THE $T(b)$ THEOREM AND ITS VARIANTS

T. TAO

ABSTRACT. We survey the various local and global variants of $T(1)$ and $T(b)$ theorems which have appeared in their literature, and outline their proofs based on a local wavelet coefficient perspective. This survey is based on recent work in [4] with Pascal Auscher, Steve Hofmann, Camil Muscalu, and Christoph Thiele.

1. INTRODUCTION

The purpose of this expository article is to survey the various types of $T(1)$ and $T(b)$ theorems which have been used to prove boundedness of Calderón-Zygmund kernels. These theorems have many applications to Cauchy integrals and analytic capacity, and also to elliptic and accretive systems; see e.g. [1], [6], [9], [13]. However we will not concern ourselves with applications here, and focus instead on the mechanics of proof of these theorems. In particular we present a slightly non-standard method of proof, based on pointwise estimates of wavelet coefficients.

For simplicity we shall work just on the real line $\mathbb{R}$, with the standard Lebesgue measure dx. For applications one must consider much more complicated settings, for instance one-dimensional sets endowed with Hausdorff measure, but we will not discuss these important generalizations in this article[1], and instead concentrate our attention on the distinction between $T(1)$ and $T(b)$ theorems, and between local and global variants of these theorems.

To avoid technicalities (in particular, in justifying whether integrals actually converge) we shall restrict all functions and operators to be real-valued, and only consider Calderón-Zygmund operators T of the form

$$Tf(x) := \int K(x,y)f(y)\ dy$$

[1]In particular we will not discuss the important recent extensions of the $T(1)$ and $T(b)$ theory to non-doubling measures, see e.g. [24], [23].

where $K(x,y)$ is a smooth, compactly supported kernel obeying the estimates

$$|K(x,y)| \lesssim 1/|x-y| \tag{1}$$

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \lesssim |x-x'|/|x-y|^2 \tag{2}$$

for all distinct x, x', y. Here we use $A \lesssim B$ or $A = O(B)$ to denote the estimate $A \leq CB$ for some constant C. An example of such an operator would be the Hilbert transform $Hf(x) := p.v. \int \frac{1}{x-y} f(y)\, dy$, after first truncating the kernel $p.v. \frac{1}{x-y}$ to be smooth and compactly supported.

Of course, since we have declared K to be smooth and compactly supported, the operator T is *a priori* bounded on $L^p(\mathbb{R})$ for $1 < p < \infty$. However, the basic problem is to find useful conditions under which one can obtain more quantitative bounds on the L^p operator norm of T (e.g. depending only on p and the implicit constants in (1), (2) and some additional data, but not on the smoothness and compact support assumptions on K).

The remarkable $T(1)$ theorem of David and Journé [12] gives a complete characterization of this problem:

Theorem 1.1 (Global $T(1)$ theorem)**.** [12] *If T is a Calderón-Zygmund operator such that*

- *(Weak boundedness property) One has $\langle T\chi_I, \chi_I \rangle = O(|I|)$ for all intervals I. Here $|I|$ denotes the length of I.*
- *We have $\|T(1)\|_{BMO} \lesssim 1$.*
- *We have $\|T^*(1)\|_{BMO} \lesssim 1$.*

Then T is bounded on L^p, $1 < p < \infty$

$$\|Tf\|_p \lesssim \|f\|_p$$

(the implicit constants depend on p) and we have the L^∞ to BMO bounds

$$\|Tf\|_{BMO} \lesssim \|f\|_\infty$$
$$\|T^*f\|_{BMO} \lesssim \|f\|_\infty.$$

The remarkable thing about this theorem is that while the conclusion asserts that T and T^* obey certain bounds for all L^p functions f (including when $p = \infty$), whereas the hypotheses only require these type of bounds for very specific L^p functions f. The point is that the kernel bounds (1), (2) impose severe constraints on the behavior of T. Informally, these bounds assert that the bilinear form $\langle Tf, g\rangle$ is small when f and g have widely separated supports, and is especially small when one also assumes that f or g has mean zero. This allows us in many cases to estimate $\langle Tf, g\rangle$ by replacing f with its mean on a certain interval (i.e. by replacing f with its projection onto constant functions), or by performing a similar replacement for g. By doing this at all scales simultaneously (by use of such tools as the wavelet decomposition), we can eventually control $\langle Tf, g\rangle$ by quantities involving $T(1)$, $T^*(1)$, or quantities such as $\langle T\chi_I, \chi_I\rangle$. In more precise terms, the standard proof of Theorem 1.1 proceeds by decomposing T into an easily estimated "diagonal" component (controlled by $\langle T\chi_I, \chi_I\rangle$, and two paraproducts, one related to $T(1)$ and one related to $T^*(1)$, which can then be estimated with the aid of the Carleson embedding theorem. In this note we give an alternate proof (in a model case) based on pointwise estimates of wavelet coefficients.

Despite giving a complete answer to our problem, Theorem 1.1 is not completely satisfactory for two reasons. Firstly, the conditions $\|T(1)\|_{BMO} \lesssim 1$ and $\|T^*(1)\|_{BMO} \lesssim 1$ are global rather than local, in that the constant function 1 and the nature of the BMO norm are both spread out over all intervals of space simultaneously, in a way that the weak boundedness property is not. Thus if one were only interested in whether T is bounded on acertain subset Ω of $\mathbb{R}$, the hypotheses of this theorem would be inappropriate. Secondly, the theorem is rather inflexible in that one must test T and T^* against the function 1, rather than some other functions which might be more convenient to apply T and T^* to.

The first objection is easily addressed. Indeed, it is quite straightforward (and well-known) to see that the global $T(1)$ theorem is equivalent to the following local version:

Theorem 1.2 (Local $T(1)$ theorem)**.** *If T is a Calderón-Zygmund operator such that*

- $\int_I |T\chi_I(x)|^2\, dx \lesssim |I|$ *for all intervals* I.
- $\int_I |T^*\chi_I(x)|^2\, dx \lesssim |I|$ *for all intervals* I.

Then the conclusions of Theorem 1.1 hold.

We sketch the derivation of the global $T(1)$ theorem from the local $T(1)$ theorem as follows. Suppose that $T(1) \in BMO$, then for any interval I we would have

$$\|T(1) - [T(1)]_I\|_{L^2(I)} \lesssim |I|^{1/2},$$

where $[f]_I := \frac{1}{|I|}\int_I f$ denotes the mean of f on I. On the other hand, from (1), (2) and a direct calculation one can show that

$$\|T(1-\chi_I) - [T(1-\chi_I)]_I\|_{L^2(I)} \lesssim |I|^{1/2},$$

and hence that

$$\|T(\chi_I) - [T(\chi_I)]_I\|_{L^2(I)} \lesssim |I|^{1/2}.$$

On the other hand, from the weak boundedness property we have $[T(\chi_I)]_I = O(1)$, and thus we have verified the first property of Theorem 1.2. The second property is verified similarly and we are done. The reverse implication (that the global theorem implies the local version) is similar and we do not detail it here.

As the name suggests, the local $T(1)$ theorem can be localized. Here is one sample localization result:

Theorem 1.3 (Localized $T(1)$ theorem)**.** *Let $\Omega \subset \mathbb{R}$ be an open set. If T is a Calderón-Zygmund operator such that*

- $\int_I |T\chi_I(x)|^2 \, dx \lesssim |I|$ *for all intervals I such that $I \cap \Omega \neq \emptyset$.*
- $\int_I |T^*\chi_I(x)|^2 \, dx \lesssim |I|$ *for all intervals I such that $I \cap \Omega \neq \emptyset$.*

Then T is bounded on $L^p(\Omega)$ for $1 < p < \infty$, in the sense that

$$\|Tf|_\Omega\|_{L^p(\Omega)} \lesssim \|f\|_{L^p(\Omega)}$$

for all $f \in L^p(\Omega)$.

Corresponding L^p and $L^\infty \to BMO$ estimates can also be made, but are a little technical due to the need to define BMO on domains.

We prove this theorem in a model case in Section 3. Unfortunately, this localization is imperfect, since the intervals I in the hypotheses are allowed to partially extend outside Ω; we do not know how to resolve this issue to make the localization more satisfactory.

Now we discuss the second objection to the $T(1)$ theorem, that the choice of function 1 is fixed. Certainly we cannot replace 1 by an arbitrary L^∞ function, since if one replaces 1 by (for instance) the zero function 0 then the hypotheses become trivial, and the theorem absurd. Or if one replaces 1 by a function which vanishes on a large set, then the hypothesis yields us very little information about T on that set, and so one would not expect to conclude boundedness on T inside that set. However, if one replaces 1 by functions which in some sense "never vanish", then one can generalize the $T(1)$ theorem:

Theorem 1.4 (Global $T(b)$ theorem)**.** [13] *Let b, b' be bounded in $L^\infty(\mathbb{R})$ and such that we have the pseudo-accretivity condition*

$$|[b]_I|, |[b']_I| \gtrsim 1$$

for all intervals I. If T is a Calderón-Zygmund operator such that

- *(Modified weak boundedness property) One has $\langle T(b\chi_I), (b'\chi_I)\rangle = O(|I|)$ for all intervals I.*

- *We have* $\|T(b)\|_{BMO} \lesssim 1$.
- *We have* $\|T^*(b')\|_{BMO} \lesssim 1$.

Then T *is bounded on* L^2

$$\|Tf\|_2 \lesssim \|f\|_2$$

and we have the L^∞ *to* BMO *bounds*

$$\|Tf\|_{BMO} \lesssim \|f\|_\infty$$
$$\|T^*f\|_{BMO} \lesssim \|f\|_\infty.$$

The conditions b, b' be bounded in $L^\infty(\mathbb{R})$ can be relaxed slightly to requiring b, b' to just be bounded in BMO; see [4]. This theorem can be proven by using wavelet systems adapted to b, b' respectively [5]. In the special "one-sided" case when $b' = 1$ then there is an alternate proof going through the global $T(1)$ theorem and the heuristic that when $T^*(1) \in BMO$, we have the approximation $\langle T(b), \phi_I\rangle \approx [b]_I \langle T(1), \phi_I\rangle$ for all intervals I and all bump functions ϕ_I of mean zero adapted to I; see [25] (and also [4]).

This $T(b)$ theorem can also be localized, so that instead of having a single function b supported globally, we only need a localized function b_I for each I. Also, each b_I only has to satisfy a single accretivity condition:

Theorem 1.5 (Local $T(b)$ theorem)**.** [4] *If* T *is a Calderón-Zygmund operator such that*

- *For every interval* I*, there exists a function* $b_I \in L^2(I)$ *with*

$$\int_I |b_I|^2 + |Tb_I|^2 \lesssim |I|$$

  which obeys the accretivity condition

$$|[b_I]_I| \gtrsim 1.$$

- *For every interval* I*, there exists a function* $b'_I \in L^2(I)$ *with*

$$\int_I |b'_I|^2 + |T^*b'_I|^2 \lesssim |I|$$

  which obeys the accretivity condition

$$|[b'_I]_I| \gtrsim 1.$$

Then T *is bounded on* L^2

$$\|Tf\|_2 \lesssim \|f\|_2$$

and we have the L^∞ *to* BMO *bounds*

$$\|Tf\|_{BMO} \lesssim \|f\|_\infty$$

$$\|T^* f\|_{BMO} \lesssim \|f\|_\infty.$$

Strictly speaking, the theorem in [4] was only proven in a model case in which "perfect" kernel cancellation conditions were assumed, but the argument extends without significant difficulty to general Calderón-Zygmund kernels. This theorem is a variant of an earlier local $T(b)$ theorem of [9], which was in a much more general setting but required L^∞ conditions on b_I, Tb_I, b'_I, $T^* b'_I$ rather than L^2 conditions. One can also replace the L^2 conditions with L^p and $L^{p'}$ conditions, see [4].

In the "one-sided case", when $T^* 1 \in BMO$, then one does not need the second hypothesis (involving the b'_I), and the proof can again proceed via the $T(1)$ theorem, again using the heuristic that when $T^* 1 \in BMO$, we have the approximation $\langle Tb_I, \phi_I\rangle \approx [b_I]_I \langle T(1), \phi_I\rangle$ when ϕ_I has mean zero and is adapted to I. (cf. the local $T(b)$ theorems in [6], [1], [2]). However in the general case it appears that one is forced to use tools such as adapted wavelet systems.

2. A dyadic model

In order to simplify the exposition, we shall replace our "continuous" Calderón-Zygmund operators with a "discrete" dyadic model; this model will be much cleaner to manipulate because of the absence of several minor error terms, but will already capture the essence of the arguments.

We define a *dyadic interval* to be any interval of the form $[2^j k, 2^j(k+1)]$ where j, k are integers. We will always ignore sets of measure zero, and so will not distinguish between open, closed, or half-open dyadic intervals. If I is an interval, we use $|I|$ to denote its length and I^{left} and I^{right} to denote left and right halves of I respectively, and refer to I^{left} and I^{right} as *siblings*. We define the *Haar wavelet* ϕ_I to be the L^2-normalized function

$$\phi_I := |I|^{-1/2}(\chi_{I^{left}} - \chi_{I^{right}}).$$

As is well-known, these wavelets form an orthonormal basis of $L^2(\mathbb{R})$. For any locally integrable f, we define the *wavelet transform* Wf, defined on the space of dyadic intervals, by $Wf(I) := \langle f, \phi_I\rangle$. Thus W is an isometry from L^2 to l^2.

Observe that if T obeys (1), (2), and I and J are disjoint dyadic intervals, then the quantity $\langle T\phi_I, \phi_J\rangle$ decays quickly in the separation of I and J (for instance, we have the estimates

$$\langle T\phi_I, \phi_J\rangle \lesssim \frac{|I|}{|J|}\left(1 + \frac{\operatorname{dist}(I,J)}{|J|}\right)^{-2}$$

when $|J| \geq |I|$). We now introduce a simplified model operator in which the quantity $\langle T\phi_I, \phi_J\rangle$ not only decays, but in fact *vanishes* when I, J are disjoint.

Definition 2.1. *A perfect dyadic Calderón-Zygmund operator T is an operator of the form*

$$Tf(x) := \int K(x,y)f(y)\ dy$$

where $K(x,y)$ is a bounded, compactly supported function which obeys the kernel condition

$$|K(x,y)| \lesssim \frac{1}{|x-y|} \tag{3}$$

and the perfect dyadic Calderón-Zygmund conditions

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| = 0 \tag{4}$$

whenever $x, x' \in I$ and $y \in J$ for some disjoint dyadic intervals I and J. Equivalently, K is constant on all rectangles $\{I \times J : I, J$ are siblings$\}$.

As a consequence of (4) we see that Tf_I is supported on I whenever f_I is supported on I with mean zero, and similarly with T replaced by T^*. In particular we have $\langle T\phi_I, \phi_J\rangle = 0$ whenever I, J are disjoint.

We define the dyadic BMO norm by

$$\|f\|_{BMO_\Delta} := \sup_I \left(\frac{1}{|I|}\int_I |f - [f]_I|^2\right)^{1/2} = \sup_I \left(\sum_{J\subseteq I} Wf(J)^2\right)^{1/2}$$

where I ranges over all dyadic intervals. This norm is very close to the usual BMO norm, see [18].

3. Pointwise wavelet coefficient estimates

The purpose of this section is to give a somewhat non-standard proof of the $T(1)$ theorem in the model case of perfect dyadic Calderón-Zygmund operators. This differs from the usual proof (involving paraproduct decomposition and Carleson embedding) in that it relies on pointwise estimates for wavelet coefficients, and also tackles the endpoint $L^\infty \to BMO$ estimates first (instead of the usual procedure of first establishing L^2 bounds.

Let T be a perfect dyadic Calderón-Zygmund operator, and let $f \in L^2(\mathbb{R})$ be a function. In order to study the L^2 or BMO norm of Tf, it makes sense to study the wavelet coefficients $W(Tf)(J)$. It turns out that one can write the coefficients of Tf rather explicitly in terms of the corresponding coefficients of f, $T\chi_J$, $T^*\chi_J$ and $fT^*\chi_J$:

Lemma 3.1. [4] *We have*

$$\begin{aligned} W(Tf)(J) = {} & [f]_J W(T\chi_J)(J) - [f]_J W(T^*\chi_J)(J) + W(fT^*\chi_J)(J) \\ & - \frac{2}{|J|}(\langle T\chi_{J_{left}}, \chi_{J_{right}}\rangle + \langle T\chi_{J_{right}}, \chi_{J_{left}}\rangle) Wf(J). \end{aligned} \tag{5}$$

From (3) we thus have the useful pointwise estimate

$$\begin{aligned} |W(Tf)(J)| \lesssim {} & |[f]_J||W(T\chi_J)(J)| + |[f]_J||W(T^*\chi_J)(J)| \\ & + |W(fT^*\chi_J)(J)| + |Wf(J)| \end{aligned} \tag{6}$$

on the wavelet coefficient of Tf. The pointwise estimates (6) are reminiscent of the standard pointwise sharp function estimates $(Tf)^\#(x) \lesssim Mf(x)$ for Calderón-Zygmund kernels bounded on L^p; see [26].

Proof First observe that if f is vanishes on J then both sides of (5) vanish (since $W(Tf)(J) = \langle f, T^*\phi_J\rangle$ and $T^*\phi_J$ is supported on J). Thus by linearity we may assume that f is supported on J.

Now suppose that f is equal to χ_J, then $Wf(J) = 0$, and (5) collapses to

$$W(T\chi_J)(J) = W(T\chi_J)(J) - W(T^*\chi_J)(J) + W(\chi_J T^*\chi_J)(J)$$

which is clearly true. By linearity we may thus assume that f is orthogonal to χ_J.

Now suppose that f is equal to ϕ_J. Then (5) simplifies to

$$\langle T\phi_J, \phi_J\rangle = \frac{1}{|J|}\langle T\chi_J, \chi_J\rangle - \frac{2}{|J|}(\langle T\chi_{J_{left}}, \chi_{J_{right}}\rangle + \langle T\chi_{J_{right}}, \chi_{J_{left}}\rangle),$$

which is easily verified by expanding out T. By linearity we may thus assume that f is also orthogonal to ϕ_J. In particular, we can now assume that f has mean zero on both J_{left} and J_{right}. After some re-arranging, the claim now collapses to

$$\langle Tf, \phi_J\rangle = \langle T(f\phi_J), \chi_J\rangle.$$

If f is supported on J_{left}, then the claim follows since ϕ_J is constant on the support of f and on Tf. Similarly if f is supported on J_{right}. The claim now follows by linearity. ∎

Of course, this identity took full advantage of the perfect cancellation (4). However in the continuous case there are still analogues of (6), but with some additional error terms which are easily treated. (Basically, the right-hand side will not just contain terms related to J, but also terms related to intervals J' which are “close” in size and location to J, but with an additional weight which decays fairly quickly as J' moves further away from J). One can also replace the Haar wavelet with smoother wavelets in order to avoid certain (harmless) logarithmic divergences which can come up in this procedure. We will not pursue the details.

With (6) in hand it is a fairly easy matter to prove the dyadic analogue of Theorem 1.2:

Theorem 3.2 (Local dyadic $T(1)$ theorem). *If T is a perfect dyadic Calderón-Zygmund operator such that*

- $\int_I |T\chi_I(x)|^2\, dx \lesssim |I|$ *for all dyadic intervals I.*
- $\int_I |T^*\chi_I(x)|^2\, dx \lesssim |I|$ *for all dyadic intervals I.*

Then T is bounded on L^2

$$\|Tf\|_2 \lesssim \|f\|_2$$

and we have the L^∞ to dyadic BMO bounds

$$\|Tf\|_{BMO_\Delta} \lesssim \|f\|_\infty$$

$$\|T^*f\|_{BMO_\Delta} \lesssim \|f\|_\infty.$$

Of course, this local theorem is equivalent to its global counterpart by the argument sketched in the introduction.

Proof First we show that T maps L^∞ into dyadic BMO. Fix $f \in L^\infty$; we may normalize $\|f\|_\infty = 1$. We need to show the Carleson measure condition

$$\sum_{J \subseteq I} |W(Tf)(J)|^2 \lesssim |J| \tag{7}$$

for all dyadic intervals I.

By a standard iteration procedure (extremely common in the study of Carleson measures or BMO; we will use it again in the next section) we may restrict the summation to those intervals J for which

$$|[T^*\chi_I]_J| \leq C \tag{8}$$

for some large constant C. We sketch the reason for this is as follows. By Cauchy-Schwartz, we see that $\int_J |T^*\chi_I|^2 \geq C^2|J|$ whenever (8) fails. Since we are assuming the bound $\int_I |T^*\chi_I|^2 \lesssim |I|$, the intervals J for which (8) fails can therefore only cover a set of measure at most $\frac{1}{100}|I|$, if C is chosen sufficiently large. One can then iterate away those contributions in the usual manner (e.g. by assuming inductively that (7) already holds for all sub-intervals of I).

Fix I; we will implicitly assume that (8) holds. Now we use (6) to decompose the wavelet coefficients $W(Tf)(J)$.

We write

$$W(fT^*\chi_J)(J) = W(fT^*\chi_I)(J) - W(fT^*\chi_{I\setminus J})(J).$$

From (4) we see that $T^*\chi_{I\setminus J}$ is constant on J, thus

$$W(fT^*\chi_{I\setminus J})(J) = ([T^*\chi_I]_J - [T^*\chi_J]_J)Wf(J).$$

From the hypotheses and Cauchy-Schwarz we have $[T^*\chi_J]_J = O(1)$. From this and (8) we thus have

$$|W(fT^*\chi_J)(J)| \lesssim |W(fT^*\chi_I)(J)| + |Wf(J)|.$$

Also, since $T^*\phi_J$ is supported on J, we have

$$W(T\chi_J)(J) = W(T\chi_I)(J),$$

and similarly for T^*. Since $[f]_J = O(\|f\|_\infty) = O(1)$, we thus see from (6) that we have the pointwise estimate

$$\begin{aligned}|W(Tf)(J)| \lesssim &|W(T\chi_I)(J)| + |W(T^*\chi_I)(J)| \\ &+ |W(fT^*\chi_I)(J)| + |W(f)(J)|.\end{aligned}$$

By hypothesis, the functions $T\chi_I$, $T^*\chi_I$, $(T^*\chi_I)f$, and f all have an $L^2(I)$ norm of $O(|I|^{1/2})$. The claim (7) follows since the ϕ_J are orthonormal in $L^2(I)$. This shows that T maps L^∞ into BMO_Δ.

A similar argument shows that T^* maps L^∞ to BMO_Δ. One then concludes L^p boundedness by Fefferman-Stein interpolation (see e.g. [26]) and a standard reiteration argument (starting, for instance, with the apriori L^2 boundedness and then improving this by repeated interpolation; cf. [29]). Alternatively, one could use L^p Calderón-Zygmund techniques as in [16] to first establish weak L^p bounds and then interpolate. We omit the details ∎

We now sketch how the above theorem localizes to a domain $\Omega \subset \mathbb{R}$. Specifically, we prove

Corollary 3.3 (Dyadic localized $T(1)$ theorem)**.** *Let $\Omega \subset \mathbb{R}$ be an open set. Let $\mathbf{I}$ be the set of dyadic intervals I such that $I \cap \Omega \neq \emptyset$. If T is a Calderón-Zygmund operator such that*

- $\int_I |T\chi_I(x)|^2\, dx \lesssim |I|$ *for all* $I \in \mathbf{I}$.
- $\int_I |T^*\chi_I(x)|^2\, dx \lesssim |I|$ *for all* $I \in \mathbf{I}$.

Then T is bounded on $L^p(\Omega)$ for $1 < p < \infty$, in the sense that

$$\|Tf|_\Omega\|_{L^p(\Omega)} \lesssim \|f\|_{L^p(\Omega)}$$

for all $f \in L^p(\Omega)$.

Proof Let Π be the projection

$$\Pi f := \sum_{I\in\mathbf{I}} Wf(I)\phi_I.$$

Equivalently, we have $\Pi f(x) = f(x)$ if $x \in \Omega$, and $\Pi f(x) = [f]_I$ when $x \notin \Omega$, where I is the smallest dyadic interval containing x and intersecting Ω.

We consider the operator $\Pi T \Pi$. It is easy to verify that this operator is also a perfect dyadic Calderón-Zygmund operator. We claim that $\Pi T \Pi$ verifies the hypotheses of Theorem 3.3. The claim then follows since $\Pi T \Pi f(x) = T \Pi f(x) = T f(x)$ for all $x \in \Omega$ and $f \in L^p(\Omega)$.

Let I be a dyadic interval. We shall verify that

$$\int_I |\Pi T \Pi \chi_I(x)|^2 \, dx \lesssim |I|.$$

First suppose that $I \in \mathbf{I}$. Then $\Pi \chi_I = \chi_I$. Also, on I one can estimate Π by the Hardy-Littlewood maximal function on I. The claim then follows from the hypotheses and the Hardy-Littlewood maximal inequality.

Now suppose that $I \notin \mathbf{I}$. Let J be the smallest dyadic interval in $\mathbf{I}$ which contains J. Then $\Pi \chi_I = \frac{|I|}{|J|} \chi_J$, and $\Pi f = [f]_J$ on I. Thus it will suffice to show that

$$\frac{1}{|J|} \int_J \left| \frac{|I|}{|J|} T \chi_J \right|^2 \lesssim |I|.$$

But this follows from hypothesis since $|I| \leq |J|$. ∎

It is also possible to establish the above corollary directly by using variants of the pointwise estimate (6), combined with the square-function wavelet characterization of L^p and some Calderón-Zygmund theory, but we will not detail this here.

The operator Π is an example of a *phase space projection*, to a region of phase space known as a "tree". Such localized estimates for Calderón-Zygmund operators are very useful in the study of multilinear objects such as the bilinear Hilbert transform and the Carleson maximal operator: see e.g. [15], [28], [19], [20], [22], etc. (For a further discussion of the connection between Carleson measures, Calderón-Zygmund theory, and phase space analysis, see [4]).

3.1. **Adapted Haar bases, and $T(b)$ theorems.** We now turn to $T(b)$ theorems, starting with the global $T(b)$ theorem in Theorem 1.4. We shall consider the dyadic model case when T is a perfect dyadic Calderón-Zygmund operator, and assume that b, b' are bounded and obey the pseudo-accretivity conditions

$$|[b]_I|, |[b']_I| \gtrsim 1$$

for all dyadic intervals I.

Ideally, one would like to have an estimate like (6), but with $T(\chi_J)$ and $T^*(\chi_J)$ replaced by $T(b\chi_J)$ and $T^*(b'\chi_J)$:

$$\begin{aligned}|W(Tf)(J)| \lesssim |[f]_J||W(T(b\chi_J))(J)| + |[f]_J||W(T^*(b'\chi_J))(J)| \\ + |W(fT^*(b'\chi_J))(J)| + |Wf(J)|.\end{aligned} \tag{9}$$

This is because one can then approximate $T(b\chi_J)$ and $T^*(b'\chi_J)$ by $T(b)$ and $T^*(b')$ by arguments similar to that used to prove Theorem 3.2.

Unfortunately, such an estimate does not seem to hold in general. However, if one replaces the Haar wavelet system by modified Haar wavelet systems adapted to b and b', then one can recover a formula similar to (9), as follows.

For each interval I, we define the adapted Haar wavelet ϕ_I^b (introduced in [10]; see also [5]) by

$$\phi_I^b := \phi_I - \frac{Wb(I)}{[b]_I}\frac{\chi_I}{|I|}. \tag{10}$$

The wavelet ϕ_I^b no longer has mean zero, but still obeys the weighted mean zero condition

$$\int b\phi_I^b = 0. \tag{11}$$

As a consequence we see that

$$\int \phi_I^b b \phi_J^b = 0 \text{ for all } I \neq J. \tag{12}$$

A calculation gives the identity

$$\int \phi_I^b b \phi_I^b = \frac{2}{[b]^{-1}_{I_{left}} + [b]^{-1}_{I_{right}}}.$$

From the pseudo-accretivity and boundedness properties, we thus see that ϕ_P^b has the non-degeneracy property

$$\left| \int \phi_P^b b \phi_P^b \right| \sim 1. \tag{13}$$

Define the dual adapted Haar wavelet ψ_I^b by

$$\psi_I^b := \frac{\phi_I^b b}{\int \phi_I^b b \phi_I^b}.$$

By (12), (13) we thus have that $\langle \psi_I^b, \phi_J^b \rangle = \delta_{IJ}$ where δ is the Kronecker delta. In particular we have the representation formula

$$f = \sum_I W_b f(I) \psi_I^b \tag{14}$$

(formally, at least), where the adapted wavelet coefficients $W_b f(I)$ are defined by

$$W_b f(I) := \langle f, \phi_I^b \rangle.$$

From some computation and the Carleson embedding theorem one can eventually verify the orthogonality property

$$\Big(\sum_I |W_b f(I)|^2 \Big)^{1/2} \sim \|f\|_2 \tag{15}$$

(see e.g. [26], or [4]). Thus the adapted wavelet transform W_b has most of the important properties that W has.

The analogue of (9) is

Lemma 3.4. [4] *If f is bounded in L^∞, then*

$$\begin{aligned} |W_b(bT^*f)(I)| \lesssim\, & |W_b(bT^*(b'\chi_I))(I)| + |W_b(b'T(b\chi_I))(I)| \\ & + |W_b(fT(b\chi_I))(I)| + |Wf(I)| + |W(b')(I)|. \end{aligned}$$

Proof We can write

$$Wb(bT^*f)(I) = \langle f, T(b\phi_I^b) \rangle = \langle f, T\psi_I^b \rangle.$$

Since ψ_I^b has mean zero, $T\psi_I^b$ is supported on I. Thus the left-hand side does not depend on the values of f outside I. Similarly for the right-hand side. Thus we may assume that f is supported on I.

The claim is clearly true when f is equal to a bounded multiple of $b'\chi_I$ (since the left-hand side is then bounded by the first term on the right-hand side). Since $|[b']_I| \sim 1$, every bounded function f on I can be split as the sum of a bounded multiple of $b'\chi_I$ and a bounded function on I with mean zero. Thus it will suffice to prove the estimate

$$|W_b(bT^*f)(I)| \lesssim |W_b(fT(b\chi_I))(I)| + |Wf(I)|$$

for functions f of mean zero on I.

We can write

$$W_b(bT^*f)(I) = \langle \phi_I^b T^*(f), b\chi_I \rangle.$$

Since

$$\langle T^*(\phi_I^b f), b\chi_I \rangle = \langle fT(b\chi_I), \phi_I^b \rangle = W_b(fT(b\chi_I))(I)$$

it will suffice to prove the commutator estimate

$$\langle \phi_I^b T^*(f), b\chi_I \rangle - \langle T^*(\phi_I^b f), b\chi_I \rangle = O(|Wf(I)|).$$

Recall that ϕ_I^b is constant on I_{left} and I_{right}. Thus if f is supported on I_{left} with mean zero, then the commutator vanishes (since $T^*(\phi_I^b f) = \phi_I^b(I_{left})T^*(f)$ is supported on I_{left}). Similarly if f is supported on I_{right} with mean zero. Since we are already assuming that f is supported on

I with mean zero, it thus suffices to verify the estimate when f is the standard Haar wavelet ϕ_I:

$$\langle \phi_I^b T^*(\phi_I), b\chi_I\rangle - \langle T^*(\phi_I^b \phi_I), b\chi_I\rangle = O(1).$$

Throwing the T^* on the other side, we see the claim will follow from Cauchy-Schwarz and the estimates

$$\int_I |T(b\chi_{I_{left}})|^2, \int_I |T(b\chi_{I_{right}})|^2, \int_I |T(b\chi_I)|^2 \lesssim |I|.$$

We just show the last estimate, as the first two are proven similarly. From weak boundedness we have

$$\Big|\int_I T(b\chi_I)b'\Big| \lesssim |I|^{1/2}.$$

Since b' is in L^∞ with $|[b']_I| \sim 1$, it thus suffices to prove the BMO estimate

$$\int_I |T(b\chi_I) - [T(b\chi_I)]_I|^2 \lesssim |I|.$$

On the other hand, since $T(b)$ is in BMO by hypothesis, we have

$$\int_I |T(b) - [T(b)]_I|^2 \lesssim |I|.$$

The claim follows since $T(b(1-\chi_I))$ is constant on I by (4). ∎

From Lemma 3.4 one can show (as in the proof of Theorem 1.2) that the wavelet coefficients $W_b(bT^*f)(I)$ obeys the Carleson measure estimate

$$\sum_{J\subseteq I} |W_b(bT^*f)(J)|^2 \lesssim |I|.$$

(As in the proof of Theorem 1.2, it will be convenient to first remove the intervals where the mean of $T(b\chi_I)$ is large, in order to estimate $T(b\chi_J)$ by $T(b\chi_I)$ effectively.)

From this Carleson measure estimate, (15) and the mean-zero property of the wavelets $\psi_J^b = b\phi_J^b$, we can show that $T^*f \in BMO$. Thus T^* maps L^∞ to BMO. Similar arguments give the bound for T, and the L^p bounds then follow from interpolation as before. This allows one to prove the global $T(b)$ theorem, Theorem 1.4.

We now briefly indicate how the above proof of the global $T(b)$ theorem can be localized to yield Theorem 1.5; the details though are rather complicated and can be found in [4]. In the above argument we gave pointwise estimates on wavelet coefficients $W_b(bT^*f)(J)$, which were in turn used to prove Carleson measure estimates. Now, however, we do not have a single b to work with, instead having a localized b_I assigned to each interval I. We would now like to have a Carleson measure estimate of the form

$$\sum_{J \subseteq I} |W_{b_I}(b_I T^* f)(J)|^2 \lesssim |I|. \tag{16}$$

Suppose for the moment that we could attain (16). If b_I obeyed the pseudo-accretivity condition $|[b_I]_J| \gtrsim 1$ for all $J \subseteq I$, and the boundedness condition $\int_J |b_I|^2 \lesssim |J|$ for all $J \subseteq I$, then one could use (15) to then show the BMO estimate

$$\int_I |T^* f - [T^* f]_I|^2 \lesssim |I| \tag{17}$$

which would show that T^* mapped L^∞ to BMO as desired.

Unfortunately, our assumptions only give us that b_I has large mean on all of I:

$$|[b_I]_I| \gtrsim 1.$$

This does not preclude the possibility of b_I having small mean on some sub-interval J of I (for instance, b_I could vanish on some sub-interval). However, because we have the L^2 bound $\int_I |b_I|^2 \lesssim I$, it does mean that $[b_I]_J$ cannot vanish for a large proportion of intervals J. To make this more rigorous, fix $0 < \varepsilon \ll 1$, and let $\Omega_I \subseteq I$ be the union of all the intervals $J \subseteq I$ for which $|[b_I]_J| \lesssim \varepsilon$. Then by a simple covering argument (splitting Ω_I as the disjoint union of the maximal dyadic intervals J in Ω_I) we have

$$\Big| \int_{\Omega_I} b_I \Big| \lesssim \varepsilon |\Omega_I| \lesssim \varepsilon |I|$$

and thus (if ε is sufficiently small)

$$\Big| \int_{I \backslash \Omega_I} b_I \Big| \gtrsim |I|.$$

From Cauchy-Schwarz and the L^2 bound on b_I we thus have

$$|I \backslash \Omega_I| \gtrsim |I|.$$

Thus there is a significant percentage of I for which one does not have the problem of b_I having small mean. Furthermore, by a similar argument one can also show that on a slightly smaller significant percentage of I, one also does not have the problem of $\int_J |b_I|^2$ being much larger than J. Thus there is a non-zero "good" portion of I where the function b_I behaves as if itwere pseudo-accretive, and on which Carleson estimates such as (16) would yield BMO bounds. The proof of (17) then proceeds by first estimating the integral on this "good" portion of I, and using a standard iteration argument to deal with the remaining "bad" portion of I, which has measure at most $c|I|$ for some $c < 1$; the point being that the geometric series $1 + c + c^2 + \dots$ converges. Details can be found in [4]; similar arguments can also be found in [1], [6] (and indeed a vector-valued, higher-dimensional version of this trick of removing the small-mean intervals is crucial to the resolution of the Kato square root problem in higher dimensions [1], [2]).

It remains to prove (16), at least when J is restricted to the set where b_I behaves like a pseudo-accretive function. Obviously one would like to have Lemma 3.4 holding (but with b, b' replaced by b_I, b'_I of course). However, an inspection of the proof shows that to do this one needs bounds of the form $\int_J |b_I|^2 \lesssim |J|$, $\int_J |b'_I|^2 \lesssim |J|$, and $|[b'_I]_J| \gtrsim 1$; in other words b_I and b'_I have to behave like pseudo-accretive functions[2] on J. Fortunately, by arguments similar to the ones given above, one can show that the bad intervals J, for which the above bounds fail, do not cover all of I, and there is a fixed proportion of I which is "good", in the sense that (16) holds when localized to this good set. One then performs yet another iteration argument on the "bad" portion of I to conclude (16).

This concludes the sketch of the proof of Theorem 1.5 in the case of perfect dyadic Calderón-Zygmund operators; further details can be found in [4]. It is an interesting question as to whether this proof technique can also produce localized estimates, perhaps similar to Theorem 1.3. One model instance of such a theorem is Theorem 29 of [9], which says informally that if E is a compact subset of an Ahlfors-David regular set with positive analytic capacity, then E has large intersection with another Ahlfors-David regular set for which the Cauchy integral is bounded. (This is somewhat analogous with a hypothesis that $T(b), T^*(b) \in BMO$ for some b which is sometimes, but not always, pseudo-accretive, and a conclusion that T is bounded when restricted to a reasonably large set Ω).

4. Acknowledgements

This expository paper is based on joint work [4] with Pascal Auscher, Steve Hofmann, Camil Muscalu, and Christoph Thiele. The author also thanks Mike Christ for some very helpful suggestions. The author is a Clay Prize Fellow and is supported by the Packard Foundation.

[2]Actually, one also needs lower bounds for the magnitude of of $[b'_I]_{J_{left}}$ and $[b'_I]_{J_{right}}$. This introduces some technicalities involving "buffer" intervals - intervals J for which b_I has large mean, but such that b_I has small mean on one of the subintervals J_{left}, J_{right}. These intervals are rather annoying to deal with but are fortunately not too numerous; see [4].

References

[1] Auscher, P., Hofmann, S., M^{c}Intosh. A., Tchamitchian, P., *The Kato square root problem for higher order elliptic operators and systems on R^n*, submitted.

[2] Auscher, P., Hofmann, S., Lacey, M. M^{c}Intosh. A., Tchamitchian, P., *The solution of the Kato square root problem for second order elliptic operators on R^n*, submitted.

[3] Auscher, P., Hofmann, S., Lewis, J.. A., Tchamitchian, P., *Extrapolation of Carleson measures and the analyticity of Kato's square root operators*, to appear, Acta Math.

[4] Auscher, P., Hofmann, S., Muscalu, C., Tao, T., Thiele, C., *Carleson measures, trees, extrapolation, and $T(b)$ theorems*, submitted.

[5] Auscher, P, Tchamitchian, P., *Bases d'ondelettes sur les courbes corde-arc, noyau de Cauchy et espaces de Hardy associés*, Rev. Mat. Iberoamericana **5** [1989], 139–170.

[6] Auscher, P, Tchamitchian, P., *Square root problem for divergence operators and related topics.* Astérisque No. 249, (1998), viii+172 pp.

[7] Calderón, C., *On commutators of singular integrals*, Studia Math. **53** [1975], 139–174.

[8] Carleson, L. *On convergence and growth of partial Fourier series*, Acta. Math. **116** 135–157, [1966].

[9] Christ, M., *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral.* Colloq. Math. 60/61 [1990], 601–628.

[10] Coifman, R.R., Jones, P., and Semmes, S. *Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves.* J. Amer. Math. Soc. **2** 553–564 [1989].

[11] David, G., *Wavelets and singular integrals on curves and surfaces.* , Lecture Notes in Mathematics, 1465. Springer-Verlag, Berlin, 1991.

[12] David, G., Journé, J., *A boundedness criterion for generalized Calderón- Zygmund operators*, Ann. of Math. **120** [1984], 371–397.

[13] David, G., Journé, J., and Semmes, S., *Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation*, Rev. Mat. Iberoamericana **1** [1985], 1–56.

[14] David, G. and Semmes, S., *Analysis of and on uniformly rectifiable sets*, Mathematical Surveys and Monographs, 38. American Mathematical Society, Providence, RI, 1993. xii+356 pp. ISBN: 0-8218-1537-7

[15] Fefferman, C. *Pointwise convergence of Fourier series*, Ann. of Math. (2) 98, pp. 551–571 [1973].

[16] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36.

[17] Fefferman, C. and Stein, E. M., *H^p spaces of several variables*, Acta Math. **129** [1972], 137–193.

[18] Garnett, J., Jones, P., *BMO from dyadic BMO*, Pacific J. Math. **99** [1982], 351–371.

[19] Lacey, M. and Thiele, C. *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$.* Ann. Math. 146, pp. 693-724, [1997]

[20] Lacey, M. and Thiele, C. *On Calderon's conjecture.* Ann. Math. 149, pp. 475-496, [1999]

[21] Lacey, M. and Thiele, C. *A proof of boundedness of the Carleson operator*, Math. Res. Lett **7** 361–370, [2000]

[22] Muscalu, M., Tao, T., Thiele, C. L^p *estimates for the "biest"*, submitted, Math. Annalen.

[23] Nazarov, F., Treil, S., Volberg, A., *Accretive system Tb theorems on nonhomogeneous spaces*, preprint.

[24] Nazarov, F., Treil, S., Volberg, A., *Tb theorems on nonhomogeneous spaces*, preprint.

[25] Semmes, S., *Square function estimates and the* $T(b)$ *theorem*, Proc. Amer. Math. Soc. **110**, 721–726 [1990]

[26] Stein, E. *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton University Press, Princeton, [1993]

[27] Thiele, C. *Time-Frequency Analysis in the Discrete Phase Plane*, PhD Thesis, Yale 1995.

[28] Thiele, C. *On the Bilinear Hilbert transform.* Universität Kiel, Habilitationsschrift [1998]

[29] Wolff, T.H. *A note on interpolation spaces.* Harmonic analysis (Minneapolis, Minn., 1981), pp. 199–204, Lecture Notes in Math., 908, Springer, Berlin-New York [1982]

Department of Mathematics, UCLA, Los Angeles CA 90095-1555 USA
E-mail address: tao@math.ucla.edu