q \int_Q g_f \omega = 2(f)^2 \langle f \rangle_G \langle T_f \sigma \rangle, g_F \omega
\]

\[
\leq 2(f)^2 \| T_f \sigma \|_{L^q(\omega)} \| g_F \|_{L^q(\omega)} \leq 2(f)^2 \langle \sigma(F) \rangle^{1/p} \| g_F \|_{L^q(\omega)}^p.
\]
and hence
\[
\sum_{F \subseteq F} \sum_{\pi(Q) = (F,G)} \lambda_Q f_Q^\omega_F \int_Q g_F \leq 2 \mathcal{T} \sum_{F} (\langle f \rangle_F^\omega_F)^{1/p} \left\| g_F \right\|_{L^{q'}(\omega)}^{1/p'} \quad (6.5)
\]
\[
\leq 2 \mathcal{T} \left( \sum_{F} (\langle f \rangle_F^\omega_F)^{p} \left\| g_F \right\|_{L^{q'}(\omega)}^{p'} \right)^{1/p'}.
\]

For the first factor, using \( \sigma(F) \leq 2\sigma(E_F(F)) \) and \( \langle f \rangle_F^\omega_F \leq \inf_F M_\sigma f \), and the disjointness of the \( E_F(F) \), we see that
\[
\left( \sum_{F} (\langle f \rangle_F^\omega_F)^{p} \left\| g_F \right\|_{L^{q'}(\omega)}^{p'} \right)^{1/p'} \leq 2 \left( \int (M_\sigma f)^{p} \sigma \right)^{1/p} \leq 2p' \left\| f \right\|_{L^{p}(\sigma)},
\]
(6.6)

Using \( \left\| f \right\|_{L^{q'}} \leq \left\| f \right\|_{L^{q'}}^{p'} \) for \( q \geq p \), it only remains to estimate
\[
\left( \sum_{F} g_F \left\| g \right\|_{L^{q'}(\omega)}^{q'} \right)^{1/q'} = \left( \sum_{F} \left\| g \right\|_{E_F(F)}^{q} \right)^{1/q'} \quad (6.7)
\]
By the pairwise disjointness of the test \( E_F(F) \), it is immediate that
\[
\sum_{F} \left\| g \right\|_{E_F(F)}^{q} \left\| g \right\|_{L^{q'}(\omega)} \leq \left\| g \right\|_{L^{q'}(\omega)}^{q'}.
\]

For the remaining double sum, we use the definition of \( ch^*_F(F) \) to reorganize:
\[
\sum_{F} \sum_{F' \subseteq ch^*_F(F)} \sum_{\pi \subseteq G = F, \pi(F') = (F,G)} (\langle g \rangle_{F'}^\omega_{F})^{q'} \omega(F') \leq 2 (\sum_{F} \sum_{\pi \subseteq G = F, \pi(F') = (F,G)} (\langle g \rangle_{F'}^\omega_{F})^{q'} \omega(F') \leq 2 \sum_{F} \sum_{\pi \subseteq G = F} (\langle g \rangle_{\pi(G)}^\omega_{F})^{q'} \omega(G) \leq 2^{1+q'} \sum_{F} \sum_{\pi \subseteq G = F} (\langle g \rangle_{\pi(G)}^\omega_{F})^{q'} \omega(G) \leq 2^{1+q'} \int (\sigma(F))^q \omega \leq 2^{1+q'} (\left\| g \right\|_{L^{q'}(\omega)})^{q'}.
\]
Substituting back to (6.7), we have that
\[
\left( \sum_{F} \left\| g_F \right\|_{L^{q'}(\omega)}^{q'} \right)^{1/q'} \leq \left\| g \right\|_{L^{q'}(\omega)} + 2^{1/q'+1} q \left\| g \right\|_{L^{q'}(\omega)} \leq 5 q \left\| g \right\|_{L^{q'}(\omega)}.
\]
(6.8)

The combination of (6.3), (6.6) and (6.8) shows that the first half of \( \langle T(f \sigma), g \omega \rangle \), according to the splitting (6.3) is estimated by
\[
2 \mathcal{T} \cdot 2p' \left\| f \right\|_{L^{p}(\sigma)} \cdot 5 q \left\| g \right\|_{L^{q'}(\omega)} \leq 20 \cdot \mathcal{T} \cdot p' q \cdot \left\| f \right\|_{L^{p}(\sigma)} \left\| g \right\|_{L^{q'}(\omega)}
\]
We conclude by symmetry of the assumptions and the splitting (6.3). \( \square \)

REFERENCES
[1] D. Cruz-Uribe, J. María Martell, and C. Pérez. Sharp weighted estimates for approximating dyadic operators. Electron. Res. Announc. Math. Sci., 17:12–19, 2010.
[2] D. Cruz-Uribe, A. Reznikov, and A. Volberg. Logarithmic bump conditions and the two weight boundedness of Calderón-Zygmund operators. arXiv:1112.0676 (2011).
[3] Y. Do and M. Lacey. Weighted bounds for variational Fourier series. arXiv:1207.1150 (2012).
[4] O. Dragičević, L. Grafakos, M. C. Pereyra, and S. Petermichl. Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Mat.*, 49(1):73–91, 2005.

[5] N. Fujii. A condition for a two-weight norm inequality for singular integral operators. *Studia Math.*, 49(1):73–91, 1991.

[6] T. Hytönen. The sharp weighted bound for general Calderón–Zygmund operators. *Ann. of Math.* (2), 175(3):1473–1506, 2012.

[7] T. Hytönen and M. T. Lacey. The $A_p - A_\infty$ inequality for general Calderón–Zygmund operators. *Indiana Univ. Math. J.* (to appear).

[8] T. Hytönen, M. T. Lacey, H. Martikainen, T. Orponen, M. C. Reguera, E. T. Sawyer, and I. Uriarte-Tuero. Weak and strong type estimates for maximal truncations of Calderón-Zygmund operators on $A_p$ weighted spaces. *J. Anal. Math.* (to appear).

[9] T. Hytönen, M. T. Lacey, and C. Pérez. Sharp weighted bounds for the $q$-variation of singular integrals. *Bull. London Math. Soc.* (to appear).

[10] T. Hytönen, M. T. Lacey, M. C. Reguera, E. T. Sawyer, I. Uriarte-Tuero, and A. Vagharshakyan. Weak and strong type $A_p$ estimates for Calderón-Zygmund operators. *J. Anal. Math.* (to appear).

[11] T. Hytönen and C. Pérez. Sharp weighted bounds involving $A_\infty$. *Anal. PDE*. (to appear).

[12] T. Hytönen, C. Pérez, S. Treil, and A. Volberg. Sharp weighted estimates for dyadic shifts and the $A_2$ conjecture. *J. Reine Angew. Math.* (to appear).

[13] B. Jawerth and A. Torchinsky. Local sharp maximal functions. *J. Approx. Theory*, 43(3):231–270, 1985.

[14] M. T. Lacey, S. Petermichl, and M. C. Reguera. Sharp $A_2$ inequality for Haar shift operators. *Math. Ann.*, 348(1):127–141, 2010.

[15] M. T. Lacey, E. T. Sawyer, C.-Y. Shen, and I. Uriarte-Tuero. Two weight inequality for the Hilbert transform: A real variable characterization. arXiv:1201.4319 (2012).

[16] M. T. Lacey, E. T. Sawyer, and I. Uriarte-Tuero. Two weight inequalities for discrete positive operators. arXiv:0911.3437 (2009).

[17] A. K. Lerner. On an estimate of Calderón–Zygmund operators by dyadic positive operators. *J. Anal. Math.* (to appear).

[18] A. K. Lerner. A simple proof of the $A_2$ conjecture. *Int. Math. Res. Not. IMRN*. (to appear).

[19] A. K. Lerner. A pointwise estimate for the local sharp maximal function with applications to singular integrals. *Bull. Lond. Math. Soc.*, 42(5):843–856, 2010.

[20] K. Moen. Sharp weighted bounds without testing or extrapolation. *Arch. Math. (Basel).* (to appear).

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

[22] F. Nazarov, S. Treil, and A. Volberg. The $Tb$-theorem on non-homogeneous spaces. *Acta Math.*, 190(2):151–239, 2003.

[23] F. Nazarov and A. Volberg. A simple sharp weighted estimate of the dyadic shifts on metric spaces with geometric doubling. arXiv:1104.4893 (2011).

[24] C. Pérez, S. Treil, and A. Volberg. On $A_2$ conjecture and corona decomposition of weights. arXiv:1006.2630 (2010).

[25] S. Petermichl. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical $A_p$ characteristic. *Amer. J. Math.*, 129(5):1355–1375, 2007.

[26] S. Petermichl and A. Volberg. Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular. *Duke Math. J.*, 112(2):281–305, 2002.

[27] E. T. Sawyer. A characterization of two weight norm inequalities for fractional and Poisson integrals. *Trans. Amer. Math. Soc.*, 308(2):533–545, 1988.

[28] S. Treil. A remark on two weight estimates for positive dyadic operators. arXiv:1201.4155 (2012).

[29] A. Vagharshakyan. Recovering singular integrals from Haar shifts. *Proc. Amer. Math. Soc.*, 138(12):4303–4309, 2010.

Department of Mathematics and Statistics, P.O.B. 68 (Gustaf Hällströmin katu 2b), FI-00014 University of Helsinki, Finland
E-mail address: tuomas.hyttonen@helsinki.fi