

Carleson's theorem with quadratic phase functions

by

MICHAEL T. LACEY (Atlanta, GA)

Abstract. It is shown that the operator below maps L^p into itself for $1 < p < \infty$.

$$Cf(x) := \sup_{a,b} \left| \text{p.v.} \int f(x-y) e^{i(ay^2+by)} \frac{dy}{y} \right|.$$

The supremum over b alone gives the famous theorem of L. Carleson [2] on the pointwise convergence of Fourier series. The supremum over a alone is an observation of E. M. Stein [12]. The method of proof builds upon Stein's observation and an approach to Carleson's theorem jointly developed by the author and C. M. Thiele [7].

1. The main result. Consider the maximal operator

$$\mathcal{C}_d f(x) = \sup_{\deg(p)=d} \left| \text{p.v.} \int f(x-y) e(p(y)) \frac{dy}{y} \right|$$

in which d is an integer, p is a polynomial of degree d , $e(u) := e^{\pi i u}$, f is a Schwartz function and the integral is understood in the principal value sense. This definition is motivated principally by the case $d = 1$. $\mathcal{C}_1 f$ controls the maximal partial Fourier integrals of f and it extends to a bounded map from L^p into itself for $1 < p < \infty$. The critical contribution here is L. Carleson's proof [2] of the boundedness of \mathcal{C}_1 from L^2 into weak L^2 . The L^p version was established by R. Hunt [4]. Also see [3, 7].

It is natural to ask if the same results hold for larger values of d —this is a conjecture due to E. M. Stein. The point of this paper is to demonstrate that this is the case for $d = 2$.

1.1. THEOREM. \mathcal{C}_2 extends to a bounded map from L^p into itself for all $1 < p < \infty$.

To prove the theorem, it suffices to show that \mathcal{C}_2 maps L^2 into weak L^2 as our proof can be modified to treat all $1 < p < \infty$. We briefly indicate how to do this in the next section.

2000 *Mathematics Subject Classification*: 42B20, 42A20.

This work has been supported by an NSF grant, DMS-9706884.

Our theorem has these further antecedents. A prevalent research theme associated to E. M. Stein, F. Ricci and S. Wainger is the study of a wide class of oscillatory singular integrals. It is of interest to know that many bounds for oscillatory operators with polynomial phase can be taken to depend only on the degree of a phase function and not otherwise on the coefficients in the phase function. Among several articles we could cite, see Stein and Wainger [13, 14] and Ricci and Stein [11].

Of particular relevance to this paper is an observation of E. M. Stein [12] concerning the distribution $e(y^2)/y$. Its Fourier transform has an easily calculable form, one that permits analysis of the maximal operator formed from dilations of this distribution. We rely very much on this observation and briefly recall it in the next section.

As well, we now have a much richer understanding of Carleson's theorem as presented in papers of Fefferman, Lacey and Thiele, and Thiele [3, 7, 15], and of some related issues [8, 9, 6]. We invoke some of these elements to provide a proof of our main theorem. Stein's argument and the overview of our proof are laid out in the next section. The main inequality described in that section requires a careful analysis in time and frequency variables, that being carried out in the remaining sections of the paper.

After this note was prepared and submitted, Stein and Wainger informed me of their paper [14], which addressed the maximal operator \mathcal{C}_d above, with however the important restriction that the polynomials p over which the supremum is formed have zero coefficient for the linear term. In this instance, they supply an elegant extension of the Kolmogorov–Silvestrov method to show that this maximal operator is bounded on L^2 . Their paper also addresses natural higher dimensional analogues of \mathcal{C}_d .

It may be the case that elements of the argument of this paper can be combined with the arguments of [14] to provide a complete proof of the conjecture that \mathcal{C}_d is bounded on L^2 . We hope to return to this in a future paper.

The elegant results of K. Oskolkov [10] are of the same genre as ours.

T. Tao and J. Wright informed me of this problem. G. Mockenhaupt brought Stein's article [12] to my attention. Part of this work was completed at the Centre for Mathematics and its Applications at the Australian National University. I am indebted to all.

Notations. The Fourier transform is taken to be

$$\widehat{f}(\xi) := \int f(x)e(-2x\xi) dx.$$

The norm of an operator T from L^p into L^p is written as $\|T\|_{p \rightarrow p}$ with a corresponding notation for the weak type norm. By $A \lesssim B$ we mean that there is an absolute constant K so that $A \leq KB$; by $A \simeq B$ we mean

$A \lesssim B$ and $B \lesssim A$; $c(J)$ is the center of the interval J ; and $\chi_J(x) = (1 + \text{dist}(x, J)|J|^{-1})^{-1}$.

2. The overview of the proof. The supremum we wish to bound admits a description as a supremum over dilations, for which there are a wealth of techniques to use, and a supremum over modulations in frequency, which is the domain of Carleson's theorem. It is useful to formalize these aspects with a couple of definitions.

A distribution $K(y)$ determines two maximal functions of interest to us. They are

$$\begin{aligned} \mathcal{D}[K]f(x) &:= \sup_{a>0} \left| \int f(x-y)a^{-1}K(a^{-1}y) dy \right|, \\ \mathcal{C}[K]f(x) &:= \sup_{b \in \mathbb{R}} |\mathcal{D}[K](e(b \cdot)f)(x)| \\ &= \sup_{b \in \mathbb{R}} \sup_{a>0} \left| \int f(x-y)e(by)a^{-1}K(a^{-1}y) dy \right|. \end{aligned}$$

Thus if $K(y) = y^{-1}1_{\{0<|y|\leq 1\}}$, the $\mathcal{D}[K]$ is the maximal truncation of the Hilbert transform and $\mathcal{C}[K]$ is Carleson's maximal operator.

For the remainder of this section we set $K(y) := e(y^2/4)/y$. To prove our theorem we show that $\mathcal{C}[K]$ maps L^2 into $L^{2,\infty}$.

We recall Stein's argument [12] that $\mathcal{D}[K]$ maps L^2 into L^2 . The Fourier transform of K is a smooth odd function satisfying

$$(2.1) \quad \widehat{K}(\xi) = c_0 + e(\xi^2)\{c_1/\xi + c_2/\xi^2 + \dots\} \quad \text{as } \xi \rightarrow \infty$$

for some choice of constants $c_j, j \geq 1$. We should emphasize that the constants c_j are of the order $j!$. As is usual in these considerations, rigorous results are obtainable from a finite expansion, with an asymptotically small error term.

Indeed $\partial \widehat{K}(\xi) = \int e(x^2/4 - x\xi) dx = ce(\xi^2)$. Moreover \widehat{K} is odd as K is odd, hence

$$c^{-1}\widehat{K}(\xi) = \int_0^\xi e(y^2) dy = \int_0^\infty e(y^2) dy - \int_\xi^\infty e(y^2) dy,$$

and the assertion follows since

$$\int_\xi^\infty e(y^2) dy = e(\xi^2)\{c'_1/\xi + c'_2/\xi^2 + \dots\} \quad \text{as } \xi \rightarrow \infty.$$

With (2.1) established we can write $\widehat{K} = \widehat{H} + \sum_{j=1}^\infty 2^{-j}\widehat{m}_j$ where \widehat{H} is smooth odd and equals $c_0\text{sign}(\xi)$ for $|\xi| > 2$. The multipliers \widehat{m}_j are of the form

$$(2.2) \quad \widehat{m}_j(\xi) = \psi_j(2^{-j}\xi)e(\xi^2)$$

where ψ_j are C^∞ functions with support in $1/2 \leq |\xi| \leq 2$, and with uniformly bounded first and second derivatives. (The ψ_j will in fact converge to a fixed function in any C^s space.) These multipliers are our main concern.

Now, the term arising from H is governed by the Hilbert transform. In particular $\mathcal{D}[H]$ is bounded from L^2 into L^2 as this is the maximal truncation of the Hilbert transform. In a like manner we have the estimate $\|\mathcal{C}[H]\|_{2 \rightarrow 2, \infty} \lesssim 1$ by Carleson’s theorem [2].

Stein has shown that $\|\mathcal{D}[m_j]\|_{2 \rightarrow 2} \lesssim 2^{j/2}$, which then completes the proof of the bound on $\mathcal{D}[K]$. To prove our main theorem, we demonstrate that

$$(2.3) \quad \|\mathcal{C}[m_j]\|_{2 \rightarrow 2, \infty} \lesssim 2^{\gamma j} \quad \text{for some } 0 < \gamma < 1.$$

In fact $\gamma = 8/9$ will work. [By optimizing our argument we could establish this estimate for any $\gamma > 1/2$.]

Stein’s argument is crucial to our own and so we recall it here. To bound $\|\mathcal{D}[m_j]\|_{2 \rightarrow 2}$ it suffices to prove that

$$(2.4) \quad \left\| \sup_{1 \leq a \leq 2} \left| \int m_{j,a}(y) f(x - y) dy \right| \right\|_2 \lesssim 2^{j/2} \|f\|_2$$

where $m_{j,a}(y) := a^{-1} m_j(a^{-1}y)$. And to this end the method of TT^* is invoked. Observe that

$$(2.5) \quad \sup_{1 \leq a, b \leq 2} |m_{j,a} * \overline{m_{j,b}(y)}| \lesssim \Phi_j(y), \quad y \in \mathbb{R},$$

where

$$(2.6) \quad \Phi_j(y) = \begin{cases} 2^{j/2} |y|^{-1/2}, & |y| \leq c2^j, \\ 2^j (1 + 2^j |y|)^{-2}, & |y| \geq c2^j. \end{cases}$$

We can take $c = 16$. Note that Φ_j is non-decreasing and $\int \Phi_j dy \lesssim 2^j$, which proves (2.4).

After taking dilation into account, (2.5) amounts to the estimate

$$(2.7) \quad \sup_{0 \leq b \leq 4} \left| \int \tilde{\psi}(2^{-j}\xi) e(b\xi^2 + 2\xi y) d\xi \right| \lesssim \Phi_j(y).$$

Here $\tilde{\psi}$ is another Schwartz function with support in $1/2 \leq |\xi| \leq 2$. Set $p(\xi) = b\xi^2 + 2\xi y$. If $|y| \geq cb2^j$ observe that the derivative of p with respect to ξ exceeds $c|y|$ on the support of $\tilde{\psi}(2^{-j}\xi)$. Thus repeated integration by parts will prove the estimate. If $|y| \leq cb2^j$ we can use the van der Corput second derivative test. It provides the estimate of the integral as $\lesssim b^{-1/2} \lesssim 2^{j/2} |y|^{-1/2}$. Thus the inequality holds.

The remainder of the paper is devoted to a proof of (2.3).

To do so we use the time frequency analysis of Lacey–Thiele [7] with some further ideas drawn from Fefferman [3] and Thiele [15]. A central conceptual problem arises from the fact that \widehat{m}_j is supported in an interval of length 2^j but m_j has (approximate) spatial support in an interval of the same length.

That is, classical Fourier uncertainty is not observed. Treating this issue is probably the main novelty of this paper.

For our subsequent use observe these points. First, in the definition of (2.2) we can assume that ψ is supported in $1 - 1/40 \leq \xi \leq 1 + 1/40$ (as opposed to $1/2 \leq |\xi| \leq 2$). As a matter of notational convenience, we shall assume that ψ is independent of j .

Second, (2.7) implies that $\int |m_j| dy \lesssim 2^j$, hence $\|\mathcal{D}[m_j]\|_{\infty \rightarrow \infty} \lesssim 2^j$, so that by interpolation

$$(2.8) \quad \|\mathcal{D}[m_j]\|_{p \rightarrow p} \lesssim 2^{j(1-1/p)}, \quad 2 < p < \infty.$$

Thus $\mathcal{D}[K]$ maps L^p into itself for $2 < p < \infty$.

Indeed, this estimate holds for $1 < p < \infty$. In fact, we have the companion estimate to the L^∞ estimate above, $\|\mathcal{D}[m_j]\|_{1 \rightarrow 1} \lesssim 2^j$. These estimates require no cancellation, and so hold for $\mathcal{C}[m_j]$ as well. Thus to prove our main theorem and in light of the extension of Hunt of Carleson's theorem, it suffices to provide the bound we have claimed for $\mathcal{C}[m_j]$ on L^2 .

Third, there is a sharper form of Stein's observation. Namely the operator

$$(2.9) \quad D_j f(x) := \sup_{1 \leq a \leq 2} \sup_{|N| \leq 2^j} \left| \int a^{-1} m_j(a^{-1}(x-y)) e(Ny) f(y) dy \right|$$

maps L^2 into L^2 with norm bounded by $\lesssim 2^{j/2}$. Employing the same arguments as above, this amounts to the estimate

$$\sup_{0 \leq b \leq 4} \sup_{|N| \leq 2^{j+2}} \left| \int \tilde{\psi}(2^{-j}\xi) e(b\xi^2 + (N+y)\xi) dy \right| \lesssim \Gamma_j(y)$$

where

$$\Gamma_j(y) = \begin{cases} 2^{j/2} |y|^{-1/2}, & |y| \leq c2^j, \\ 2^j (1 + 2^j |y|)^{-n}, & |y| \geq c2^j. \end{cases}$$

In this definition, n is an arbitrary positive integer and $c = 32$. Details are a modification of the earlier argument. In fact we have $\|D_j\|_p \lesssim 2^{j(1-1/p)}$ for $2 \leq p < \infty$. We shall have recourse to this below.

Fourth, in proving the estimate (2.4) we follow the approach of Kolmogorov and Silvestrov, as Fefferman [3] has demonstrated that this is a powerful technique in issues related to Carleson's theorem. We show that there is a $0 < \gamma < 1$ so that for all j , measurable functions $N : \mathbb{R} \rightarrow \mathbb{R}$, $\ell : \mathbb{R} \rightarrow \mathbb{Z}$ and $a : \mathbb{R} \rightarrow [1, 1 + 1/40]$,

$$(2.10) \quad \left\| \int m_{j,a(x)2^{\ell(x)}}(x-y) f(y) e(N(x)y) dy \right\|_{2,\infty} \lesssim 2^{\gamma j} \|f\|_2.$$

We will do this with $\gamma = 8/9$. This inequality is sufficient for our purposes.

3. The discrete operator. Let \mathbf{D} be a collection of dyadic intervals in the real line. Let \mathbf{P}_{fat} be the set of rectangles $s = I_s \times \omega_s \in \mathbf{D} \times \mathbf{D}$ which have area $|I_s| |\omega_s| = 2^{2j}$. We call these *fat tiles* and we generically

write s, s', s'' for fat tiles. Let ω_{s1} (resp. ω_{s2}) be the left (resp. right) half of ω_s . This definition is chosen in accordance with the frequency and spatial localizations of the kernel m_j , its dilates and modulations.

Let \mathbf{P}_{thin} be the set of rectangles $\sigma \in \mathbf{D} \times \mathbf{D}$ of area 1. We call these *thin tiles* and we generically write $\sigma, \sigma', \sigma''$ for thin tiles. Set $\text{thin}(s) := \{\sigma \in \mathbf{P}_{\text{thin}} : I_\sigma = I_s, \omega_\sigma \subset \frac{3}{4}\omega_{s2}\}$. For $\sigma \in \text{thin}(s)$, set $\omega_{\sigma j} := \omega_{s j}$ for $j = 1, 2$. See Figure 1. [Actually, “thin tiles” obey classical Fourier uncertainty and so are thin only in contrast to fat tiles.]

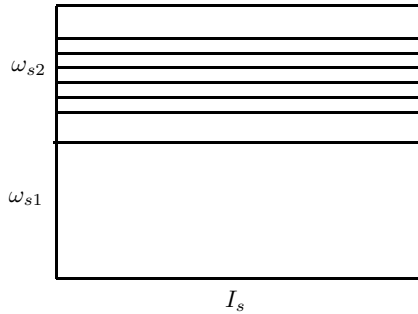


Fig. 1. A fat tile s and thin tiles $\sigma \in \text{thin}(s)$

Fix a Schwartz function φ with $1_{[-1/80, 1/80]} \leq \widehat{\varphi} \leq 1_{[-1/40, 1/40]}$. For a rectangle $\sigma = I_\sigma \times \omega_\sigma$ of area 1 (not necessarily a thin tile) define

$$\varphi_\sigma(x) = \frac{e(2c(\omega_\sigma)x)}{\sqrt{|I_\sigma|}} \varphi\left(\frac{x - c(I_\sigma)}{|I_\sigma|}\right).$$

In this display and throughout, $c(J)$ is the center of the interval J .

Fix the data $j \geq 1, f \in L^2$ of norm one, functions N, ℓ and a as in (2.10). For $s \in \mathbf{P}_{\text{fat}}, \sigma \in \text{thin}(s)$ and integer l with $2^{-l} = |\omega_{s2}|$, define

$$E(s) = E(\sigma) := \{x \in \mathbb{R} : N(x) \in \omega_{s1}, \ell(x) = l\},$$

$$\phi_\sigma(x) = 1_{E(\sigma)}(x) \int \varphi_\sigma(y) e(-2N(x)y) a(x)^{-1} 2^{-l-j} m_j(a(x)^{-1} 2^{-l-j}(x-y)) dy,$$

$$M_j f(x) = \sum_{s \in \mathbf{P}_{\text{fat}}} \sum_{\sigma \in \text{thin}(s)} \langle f, \varphi_\sigma \rangle \phi_\sigma(x).$$

A principal motivation for these definitions is the proof of Lemma 4.3 below. At this point we simply observe that the support of the integral in the definition of ϕ_σ is in $E(\sigma)$. \widehat{m}_j is supported in a small neighborhood of 2^j so that the second function in the last integral has frequency support in a small interval around 2^{-l} . Moreover, $\widehat{\varphi}_\sigma$ is supported in a small interval around $c(\omega_\sigma)$ with $\omega_\sigma \subset \frac{3}{4}\omega_{s2}$. So $N(x)$ must be in ω_{s1} in order for the integral to be non-zero.

We claim that the following inequality is sufficient for (2.10):

$$(3.1) \quad \|M_j f\|_{2,\infty} \lesssim 2^{\gamma j}, \quad \gamma = 8/9.$$

In the proof of this inequality, we only consider sums over finite subsets $S_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$. We fix data $f \in L^2$ of norm one and the functions N, ℓ and a . Let M_j be the sum restricted to this new smaller class of tiles. Then, by dilation invariance, (3.1) is implied by this inequality:

$$(3.2) \quad |\{M_j f > 1\}| \lesssim 2^{2\gamma j}, \quad \gamma = 8/9,$$

the inequality holding for all functions f of L^2 norm one.

Proof of sufficiency of (3.1). A convexity argument can be used to show that (3.1) implies the inequality (2.3). Indeed arguments like this are a standard part of the theory related to Carleson's theorem and multilinear operators, having been used many times in related papers, for instance [3, 7, 8]. Accordingly, we will be somewhat brief.

Let us give the convexity argument in an elemental form. For our subsequent use, let us define translation and modulation operators by

$$\text{Tr}_t f = f(x - t) \quad \text{and} \quad \text{Mod}_t f(x) = e^{-ixt} f(x) \quad \text{for } t \in \mathbb{R}.$$

Observe that the sum

$$\sum_{n \in \mathbb{Z}} \langle f, \text{Tr}_n \varphi \rangle \text{Tr}_n \varphi$$

could be written as a sum over tiles. More importantly,

$$\int_0^1 \sum_{n \in \mathbb{Z}} \langle f, \text{Tr}_{n+t} \varphi \rangle \text{Tr}_{n+t} \varphi dt = \int_{-\infty}^{\infty} \langle f, \text{Tr}_t \varphi \rangle \text{Tr}_t \varphi dt = \psi * f$$

where $\psi(x) = \int \overline{\varphi(-y)} \varphi(x-y) dy$. Recall that we specified φ to be a Schwartz function with $1_{[-1/80, 1/80]} \leq \widehat{\varphi} \leq 1_{[-1/40, 1/40]}$, so that ψ satisfies a similar set of inequalities.

Elaborating on this theme, observe that the sum

$$(3.3) \quad \sum_{m, n \in \mathbb{Z}} \langle f, \text{Mod}_m \text{Tr}_n \varphi \rangle \text{Mod}_m \text{Tr}_n \varphi$$

could be written as a sum over tiles. Define

$$\begin{aligned} Af &= \int_0^1 \int_0^1 \sum_{m, n \in \mathbb{Z}} \langle f, \text{Mod}_{m+\tau} \text{Tr}_{n+t} \varphi \rangle \text{Mod}_{m+\tau} \text{Tr}_{n+t} \varphi d\tau dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle f, \text{Mod}_\tau \text{Tr}_t \varphi \rangle \text{Mod}_\tau \text{Tr}_t \varphi d\tau dt. \end{aligned}$$

This is a multiple of the identity, as is easy to see. Af is a bounded linear operator on L^2 that commutes with translations and modulations. It is furthermore positive definite, whence a multiple of the identity.

By periodicity, Af is also equal to

$$(3.4) \quad Af = \lim_{\tau \rightarrow \infty} \lim_{t \rightarrow \infty} (\tau t)^{-1} \iint_{00}^{t\tau} \sum_{m,n \in \mathbb{Z}} \langle f, \text{Mod}_{m+\tau'} \text{Tr}_{n+t'} \varphi \rangle \text{Mod}_{m+\tau'} \text{Tr}_{n+t'} \varphi \, d\tau' \, dt'.$$

This concludes our general remarks on resolutions of the identity and the use of convexity.

Let us turn to the operator M_j . Define, for an integer l ,

$$P_{j,l}f = \sum_{\substack{s \in \mathbf{P}_{\text{fat}} \\ |\omega_{s2}| = 2^{-l}}} \sum_{\sigma \in \text{thin}(s)} \langle f, \varphi_\sigma \rangle \varphi_\sigma,$$

and observe that this sum is similar to (3.3). We may average these operators over modulations and translations to obtain a multiple of the identity. This can be done in a way that is independent of $l \in \mathbb{Z}$ and of $j \geq 1$. We shall return to this point momentarily.

To make the connection with our operator M_j more directly, observe that with the notation used in the definition of M_j ,

$$M_j f(x) = (\text{Mod}_{2N(x)} \text{Dil}_{2^{\ell(x)-j} a(x)} m_j) * P_{j,\ell(x)} f(x)$$

where $\text{Dil}_\delta g(x) = \delta^{-1} g(x\delta^{-1})$.

Thus the main point is that we can recover the identity operator from $P_{j,l}$ in a way that is independent of l and j and does not affect the assumed inequality (3.1).

But certainly translation and modulation do not affect the distributional inequality. Moreover, we can obtain the identity operator from the $P_{j,l}$ in this way. Recall that the tiles depend upon choices of dyadic grids \mathbf{D} and \mathbf{D}' . A translation of \mathbf{D} (resp. \mathbf{D}') corresponds to an application of Tr_t (resp. Mod_τ) to the functions φ_σ . Thus the assumed inequality applies to any M_j obtained from translations of either grid. Finally, the periodicity property (3.4) shows that the identity operator can be obtained in a way that is independent of l . This completes the proof. ■

4. Trees and size. The principal definitions and lemmas are stated in this section. We show how they prove (3.2), and prove the lemmas in the following section. We begin with requisite definitions.

A subset $S_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$ has scales separated by factor J if for all $s, s' \in S_{\text{fat}}$, the inequality $|I_s| < |I_{s'}|$ implies that $2^J |I_s| < |I_{s'}|$. We will apply the same terminology to collections of thin tiles.

For $s, s' \in \mathbf{P}_{\text{fat}}$ say that $s < s'$ iff $I_s \subset I_{s'}$ and $\omega_s \supset \omega_{s'}$. Say that $T_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$ is a tree if there is an $I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}} \in \mathbf{P}_{\text{fat}}$ with $s < I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}}$ for all $s \in T_{\text{fat}}$. The top of a tree need not be unique.

A subset $T_{\text{thin}} \subset \mathbf{P}_{\text{thin}}$ is a *tree* if it is a subset of $\text{thin}(T_{\text{fat}})$ for some tree $T_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$. We denote the top of the tree by $I_{T_{\text{thin}}} \times \omega_{T_{\text{thin}}} := I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}}$. A tree T_{thin} is a *1-tree* (resp. *2-tree*) if for all $\sigma \neq \sigma' \in T_{\text{thin}}$ either $\omega_\sigma = \omega_{\sigma'}$ or $\omega_\sigma \cap \omega_{\sigma'} = \emptyset$ (resp. $\omega_\sigma \cap \omega_{\sigma'} \neq \emptyset$).

Note that if T_{thin} has scales separated by factor $2j$, then T_{thin} can be uniquely decomposed as a union of a 1-tree and a 2-tree. Also note that the definition of 1-trees and 2-trees is formulated in terms of ω_σ and not $\omega_{\sigma 2}$.

For $S_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$, define the *size* of S_{fat} to be

$$\text{size}(S_{\text{fat}}) := \sup_{T_{\text{thin}} \subset \text{thin}(S_{\text{fat}})} \left[|I_{T_{\text{thin}}}|^{-1} \sum_{\sigma \in T_{\text{thin}}} |\langle f, \varphi_\sigma \rangle|^2 \right]^{1/2}$$

where the supremum is taken over all 1-trees

$$T_{\text{thin}} \subset \text{thin}(S_{\text{fat}}) := \bigcup_{s \in S_{\text{fat}}} \text{thin}(s).$$

The central lemma concerning size is

4.1. LEMMA. *A finite collection $S_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$ is a union of collections $S_{\text{fat}}(n)$, $n \in \mathbb{Z}$, for which $\text{size}(S_{\text{fat}}(n)) \leq j2^n$ and*

$$(4.2) \quad \sum_{s \in S_{\text{fat}}(n)^*} |I_s| \lesssim 2^{-2n},$$

where $S_{\text{fat}}(n)^*$ consists of the maximal $s \in S_{\text{fat}}(n)$.

Observe that j (that is a measure of how fat the tiles are) enters into this lemma, albeit in a weak fashion.

For $S_{\text{thin}} \subset \mathbf{P}_{\text{thin}}$ set

$$M_j^{S_{\text{thin}}} = \sum_{\sigma \in S_{\text{thin}}} \langle f, \varphi_\sigma \rangle \phi_\sigma.$$

If $S_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$, define $M_j^{S_{\text{fat}}}$ to be $M_j^{\text{thin}(S_{\text{fat}})}$. Concerning trees, our central lemma is

4.3. LEMMA. *For all trees T_{fat} ,*

$$(4.4) \quad \|M_j^{T_{\text{fat}}}\|_p \lesssim \text{size}(T_{\text{fat}}) j 2^{j(3-5/p)} |I_{T_{\text{fat}}}|^{1/p}, \quad 2 < p < \infty.$$

$$(4.5) \quad |M_j^{T_{\text{fat}}}(x)| \lesssim 2^{2j} \text{size}(T_{\text{fat}}) \chi_{I_{T_{\text{fat}}}}(x)^m, \quad x \notin 2I_{T_{\text{fat}}}, \quad m \geq 1,$$

where $\chi_J(x) = (1 + \text{dist}(x, J)|J|^{-1})^{-1}$.

Notice that the first estimate should be compared to Stein's estimate for $\mathcal{D}[M_j]$, and is only slightly worse than that estimate if $p = 2$. That the (large) factor of 2^{2j} enters into the second estimate is completely harmless.

Set $\varepsilon = (200)^{-1}$, $p = 9/4$, $\mu = 7/9$ and $\gamma = 8/9$. We are to prove the distributional estimate (3.2).

For $n > -\gamma j$, we in essence rely upon the fact that $M_j^{S_{\text{fat}}(n)}$ is supported on a set of small measure. To make this precise, let $E_n = \bigcup_{s \in S_{\text{fat}}(n)^*} 2^{\varepsilon j} I_s$ and $F_0 = \bigcup_{n > -\gamma j} E_n$. The latter set has measure

$$|F_0| \leq \sum_{n > -\mu j} |E_n| \lesssim \sum_{n > -\mu j} 2^{\varepsilon j - 2n} \lesssim 2^{2\gamma j}.$$

We do not need to estimate M_j on this set.

Using (4.5), with m in that inequality being $m = 400$, we see that

$$\|M_j^{S_{\text{fat}}(n)}\|_{L^1(F_0^c)} \lesssim 2^{-100j} \sum_{s \in S_{\text{fat}}(n)^*} |I_s| \lesssim 2^{-100j - 2n}.$$

Bringing these estimates together, we see that for the collection $\tilde{S}_{\text{fat}} = \bigcup_{n > -\mu j} S_{\text{fat}}(n)$, we have $|\{M_j^{\tilde{S}_{\text{fat}}} > 1\}| \lesssim 2^{2\gamma j}$, as is required in (3.2).

For $n \leq -\gamma j$, we need a more involved argument. We encode some of the necessary combinatorics into this lemma:

4.6. LEMMA. *For $n \leq -\gamma j$ there is a set $E_n \subset \mathbb{R}$ with $|E_n| \lesssim 2^n$ so that the collection $\tilde{S}_{\text{fat}}(n) = \{s \in S_{\text{fat}}(n) : I_s \not\subset E_n\}$ is a union of collections $U_{\text{fat}}(n, k)$, $1 \leq k \leq -500n$, which satisfies these properties. For each $1 \leq k \leq -500n$,*

(i) $U_{\text{fat}}(n, k)$ is uniquely decomposable into maximal disjoint trees $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)$.

(ii) $\|\sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)} 1_{I_{T_{\text{fat}}}}\|_{\infty} \lesssim 2^{-10n}$.

(iii) $\{2^{-\varepsilon n} I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}} : T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)\}$ are pairwise disjoint rectangles.

(iv) For all $s \in U_{\text{fat}}(n, k)$,

$$I_s \not\subset \bigcup_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)} \{x : \text{dist}(x, \partial I_{T_{\text{fat}}}) < 2^{10n} |I_{T_{\text{fat}}}| \}.$$

(v) Either $T_{\text{fat}} = \{I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}}\}$ for all $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)$ or $I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}} \notin T_{\text{fat}}$ for all $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)$.

(vi) The collection of tiles $U_{\text{fat}}(n, k) \cup \bigcup \{I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}} : T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)\}$ has scales separated by a factor of $200n$.

We do not estimate our operator M_j on the set $F_1 = \bigcup_{n \leq -\mu j} E_n$. As this set has measure $|F_1| \lesssim 2^{-\mu j} \lesssim 2^{-2\gamma j}$, there is no harm in doing this.

Off this set, our lemma permits the following construction. For all $n \leq -\mu j$, $0 \leq k \leq -10,000n$ and $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)$ there are functions $N^{T_{\text{fat}}}$ for which

$$(4.7) \quad |N^{T_{\text{fat}}}(x) - M_j^{T_{\text{fat}}}(x)| \lesssim 2^{10n} \chi_{I_{T_{\text{fat}}}}(x)^{100}, \quad x \in \mathbb{R},$$

(4.8) the functions $N^{T_{\text{fat}}}$ are disjointly supported in $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n, k)$.

But then we can estimate, by (4.4),

$$\begin{aligned} \left\| \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n,k)} N^{T_{\text{fat}}} \right\|_p^p &= \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n,k)} \|N^{T_{\text{fat}}}\|_p^p \\ &\lesssim j^p 2^{np+(3-5/p)pj} \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}(n,k)} |I_{T_{\text{fat}}}| \\ &\lesssim j^5 2^{7j/4+n/4}. \end{aligned}$$

Thus certainly

$$|\{M_j^{U_{\text{fat}}(n,k)} \geq 2^{n/18}\}| \lesssim 2^{7j/4+n/8}.$$

This is summable over $n \leq -\mu j$ and $0 \leq k \leq -500n$ and so completes our proof of (3.2).

[This interplay between L^2 and L^p estimates is due to C. Thiele [15] and contrasts with the argument of Lacey and Thiele [7]. The latter paper uses two notions of “energy” (the current “size”) and “mass”, which are in some sense dual to each other. The notion of “mass” cannot play the same role in our paper: “Mass” can be exploited through devices linked to the Hardy–Littlewood maximal function, but our kernels bear no close connection to that maximal function.]

The construction relies on an argument from [6]. Fix n, k , set $U_{\text{fat}} := U_{\text{fat}}(n, k)$ and $\mathbf{T}_{\text{fat}} := \mathbf{T}_{\text{fat}}(n, k)$. For each $s \in U_{\text{fat}}$ we construct a set G_s as follows. Recall (v) from Lemma 4.6. If each $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}$ consists only of a top we set $G_s = 2^{-\varepsilon n} I_s$ where s is the top of the tree and $N^{T_{\text{fat}}} = 1_{G_s} \sum_{\sigma \in \text{thin}(s)} \langle f, \varphi_\sigma \rangle \phi_\sigma$. Then (4.8) follows from (iii) and (4.7) follows from (4.5).

We thus assume that no tree $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}$ contains its top. We then make the following definitions for $s \in T_{\text{fat}}$:

$$\begin{aligned} G_s &:= I_{T_{\text{fat}}} - \bigcup_{T'_{\text{fat}} \in \mathbf{T}_{\text{fat}}(s)} I_{T'_{\text{fat}}}, \\ \mathbf{T}_{\text{fat}}(s) &:= \{T'_{\text{fat}} \in \mathbf{T}_{\text{fat}} - \{T_{\text{fat}}\} : \omega_{s1} \supset \omega_{T'_{\text{fat}}}, I_{T'_{\text{fat}}} \subset I_{T_{\text{fat}}}\}, \\ N^{T_{\text{fat}}} &:= \sum_{s \in T_{\text{fat}}} 1_{G_s} \sum_{\sigma \in \text{thin}(s)} \langle f, \varphi_\sigma \rangle \phi_\sigma. \end{aligned}$$

We verify (4.8). Since the support of ϕ_σ is in $\{x : N(x) \in \omega_{s1}\}$, where $\sigma \in \text{thin}(s)$, (4.8) is a consequence of the observation that if $G_s \times \omega_{s1} \cap G_{s'} \times \omega_{s'1} \neq \emptyset$ then s and s' are in the same tree. Indeed, write $s \in T_{\text{fat}}$ and $s' \in T'_{\text{fat}}$ and assume say $\omega_{s'1} \subset \omega_{s1}$. If $\omega_{s'1} = \omega_{s1}$ and the two trees are distinct then $I_{T_{\text{fat}}}$ and $I_{T'_{\text{fat}}}$ are disjoint by (i). Assume $\omega_{s'1} \subsetneq \omega_{s1}$ and $G_s \cap G_{s'} \neq \emptyset$. Then $\omega_{T_{\text{fat}}}, \omega_{T'_{\text{fat}}} \subset \omega_{s1}$. See Figure 2. The tops $I_{T_{\text{fat}}}$ and $I_{T'_{\text{fat}}}$ must intersect. Assuming $I_{T_{\text{fat}}} \subset I_{T'_{\text{fat}}}$ implies $s < I'_{\text{fat}} \times \omega'_{\text{fat}}$. But then

condition (i) forces $T_{\text{fat}} = T'_{\text{fat}}$. Thus we must have $I_{T'_{\text{fat}}} \subset I_{T_{\text{fat}}}$, which by definition means that $I_{T'_{\text{fat}}} \cap G_s = \emptyset$, so by definition $G_s \cap G_{s'} = \emptyset$. This is a contradiction and so proves (4.8).

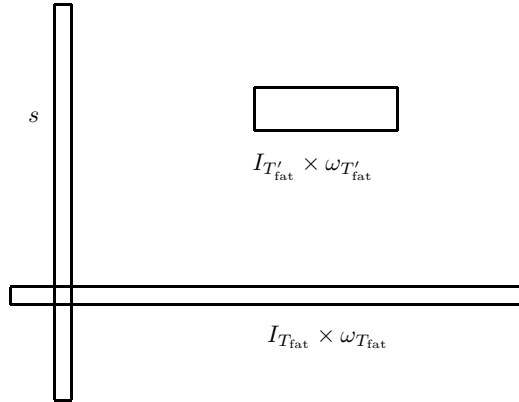


Fig. 2. The top of the tree T'_{fat} does not intersect the set G_s since the intervals $I_{T_{\text{fat}}}$ and $I_{T'_{\text{fat}}}$ intersect.

We verify (4.7). In the case of $x \notin I_{T_{\text{fat}}}$ this follows from the fact that the top is not in the tree, (4.5) and conditions (iv) and (vi) of Lemma 4.6. We do not comment further.

For $x \in I_{T_{\text{fat}}}$ we in fact have $N^{T_{\text{fat}}}(x) = M_j^{T_{\text{fat}}}(x)$ unless $x \in I_{T'_{\text{fat}}}$ and $T'_{\text{fat}} \in T_{\text{thin}}(s)$ for some $s \in T_{\text{fat}}$. Indeed, with T_{fat} fixed we can assume that $I_{T'_{\text{fat}}} \subset I_{T_{\text{fat}}}$ for all $T'_{\text{fat}} \in \mathbf{T}_{\text{fat}}$. Then we shall just reverse the order of summation below.

$$|N^{T_{\text{fat}}}(x) - M_j^{T_{\text{fat}}}(x)| \leq \sum_{T'_{\text{fat}} \in \mathbf{T}_{\text{fat}} - \{T_{\text{fat}}\}} \sum_{\sigma \in \text{thin}(\mathbf{T}_{\text{fat}}(T'_{\text{fat}}))} |\langle f, \varphi_\sigma \rangle \phi_\sigma(x)|,$$

where $\mathbf{T}_{\text{fat}}(T'_{\text{fat}}) := \{s \in T_{\text{fat}} : T'_{\text{fat}} \in \mathbf{T}_{\text{fat}}(s)\}$ and $\mathbf{T}_{\text{fat}}(s)$ was used to define G_s . But again conditions (iv) and (vi) imply that

$$\sum_{\sigma \in \text{thin}(\mathbf{T}_{\text{fat}}(T'_{\text{fat}}))} |\langle f, \varphi_\sigma \rangle \phi_\sigma(x)| \lesssim 2^{200n} 1_{T'_{\text{fat}}}(x)$$

and (4.7) follows from condition (ii).

Our proof of (3.2) is complete modulo the proofs of the lemmas, which are taken up in the next section.

5. Proofs of the lemmas

Proof of Lemma 4.1. The argument is a variant of one in [5, 1] and has been used several times since. We give the details, although only small

changes are needed to account for the disparity between fat and thin tiles. The most expedient treatment requires a new definition of a tree.

Fix a choice of integer $0 \leq k < 200j$. For a 1-tree T_{thin} call a subset $T_{\text{thin}}^\ell \subset T_{\text{thin}}$ a *left-tree* (resp. *right-tree*) if there is a $\xi_{T_{\text{thin}}} \in \omega_{T_{\text{thin}}}$ with $\xi_{T_{\text{thin}}}$ to the left (resp. right) of every ω_s , $s \in T_{\text{thin}}^\ell$. In addition require that for all $s \in T_{\text{thin}}^\ell \cup \{I_{T_{\text{thin}}^\ell} \times \omega_{T_{\text{thin}}^\ell}\}$, $\log_2 |I_s| \in k + 200j\mathbb{Z}$. Define the *left size* of S_{fat} , or $\ell\text{-size}(S_{\text{fat}})$, as

$$\ell\text{-size}(S_{\text{fat}}) = \sup \left\{ \left[|I_{T_{\text{fat}}} |^{-1} \sum_{s \in T_{\text{fat}}^\ell} |\langle f, \varphi_\sigma \rangle|^2 \right]^{1/2} \right\}$$

where the supremum is over all left-trees T_{fat}^ℓ with $T_{\text{fat}} \subset S_{\text{fat}}$.

We prove this statement. For any finite $S_{\text{fat}} \subset \mathbf{P}_{\text{fat}}$ set $\varepsilon = \ell\text{-size}(S_{\text{fat}})$. Then $S_{\text{fat}} = S_{\text{lo}} \cup S_{\text{hi}}$ with $\ell\text{-size}(S_{\text{lo}}) \leq \varepsilon/4$ and S_{hi} is a union of trees $T_{\text{fat}} \subset \mathbf{T}_{\text{fat}}$ with

$$(5.1) \quad \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}} |I_{T_{\text{fat}}} | \lesssim \varepsilon^{-2}.$$

An inductive application of this statement proves Lemma 4.1 with $\text{size}(S_{\text{fat}})$ replaced by $\ell\text{-size}(S_{\text{fat}})$, with however a small improvement: The factor j does not enter into this statement of the lemma in terms of $\ell\text{-size}$. The same statement is true for right size. Letting k vary from 0 to $200j$ proves the lemma as stated.

The construction of S_{hi} and \mathbf{T}_{fat} is inductive. The construction also associates to each $T_{\text{fat}} \in \mathbf{T}_{\text{fat}}$ a particular left-tree T_{thin}^ℓ which is used to prove (5.1). Initially set $S_{\text{fat}}^{\text{stock}} := S_{\text{fat}}$. Select a tree $T_{\text{fat}} \subset S_{\text{fat}}^{\text{stock}}$ so that

- (a) T_{fat} contains a left-tree T_{thin}^ℓ with

$$\sum_{\sigma \in T_{\text{thin}}^\ell} |\langle f, \varphi_\sigma \rangle|^2 \geq (4\varepsilon)^{-2} |I_{T_{\text{fat}}} |.$$

- (b) $I_{T_{\text{fat}}}$ is maximal among trees satisfying condition (a) and T_{fat} is the maximal tree in $S_{\text{fat}}^{\text{stock}}$ with that top.

- (c) $\xi_{T_{\text{thin}}^\ell}$ is rightmost among trees satisfying (a) and (b).

Then add T_{fat} to \mathbf{T}_{fat} , set $S_{\text{fat}}^{\text{stock}} := S_{\text{fat}}^{\text{stock}} - T_{\text{fat}}$. Repeat this procedure until there is no tree satisfying (a). Then set $S_{\text{lo}} := S_{\text{fat}}^{\text{stock}}$. By definition, $\ell\text{-size}(S_{\text{lo}}) \leq \varepsilon/4$.

The left-trees we have constructed satisfy this disjointness property: For $T_{\text{fat}} \neq T'_{\text{fat}} \in \mathbf{T}_{\text{fat}}$ and $\sigma \in T_{\text{thin}}^\ell$ and $\sigma' \in T'_{\text{thin}}^{\ell'}$,

$$(5.2) \quad \text{if } \omega_\sigma \not\subset \omega_{\sigma'} \text{ then } I_{T_{\text{fat}}} \cap I_{\sigma'} = \emptyset.$$

Indeed, $\omega_{T_{\text{fat}}} \subset \omega_{\sigma'}$ so that $\xi_{T'_{\text{fat}}} < \xi_{T_{\text{fat}}}$. Thus the tree T_{fat} was constructed before T'_{fat} . But if $I_{T_{\text{fat}}} \cap I_{\sigma'} \neq \emptyset$ we see that $s' < I_{T_{\text{fat}}} \times \omega_{T_{\text{fat}}}$ where $\sigma \in$

thin(s'). Hence $s' \in T_{\text{fat}}$, which is a contradiction. See Figure 3.

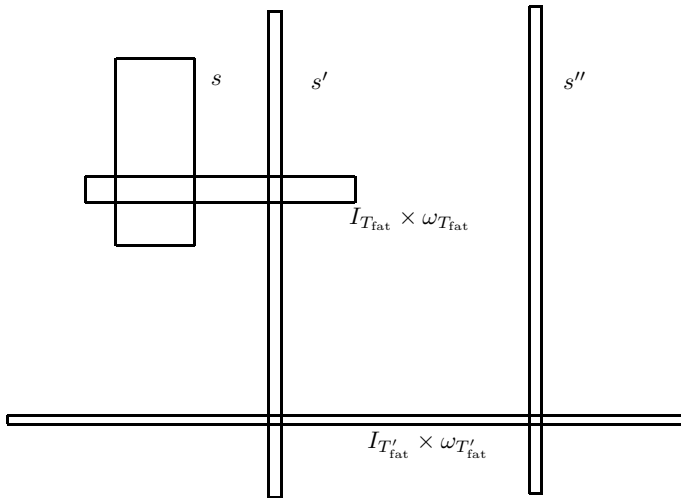


Fig. 3. By the manner in which the trees are constructed, the tree T_{fat} was constructed before T'_{fat} . Hence the tile s' must be in T_{fat} . But the tile s'' is a member T'_{fat} .

Let T^{red} be those $\sigma \in T_{\text{thin}}^\ell$ for which if $|I_\sigma| < |I_{T^{\text{red}}}|$ then $\text{dist}(I_\sigma, \partial I_{T^{\text{red}}}) \geq \frac{1}{32}|I_{T^{\text{red}}}|$. [“red” is for “reduced.” Note that the top is permitted to be in T^{red} , and if $|I_\sigma| < |I_{T^{\text{red}}}|$ then $|I_\sigma|$ is in fact much smaller than $|I_{T^{\text{red}}}|$.] As $\ell\text{-size}(S_{\text{fat}}) \leq \varepsilon$, it follows that

$$\sum_{\sigma \in T^{\text{red}}} |\langle f, \varphi_\sigma \rangle|^2 \geq \frac{\varepsilon^2}{32} |I_{T^{\text{red}}}|.$$

Set $S_{\text{thin}} = \bigcup_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}} T^{\text{red}}$ and

$$B := \left\| \sum_{\sigma \in S_{\text{thin}}} \langle f, \varphi_\sigma \rangle \varphi_\sigma \right\|_2.$$

Observe that by Cauchy–Schwarz and $\|f\|_2 = 1$,

$$(5.3) \quad \frac{\varepsilon^2}{32} \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}} |I_{T_{\text{fat}}}| \leq \sum_{\sigma \in S_{\text{thin}}} |\langle f, \varphi_\sigma \rangle|^2 = \left\langle f, \sum_{\sigma \in S_{\text{thin}}} \langle f, \varphi_\sigma \rangle \varphi_\sigma \right\rangle \leq B.$$

To conclude (5.1) we show that

$$(5.4) \quad B^2 \lesssim \varepsilon^2 \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}} |I_{T_{\text{fat}}}|.$$

By expanding the L^2 norm, $B^2 \leq 2(B_1^2 + B_2^2)$ where we define

$$B_j^2 := \sum_{\sigma \in S_{\text{thin}}} \langle f, \varphi_\sigma \rangle \sum_{\sigma' \in S_{\text{thin}}^j(\sigma)} \langle \varphi_\sigma, \varphi_{\sigma'} \rangle \langle \varphi_{\sigma'}, f \rangle$$

$$S_{\text{thin}}^1(\sigma) := \{\sigma' \in S_{\text{thin}} : \omega_\sigma = \omega_{\sigma'}\}, \quad S_{\text{thin}}^2(\sigma) := \{\sigma' \in S_{\text{thin}} : \omega_\sigma \subsetneq \omega_{\sigma'}\}.$$

Note that if $\omega_\sigma \subset \omega_{\sigma'}$ we have

$$(5.5) \quad |\langle \varphi_\sigma, \varphi_{\sigma'} \rangle| \lesssim \sqrt{\frac{|I_{\sigma'}|}{|I_\sigma|}} \chi_{I_\sigma}(c(I_{\sigma'}))^{100}.$$

To bound B_1 fix a dyadic interval ω . This last estimate and Cauchy–Schwarz show that

$$\left| \sum_{\substack{\sigma \in S_{\text{thin}} \\ \omega_\sigma = \omega}} \langle f, \varphi_\sigma \rangle \sum_{\sigma' \in S_{\text{thin}}^1(s)} \langle \varphi_\sigma, \varphi_{\sigma'} \rangle \langle \varphi_{\sigma'}, f \rangle \right| \lesssim \sum_{\substack{\sigma \in S_{\text{thin}} \\ \omega_\sigma = \omega}} |\langle f, \varphi_\sigma \rangle|^2.$$

Hence by (5.3) and summing over ω ,

$$B_1^2 \lesssim \sum_{\sigma \in S_{\text{thin}}} |\langle f, \varphi_\sigma \rangle|^2 \lesssim \varepsilon^2 \sum_{T_{\text{fat}} \in \mathbf{T}_{\text{fat}}} |I_{T_{\text{fat}}}|$$

as $\ell\text{-size}(S_{\text{thin}}) = \varepsilon$. This is the first step in establishing (5.4).

To control B_2^2 we must use the disjointness property (5.2). Fix a tree T^{red} and consider $\sigma \in T^{\text{red}}$. Then the intervals $\{I_{\sigma'} : \sigma' \in S_{\text{thin}}(s)\}$ are pairwise disjoint and contained in $(I_{T_{\text{fat}}})^c$. To see this note that for all $\sigma', \sigma'' \in S_{\text{thin}}(s)$ we have $\omega_\sigma \subset \omega_{\sigma'} \cap \omega_{\sigma''}$. So (5.2) implies $I_{\sigma'} \cap I_{\sigma''} = \emptyset$. Then we can estimate

$$\begin{aligned} \sum_{\sigma' \in S_{\text{thin}}(\sigma)} |\langle f, \varphi_\sigma \rangle \langle \varphi_\sigma, \varphi_{\sigma'} \rangle \langle \varphi_{\sigma'}, f \rangle| &\lesssim \varepsilon^2 \sum_{\sigma' \in S_{\text{thin}}(s)} \chi_{I_\sigma}(c(I_{\sigma'}))^{100} |I_{\sigma'}| \\ &\lesssim \varepsilon^2 \int_{(I_{T_{\text{fat}}})^c} \chi_{I_\sigma}(x)^{90} dx \\ &\lesssim \varepsilon^2 \left(\frac{|I_\sigma|}{|I_{T^{\text{red}}}|} \right)^{10} |I_\sigma|. \end{aligned}$$

Here, we have in addition relied upon the estimate

$$|\langle f, \varphi_\sigma \rangle| \leq \varepsilon \sqrt{|I_\sigma|}.$$

Finally, we have the estimate

$$\sum_{\sigma \in T_{\text{thin}}^\ell} \int_{(I_{T_{\text{fat}}})^c} \chi_{I_\sigma}(x)^{90} dx \lesssim |I_{T_{\text{fat}}}|$$

as I_σ is both much smaller than $I_{T^{\text{red}}}$ and not close to the boundary of $I_{T^{\text{red}}}$. This completes the proof of (5.4). ■

Proof of Lemma 4.3. We begin by verifying (4.5). For any $\sigma \in \mathbf{P}_{\text{thin}}$ and $m \geq 0$ observe that

$$(5.6) \quad |\phi_\sigma(x)| \lesssim |I_\sigma|^{-1/2} \chi_{I_\sigma}(x)^m, \quad x \in \mathbb{R}.$$

Indeed, after taking dilation and translation into account this estimate reduces to

$$\left| 2^{j/2} \int \tilde{\psi}(2^j \xi + \xi_0) e(\xi^2 + \xi y) d\xi \right| \lesssim 2^{-j/2} (1 + 2^j |y|)^{-m}, \quad y \in \mathbb{R}.$$

Here, $\tilde{\psi}$ is a Schwartz function supported in $1/2 \leq |\xi| \leq 2$ and $2^{j-1} \leq \xi_0 \leq 2^{j+1}$. But then at most $\lesssim 1$ oscillations of $e(\xi^2)$ are relevant to the integral, so the estimate follows by a repeated integration by parts. Then (5.6) plus a routine argument proves (4.5).

Turning to the estimate (4.4), we can assume that the tree has scales separated by a factor $4j$. Then the tree is a union of a 1-tree and a 2-tree. It suffices to prove (4.4) without the leading factor of j on the right for 1-trees and 2-trees.

We consider first the case of a 2-tree T_{thin} . In this case, because of the separation of scales, the sets $\omega_{\sigma 1}$ for $\sigma \in T_{\text{thin}}$ are disjoint and for $\sigma \neq \sigma' \in T_{\text{thin}}$, we can have $|I_{\sigma}| \neq |I_{\sigma'}|$, in which case ϕ_{σ} and $\phi_{\sigma'}$ are disjointly supported. Thus it suffices to consider the case of $|I_{\sigma}| = |I_{\sigma'}|$. Then $I_{\sigma} \cap I_{\sigma'} = \emptyset$, and the decay specified in (5.6) shows that

$$|M_j^{T_{\text{thin}}}(x)| \lesssim \text{size}(T_{\text{thin}}) \chi_{I_{T_{\text{thin}}}}(x)^8.$$

In either case, we have

$$|M_j^{T_{\text{thin}}}(x)| \lesssim \text{size}(T_{\text{thin}}) \chi_{I_{T_{\text{thin}}}}(x)^8.$$

That is, (4.4) is trivially satisfied in this case. [This argument is the key motivation for the definitions of \mathbf{P}_{fat} and \mathbf{P}_{thin} as well as those of 1-trees and 2-trees.]

We now turn to the case of a 1-tree T_{thin} . A specific case unlocks the general case. Suppose that T_{thin} is a tree with $|I_{\sigma}| = |I_{\sigma'}|$ for all $\sigma, \sigma' \in T_{\text{thin}}$ and $0 = c(\omega_{\sigma, 1})$. Then from (2.9) we have, for all $2 < p < \infty$,

$$\begin{aligned} (5.7) \quad \|M^{T_{\text{thin}}}\|_p &\leq \left\| D_j \left(\sum_{\sigma \in T_{\text{thin}}} \langle f, \varphi_{\sigma} \rangle \varphi_{\sigma} \right) \right\|_p \\ &\lesssim 2^{j(1-1/p)} \left\| \sum_{\sigma \in T_{\text{thin}}} \langle f, \varphi_{\sigma} \rangle \varphi_{\sigma} \right\|_p \\ &\lesssim 2^{j(1-1/p) + 2j(1-2/p)} \text{size}(T_{\text{thin}}) |I_{T_{\text{thin}}}|^{1/p}. \end{aligned}$$

The last line follows as there are 2^{2j} tiles σ in any $\text{thin}(s)$ for $s \in \mathbf{P}_{\text{fat}}$.

More generally, for any 1-tree T_{thin} , observe that there is a connection to the space of functions of bounded mean oscillation. The distinction between fat and thin tiles must enter into this relationship however. In particular,

$$(5.8) \quad \left\| \sum_{\sigma \in T_{\text{thin}}} \langle f, \varphi_{\sigma} \rangle \varphi_{\sigma} \right\|_{\text{BMO}} \lesssim 2^{2j} \text{size}(T_{\text{thin}}).$$

To verify this estimate, we need only consider the estimate for a 1-tree T_{thin} with $0 \in \omega_{T_{\text{thin}}}$, and then, only consider the mean oscillation over an interval of length $\simeq 1$. We may then restrict ourselves to the case in which $|I_\sigma| \geq 1$ for all $\sigma \in T_{\text{thin}}$. The point is then that

$$\left| \frac{d}{dx} \sum_{\sigma \in T_{\text{thin}}} \langle f, \varphi_\sigma \rangle \varphi_\sigma(x) \right| \lesssim 2^{2j \text{size}(T_{\text{thin}})}, \quad x \in \mathbb{R}.$$

Let $\omega_{T_{\text{thin}}} = \omega^1 \not\subset \omega^2 \not\subset \dots$ be the maximal sequence of dyadic intervals containing $\omega_{T_{\text{thin}}}$. Let $T_{\text{thin}}^l = \{s \in T_{\text{thin}} : \omega_s = \omega^l\}$. Then the functions $M^{T_{\text{thin}}^l}$ are disjointly supported in l . Hence,

$$\begin{aligned} \|M^{T_{\text{thin}}}\|_p^p &\leq \sum_l \|M^{T_{\text{thin}}^l}\|_p^p \lesssim 2^{jp(1-1/p)} \sum_l \left\| \sum_{\sigma \in T_{\text{thin}}^l} \langle f, \varphi_\sigma \rangle \varphi_\sigma \right\|_p^p \\ &\lesssim 2^{jp(1-1/p)} \left\| \left[\sum_l \left| \sum_{\sigma \in T_{\text{thin}}^l} \langle f, \varphi_\sigma \rangle \varphi_\sigma \right|^2 \right]^{1/2} \right\|_p^p \\ &\lesssim 2^{jp(3-5/p) \text{size}(T_{\text{thin}})} |I_{T_{\text{thin}}}|^{1/p}. \end{aligned}$$

Here we rely on $p > 2$, (5.8) and (5.7). ■

Proof of Lemma 4.6. We shall show that there is a set $E_n \subset \mathbb{R}$ so that $|E_n| \lesssim 2^n$ and the collection $\tilde{S}_{\text{fat}}(n) := \{s \in S_{\text{fat}}(n) : I_s \not\subset E_n\}$ is a union of collections $\tilde{U}_{\text{fat}}(n, k)$, $0 \leq k \leq -50n$ satisfying (i)–(iii).

The last three conditions of the lemma are trivially satisfied by making further subdivisions of the subcollections $U_{\text{fat}}(n, k)$, and making a small further contribution to the exceptional set E_n . Thus, the lemma will follow in complete generality.

Fix n and set $S_{\text{fat}} = S_{\text{fat}}(n)$. Condition (ii) is also easy to satisfy. For the first contribution to our exceptional set, define

$$E^1 := \left\{ x : \sum_{x \in S_{\text{fat}}^*} 1_{I_s}(x) > 2^{-10n} \right\}$$

where S_{fat}^* consists of the maximal elements of S_{fat} . By (4.2), $|E^1| \lesssim 2^n$. We can assume that for all $s \in S_{\text{fat}}$, $I_s \not\subset E^1$. Then certainly (ii) is true.

We now show that S_{fat} is decomposable into subcollections $U_{\text{fat}}(k)$, $1 \leq k \leq -10n$, which are uniquely decomposable into maximal trees. This last condition is true iff for each $s \in U_{\text{fat}}(k)$ there is a unique maximal $s^* \in U_{\text{fat}}^k$ with $s < s^*$. And this is so iff the collection $U_{\text{fat}}(k)$ does not admit a vee in the partial order on tiles. A *vee* is three tiles s, s', s'' with $s < s', s''$ but s' and s'' not comparable with respect to the partial order on tiles.

To achieve this, we employ a method of Fefferman [3]. Define a counting function

$$C(s) := \#\{s^* \in S_{\text{fat}}^* : s < s^*\}.$$

Then $C(s) \leq 2^{-10n}$ for all $s \in S_{\text{fat}}$ as (ii) is true. Take the set $U_{\text{fat}}(k)$ to be $\{s : 2^{k-1} \leq C(s) < 2^k\}$.

That these sets do not contain vees follows immediately from the observation that $C(s)$ is superadditive in this sense: If $s, s', s'' \in S_{\text{fat}}$ is a vee, then $C(s) \geq C(s') + C(s'')$. [Then if $s', s'' \in U_{\text{fat}}(k)$ we see that $C(s) \geq 2^{k-1} + 2^{k-1} = 2^k$, so it cannot be in $U_{\text{fat}}(k)$.] Indeed, there can be no maximal tile s''' larger than both s' and s'' , for this would force s' and s'' to be comparable in the partial order, as one checks immediately. Hence the maximal tiles greater than s' are disjoint from those greater than s'' , which proves the superadditivity property.

The last condition to verify is (iii), which requires another class of contributions to the exceptional set. Fix a choice of $1 \leq k \leq -10n$. Consider the maximal tiles $U_{\text{fat}}^*(k)$. We want to separate these tiles after expanding the coordinates I_s by a factor of $2^{-\varepsilon n}$. This can be done, up to an exceptional set and a further division of $U_{\text{fat}}^*(k)$, by applying Lemma 4.4 of [9] to $S = U_{\text{fat}}^*$, with $A = 2^{\varepsilon n}$. The details are omitted. ■

References

- [1] J. Barrionuevo and M. Lacey, *A weak-type orthogonality principle*, Proc. Amer. Math. Soc., to appear; math.CA/0201061.
- [2] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. 116 (1966), 135–157.
- [3] C. Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. 98 (1973), 551–571.
- [4] R. Hunt, *On the convergence of Fourier series*, in: Orthogonal Expansions and their Continuous Analogues (Edwardsville, IL, 1967), Southern Illinois Univ. Press, Carbondale, IL, 1968, 235–255.
- [5] M. Lacey, *On the bilinear Hilbert transform*, in: Proc. Internat. Congress of Mathematicians, Vol. II (Berlin, 1998), Doc. Math. 1998, Extra Vol. II, 647–656.
- [6] —, *The bilinear maximal functions map into L^p for $2/3 < p \leq 1$* , Ann. of Math. 151 (2000), 35–57.
- [7] M. Lacey and C. Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. 7 (2000), 361–370.
- [8] —, —, *On Calderón’s conjecture*, Ann. of Math. 149 (1999), 475–496.
- [9] —, —, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$* , Ann. of Math. 146 (1997), 693–724.
- [10] K. Oskolkov, *Schrödinger equation and oscillatory Hilbert transform of second degree*, J. Fourier Anal. Appl. 4 (1998), 341–356.
- [11] F. Ricci and E. M. Stein, *Harmonic analysis on nilpotent groups and singular integrals I*, J. Funct. Anal. 73 (1987), 179–194.
- [12] E. M. Stein, *Oscillatory integrals related to Radon-like transforms*, in: Proc. Conf. in Honor of Jean-Pierre Kahane (Orsay, 1993), J. Fourier Anal. Appl. 1995, Special Issue, 535–551.
- [13] E. M. Stein and S. Wainger, *The estimation of an integral arising in multiplier transformations*, Studia Math. 35 (1970), 101–104.

- [14] E. M. Stein and S. Wainger, *Oscillatory integrals related to Carleson's theorem*, Math. Res. Lett. 8 (2001), 789–800.
- [15] C. Thiele, *The quartile operator and pointwise convergence of Walsh series*, Trans. Amer. Math. Soc. 352 (2000), 5745–5766.

School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.
E-mail: lacey@math.gatech.edu
Web: <http://www.math.gatech.edu/~lacey>

Received June 11, 2001
Revised version September 9, 2002

(4757)