MAXIMAL THEOREMS FOR THE DIRECTIONAL HILBERT TRANSFORM ON THE PLANE

MICHAEL T. LACEY AND XIAOCHUN LI

ABSTRACT. For a Schwartz function f on the plane and a non-zero $v \in \mathbb{R}^2$ define the Hilbert transform of f in the direction v to be

$$\mathrm{H}_v \, f(x) = \mathrm{p.v.} \int_{\mathbb{R}} f(x - vy) \, \frac{dy}{y}.$$

Let ζ be a Schwartz function with frequency support in the annulus $1 \le |\xi| \le 2$, and $\boldsymbol{\zeta} f = \zeta * f$. We prove that the maximal operator $\sup_{|v|=1} |\mathrm{H}_v \, \boldsymbol{\zeta} f|$ maps L^2 into weak L^2, and L^p into L^p for $p > 2$. The L^2 estimate is sharp. The method of proof is based upon techniques related to the pointwise convergence of Fourier series. Indeed, our main theorem implies this result on Fourier series.

1. INTRODUCTION, PRINCIPAL THEOREM

Our interest is in the directional Hilbert transform applied in a choice of directions of the plane. Thus, for $v \in \mathbb{R}^2 - \{0\}$, set

$$\mathrm{H}_v \, f(x) = \mathrm{p.v.} \int_{\mathbb{R}} f(x - vy) \, \frac{dy}{y}.$$

This definition is independent of the length of v, and below we shall only concern ourselves with $|v| = 1$. Let ζ be a Schwartz function with frequency support in the annulus $1 \le |\xi| \le 2$, and $\boldsymbol{\zeta} f = \zeta * f$. Our theorem is

1.1. **Theorem.** *The maximal operator* $\mathrm{H}_v^* f := \sup_{|v|=1} |\mathrm{H}_v \, \boldsymbol{\zeta} f|$ *maps* L^2 *into weak* L^2, *and* L^p *into* L^p *for* $p > 2$.

This theorem is a complement to a corresponding result, due to J. Bourgain [3], for the directional maximal function, namely

$$(1.2) \qquad \mathrm{M}^* f(x) := \sup_{|v|=1} \sup_{t>0} (2t)^{-1} \int_{-t}^{t} |\boldsymbol{\zeta} f(x - yv)| \, dy.$$

See Section 3 for a discussion of norm bounds for this operator. For both this operator and H^*, the estimate of weak square integrability is sharp, as was pointed out to us by M. Christ [5]. This argument is summarized as follows. Begin with
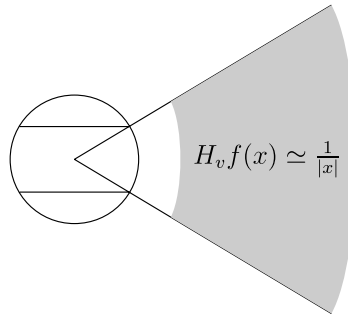
FIGURE 1. The function f is depicted by the circle centered at the origin with a planar wave in the vertical direction indicated by the two horizontal lines. In the shaded region, $H_v f \simeq \frac{1}{|x|}$.

a Schwartz function $\varphi \geq 0$ with frequency support in a small ball about the origin in the plane. Then consider $f(x_1, x_2) := e^{ix_2} \varphi(x_1, x_2)$. For any point in the plane $x = (x_1, x_2)$ with $|x_1| > 2$ and $|x_2| < \frac{1}{100}|x_1|$, consider the line that passes through x and the origin. The real part of f will not change sign on this line, so we compute the Hilbert transform with no cancellation. That is, $\mathrm{H}^* f(x) \simeq \mathrm{M}^* f(x) \simeq |x|^{-1}$, and $|x|^{-1}$ is just in weak L^2. See Figure 1.

Whereas, Bourgain's argument for the maximal estimate is not difficult, the theorem above is of necessity somewhat harder, as it implies the pointwise convergence of Fourier series in one dimension. This is an observation in the style of De Leeuw. One considers the trace of the operator in frequency variables along any line in the annulus $1 \leq |\xi| \leq 2$. This is Carleson's theorem, [4], but also see [6]. As such, we use a method which is adopted from the proof of Carleson's Theorem given by M. Lacey and C. Thiele [12].

Perhaps the principal novelty of this paper is the suitable adaptation of the time frequency analysis of Lacey and Thiele to the current setting of the plane and measurable choice of directions. In comparison to the proof of Lacey and Thiele, we find a rather precise analogy between the proofs on the real line and the proofs on the plane. Some differences arise from the notion of a *tile*, which requires some care to define, and the proof of the "size lemma", an orthogonality statement, requires a small amount of innovation.

Finally, there is an outstanding question, attributable to E. M. Stein [14], concerning the boundedness of the Hilbert transform on families of lines that are determined by, say, a Lipschitz map. Thus, for a map $v : \mathbb{R}^2 \to \{|x| = 1\}$, one wishes to know if

$$\int_{-1}^{1} f(x - yv(x)) \, \frac{dy}{y}$$

is a bounded operator on, say, L^2. Positive results for analytic and real analytic vector fields are due to, respectively, Nagel, Stein and Wainger [13], and to Bourgain [2]. In a subsequent paper, the authors [11] will prove that the operator above is bounded on L^2 if $v \in C^{1+\epsilon}$, for any positive ϵ. The results of this paper are a crucial aspect of the proof of this result.

The presentation of this paper has been substantially improved by skilled and generous referees.

2. Definitions and principle lemma

We begin with some conventions. We do not keep track of the value of generic absolute constants, instead using the notation $A \lesssim B$ iff $A \leq KB$ for some constant K. Also, $A \simeq B$ iff $A \lesssim B$ and $B \lesssim A$. We use the notation $\mathbf{1}_A$ to denote the indicator function of the set A, and the Fourier transform on $\mathbb{R}^2$ is denoted by $\widehat{f}(\xi) = \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} f(x) \, dx$, with a similar definition on the real line.

2.1. Definition. A *grid* is a collection of intervals $\mathcal{G}$ so that for all $I, J \in \mathcal{G}$, we have $I \cap J \in \{\emptyset, I, J\}$. The dyadic intervals are a grid. A grid $\mathcal{G}$ is *central* iff for all $I, J \in \mathcal{G}$, with $I \subset_{\neq} J$ we have $100I \subset J$.

The reader can find the details on how to construct such a central grid structure in [7].

Let ρ be rotation on $\mathbb{T}$ by an angle of $-\pi/2$. Coordinate axes for $\mathbb{R}^2$ are a pair of unit orthogonal vectors $(e, e_\perp)$ with $e = \rho e_\perp$.

2.2. Definition. We say that $\omega \subset \mathbb{R}^2$ is a *rectangle* if it is a product of intervals with respect to a choice of axes $(e, e_\perp)$ of $\mathbb{R}^2$. We will say that ω is an *annular rectangle* if $\omega = (-2^{l-1}, 2^{l-1}) \times (a, 2a)$ for an integer l with $2^l < a/8$, with respect to the axes $(e, e_\perp)$. The dimensions of ω are said to be $2^l \times a$. Note that the face $(-2^{l-1}, 2^{l-1}) \times a$ is tangent to the circle $|\xi| = a$ at the midpoint to the face, $(0, a)$. We say that the *scale of* ω is $\mathrm{scl}(\omega) := 2^l$ and that the *annular parameter of* ω is $\mathrm{ann}(\omega) := a$. In referring to the coordinate axes of an annular rectangle, we shall always mean $(e, e_\perp)$ as above.

It is imperative to keep in mind that the choice of basis in the definition of a tile depends upon the tile in question. These definitions are illustrated in Figure 2. Also see Figure 7.

Annular rectangles will decompose our functions in the frequency variables. But our methods must be sensitive to spatial considerations; it is this and the uncertainty principle that motivate the next definition.

2.3. Definition. Two rectangles R and R are said to be *dual* if they are rectangles with respect to the same basis $(e, e_\perp)$, thus $R = r_1 \times r_2$ and $\mathsf{R} = \mathsf{r}_1 \times \mathsf{r}_2$ for intervals $r_i, \mathsf{r}_i, \ i = 1, 2$. Moreover, $1 \leq |r_i| \cdot |\mathsf{r}_i| \leq 2$ for $i = 1, 2$. The product of two dual rectangles we shall refer to as a *phase rectangle*. The first coordinate of a phase rectangle we think of as a frequency component and the second as a spatial component.

We consider collections of phase rectangles $\mathcal{AT}$ which satisfy these conditions. For $s, s' \in \mathcal{AT}$ we write $s = R_s \times \omega_s$, and require that

$$(2.4) \qquad \omega_s \text{ is an annular rectangle,}$$

$$(2.5) \qquad R_s \text{ and } \omega_s \text{ are dual rectangles,}$$

$$(2.6) \qquad \text{the rectangles } R_s \text{ are the products of intervals from a central grid,}$$

$$(2.7) \qquad \{R \,:\, R \times \omega_s \in \mathcal{AT}\} \quad \text{partitions } \mathbb{R}^2, \text{ for all } \omega_s,$$

$$(2.8) \qquad \mathrm{ann}(\omega_s) = \mathsf{ann} \quad \text{for some fixed } \mathsf{ann},$$

$$(2.9) \qquad \sharp\{\omega_s \,:\, \mathrm{scl}(s) = 2^l, \ \mathrm{ann}(s) = \mathsf{ann}\} \geq c\,\mathsf{ann}\,2^{-l}, \qquad l \in \mathbb{Z}.$$
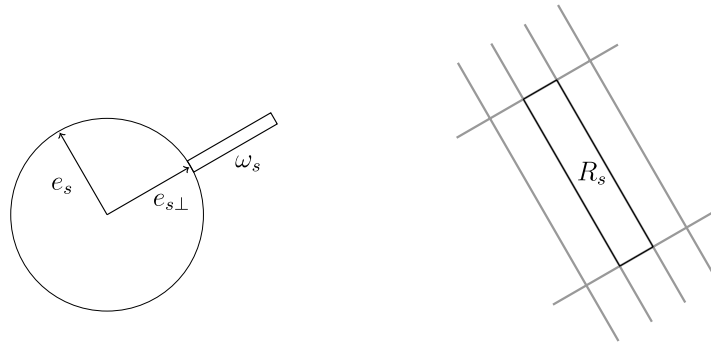
FIGURE 2. The two rectangles ω_s and R_s whose product is a tile. The gray rectangles are other possible locations for the rectangle R_s.
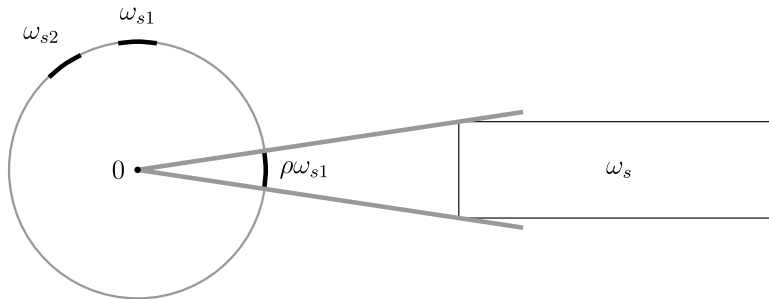

FIGURE 3. An annular rectangular ω_s, and three associated subintervals of $\rho\omega_{s1}$, ω_{s1}, and ω_{s2}.

We assume that there are auxiliary sets $\boldsymbol{\omega}_s, \boldsymbol{\omega}_{s1}, \boldsymbol{\omega}_{s2} \subset \mathbb{T}$ associated to s—or more specifically ω_s—which satisfy these properties:

$$(2.10) \qquad \boldsymbol{\Omega} := \{\boldsymbol{\omega}_s, \boldsymbol{\omega}_{s1}, \boldsymbol{\omega}_{s2} : s \in \mathcal{AT}\} \text{ is a grid in } \mathbb{T},$$

$$(2.11) \qquad \boldsymbol{\omega}_{s1} \cap \boldsymbol{\omega}_{s2} = \emptyset, \qquad \boldsymbol{\omega}_s := \text{hull}(\boldsymbol{\omega}_{s1}, \boldsymbol{\omega}_{s2}),$$

$$(2.12) \qquad \boldsymbol{\omega}_{s1} \text{ lies clockwise from } \boldsymbol{\omega}_{s2} \text{ on } \mathbb{T},$$

$$(2.13) \qquad |\boldsymbol{\omega}_s| \leq K \frac{\text{scl}(\omega_s)}{\text{ann}(\omega_s)},$$

$$(2.14) \qquad \left\{\frac{\xi}{|\xi|} : \xi \in \omega_s\right\} \subset \rho\boldsymbol{\omega}_{s1},$$

$$(2.15) \qquad \text{to each } \boldsymbol{\omega}_s \text{ there is a } \overline{\boldsymbol{\omega}}_s \in \boldsymbol{\Omega} \text{ such that } 2\boldsymbol{\omega}_s \subset \overline{\boldsymbol{\omega}}_s \subset 4\boldsymbol{\omega}_s.$$

Recall that ρ is the rotation that takes e into $e_\perp$. Thus, $e_{\omega_s} \in \boldsymbol{\omega}_{s1}$. See Figure 3.

Note that $|\boldsymbol{\omega}_s| \geq |\boldsymbol{\omega}_{s1}| \geq \text{scl}(\omega_s)/\text{ann}(\omega_s)$. Thus, e_{ω_s} is in $\boldsymbol{\omega}_{s1}$, and $\boldsymbol{\omega}_s$ serves as "the angle of uncertainty associated to R_s". Let us be more precise about the geometric information encoded into the angle of uncertainty. Let $R_s = r_s \times r_{s\perp}$ be as above. Choose another set of coordinate axes $(e', e'_\perp)$ with $e' \in \boldsymbol{\omega}_s$ and let R' be the product of the intervals r_s and $r_{s\perp}$ in the new coordinate axes. Then $K_0^{-1} R' \subset R_s \subset K_0 R'$ for an absolute constant $K_0 > 1$.

We say that *annular tiles* are collections $\mathcal{AT}(\mathsf{ann})$ satisfying the conditions (2.4)–(2.14) above. The constant $\mathsf{ann} > 0$ appears in (2.8) and (2.9). We extend the definition of e_ω, $e_{\omega\perp}$, $\mathrm{ann}(\omega)$ and $\mathrm{scl}(\omega)$ to annular tiles in the obvious way, using the notation e_s, $e_{s\perp}$, $\mathrm{ann}(s)$ and $\mathrm{scl}(s)$.

A phase rectangle will have two distinct functions associated to it. In order to define these functions, set

$$\mathrm{Tran}_y f(x) := f(x - y), \quad y \in \mathbb{R}^2 \quad \text{(Translation)},$$

$$\mathrm{Mod}_\xi f(x) := \mathrm{e}^{i\xi \cdot x} f(x), \quad \xi \in \mathbb{R}^2 \quad \text{(Modulation)},$$

$$\mathrm{Dil}_{R_1 \times R_2}^p f(x_1, x_2) := \frac{1}{(|R_1||R_2|)^{1/p}} f\left(\frac{x_1}{|R_1|}, \frac{x_2}{|R_2|}\right) \quad \text{(Dilation)}.$$

In the last display, $0 < p < \infty$, $R_1 \times R_2$ is a rectangle, and the coordinates (x_1, x_2) are those of the rectangle. Note that the definition depends only on the side lengths of the rectangle, and not the location, and that it preserves the L^p norm.

For a function φ and tile $s \in \mathcal{AT}$ set

$$\varphi_s := \mathrm{Mod}_{c(\omega_s)} \mathrm{T}_{c(R_s)} \mathrm{Dil}_{R_s}^2 \varphi,$$

where $c(J)$ is the center of J. Below, we shall consider φ to be a Schwartz function for which $\widehat{\varphi} \geq 0$ is supported in a small ball about the origin in $\mathbb{R}^2$, and is identically 1 on another smaller ball around the origin.

We introduce the tool to decompose the singular integral kernels, and a measurable function $v : \mathbb{R}^2 \to \{|x| = 1\}$ that achieves the maximum in our operator, up to a factor of say $\frac{1}{2}$. We consider a class of functions ψ_t, $t > 0$, so that

(2.16)   each ψ_t is supported in frequency in $[-\theta - \kappa, -\theta + \kappa]$,

(2.17) $|\psi_t(x)| \lesssim (1 + |x|)^{-1/\kappa}$.

In these conditions, $0 < \kappa < 1$ is a small fixed constant that need not concern us. In the top line, θ is a fixed positive constant that is only needed to ensure that (2.18) below is true. For the current paper, we will need only a single choice of function ψ, but for the authors' subsequent paper [11], countably many choices of ψ will be needed.

Define

(2.18)
$$\phi_s(x) := \int_{\mathbb{R}} \varphi_s(x - yv(x))\psi_s(y) \, dy$$

$$= \mathbf{1}_{\boldsymbol{\omega}_{s2}}(v(x)) \int_{\mathbb{R}} \varphi_s(x - yv(x))\psi_s(sy) \, dy,$$

(2.19) $\psi_s(y) := \mathrm{scl}(s)\psi_{\mathrm{scl}(s)}(\mathrm{scl}(s)y).$

An essential feature of this definition is that the support of the integral is contained in the set $\{v(x) \in \boldsymbol{\omega}_{s2}\}$, a fact which can be routinely verified, for an appropriate choice of θ in (2.16). See Figure 4. That is, we can insert the indicator $\mathbf{1}_{\boldsymbol{\omega}_{s2}}(v(x))$ without loss of generality. The set $\boldsymbol{\omega}_{s2}$ serves to localize the vector field, while $\boldsymbol{\omega}_{s1}$ serves to identify the location of φ_s in the frequency coordinate.

Remark. It seems likely that we could prove our theorem using the simpler definition $\phi_s = (\mathbf{1}_{\boldsymbol{\omega}_{s2}} \circ v)\varphi_s$. This would be the natural analog of the decomposition used by Lacey and Thiele [12]. But, in our subsequent paper on this subject [11], we will need to consider a truncation of the Hilbert transform, which suggests the definition
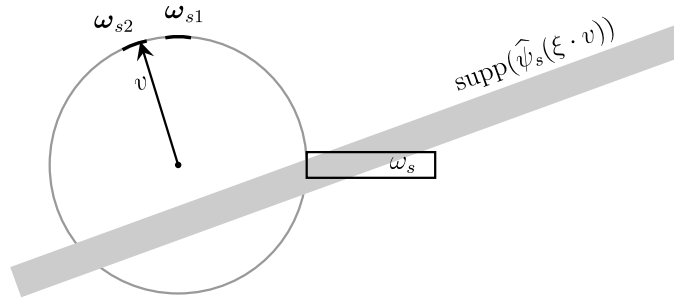
FIGURE 4. The support of ϕ_s is contained in $v^{-1}(\boldsymbol{\omega}_{s2})$.

of ϕ_s we have adopted above. We will also want to rely upon facts proved in this paper, to deduce our theorems concerning smooth vector fields.

The model operators we consider are defined by

$$\mathcal{C}_{\mathsf{ann},v}f := \sum_{s \in \mathcal{AT}(\mathsf{ann})} \langle f, \varphi_s \rangle \phi_s \,.$$

In this display, $\mathcal{AT}(\mathsf{ann}) := \{s \in \mathcal{AT} : \mathsf{ann}(s) = \mathsf{ann}\}$.

2.20. **Lemma.** *Assume that the vector field is only measurable. The operator $\mathcal{C}_{\mathsf{ann},v}$ extends to a bounded map from L^2 into weak L^2, and L^p into itself for $2 < p < \infty$. The norm of the operator is independent of ann.*

We shall prove that $\mathcal{C}_{\mathsf{ann},v}$ maps L^2 into weak L^2, with a constant independent of ann and v. By duality, it suffices to show that for all $f \in L^2 \cap L^\infty$ of L^2 norm one, and sets $F \subset \mathbb{R}^2$ of finite measure,

$$(2.21) \qquad |\langle \mathcal{C}_{\mathsf{ann},v}f, \mathbf{1}_F \rangle| \le \sum_{s \in \mathcal{AT}(\mathsf{ann})} |\langle f, \varphi_s \rangle \langle \phi_s, \mathbf{1}_F \rangle| \lesssim |F|^{1/2}.$$

By dilation invariance of the $\mathcal{C}_{\mathsf{ann},v}$ with respect to powers of 2, we can further take $1 \le |F| < 2$.

For the case of L^p, $2 < p < \infty$, we shall demonstrate the restricted-type estimate

$$|\{\mathcal{C}_{\mathsf{ann},v}\mathbf{1}_E > \lambda\}| \lesssim \lambda^{-p}|E|, \qquad 2 < p < \infty.$$

We need only consider this for $\lambda > 1$, by the weak L^2 bound. Moreover, this inequality follows from

$$|\langle \mathcal{C}_{\mathsf{ann},v}\mathbf{1}_E, \mathbf{1}_F \rangle| \le \sum_{s \in \mathcal{AT}(\mathsf{ann})} |\langle \varphi_s, \mathbf{1}_E \rangle \langle \phi_s, \mathbf{1}_F \rangle|$$

$$(2.22) \qquad\qquad \lesssim |F|\{\log(|E|/|F|)\}, \qquad 9|F| < |E|.$$

Observe that by dilation invariance we can assume that $|E| \simeq 1$. We also assume that the vector field v is defined only on F.

The proofs of (2.21) and (2.22) are given in Section 5, and follow lines of argument that are similar to the one-dimensional case, as given by Lacey and Thiele [12] for the case $p = 2$, and as in the work of Grafakos, Tao, and Terwilleger [8] for the case of $p > 2$. Also see the survey by Lacey [10].

Remark. The results for $p > 2$ are available essentially because functions that are in L^p are also locally square integrable. In one dimension, the proof of convergence of Fourier series for $1 < p < 2$ is closely linked to the boundedness of the Hardy Littlewood maximal function. In the current setting, the corresponding maximal function would be the unbounded Kakeya maximal function. The relationship between maximal functions and L^p bounds for $1 < p < 2$ has a subtle role to play in our companion paper [11].

3. A MAXIMAL FUNCTION ESTIMATE

We outline a proof of the weak-type estimate on L^2 for the directional maximal function M* given in (1.2).

3.1. Proposition. *For functions $f \in L^2$ and $\lambda > 0$, we have the estimate*

$$|\{M^* f > \lambda\}| \lesssim \lambda^{-2} \|f\|_2^2.$$

It suffices to give the same bound for the model operator

$$\mathcal{M}_{\mathsf{ann},v} f := \sup_{s \in \mathcal{AT}(\mathsf{ann})} |\langle f, \varphi_s \rangle \phi_s|.$$

We remark that this inequality is an immediate consequence of (2.20). The sum defining $\mathcal{C}_{\mathsf{ann},v}$ is in fact unconditional convergent as a sum over tiles. As a consequence the square function below, which trivially dominates $\mathcal{M}_{\mathsf{ann},v}$, satisfies the weak-type inequality:

$$\Big[\sum_{s \in \mathcal{AT}(\mathsf{ann})} |\langle f, \varphi_s \rangle \phi_s|^2 \Big]^{1/2}.$$

There is however a more direct proof. The fundamental fact, a simpler instance of the size lemma, Lemma 5.5, is this.

3.2. Lemma. *Fix $f \in L^2$ and $\lambda > 0$. Let $\mathcal{S} \subset \mathcal{AT}(\mathsf{ann})$ be a collection of tiles which are pairwise incomparable with respect to the partial order '<', and moreover for each $s \in \mathcal{S}$, $|\langle f, \varphi_s \rangle| \geq \lambda\sqrt{|R_s|}$. Then it is the case that*

$$\sum_{s \in \mathcal{S}} |R_s| \leq \lambda^{-2} \|f\|_2^2.$$

This is proved in e.g. Barrionuevo and Lacey [1], or one can inspect the proof of Lemma 5.5 for a number of simplifications which apply in this case.

It is then clear that the maximal function below maps L^2 into weak L^2:

$$A_1 f := \sup_{\mathcal{AT}(\mathsf{ann})} \frac{|\langle f, \varphi_s \rangle|}{\sqrt{|R_s|}} \mathbf{1}_{R_s}.$$

For $t > 2$, set

$$A_t f := \sup_{\mathcal{AT}(\mathsf{ann})} \frac{|\langle f, \varphi_s \rangle|}{\sqrt{|R_s|}} \mathbf{1}_{tR_s}.$$

This also maps L^2 into weak L^2 with norm at most $\lesssim t^2$. Indeed, given $\lambda > 0$, let $\mathcal{S}$ denote the maximal tiles with $|\langle f, \varphi_s \rangle| \geq \lambda\sqrt{|R_s|}$. Then, it is the case that

$$\{A_t f > \lambda\} \subset \bigcup_{s \in \mathcal{S}} KtR_s$$

for an absolute constant K. Hence, the claimed bound follows from Lemma 3.2. Then observe that

$$\mathcal{M}_{\mathsf{ann},v} f \lesssim \sum_{k=0}^{\infty} 2^{-4k} \, \mathrm{A}_{2^k} f.$$

4. Proof of Theorem 1.1

We show that (2.20) implies our main result, Theorem 1.1. Indeed, the Hilbert transform in the direction of v is, in the limit, an appropriate average of the discrete sums formed in the previous section. Let λ be a smooth radial function satisfying

(4.1) $$\mathbf{1}_{[\frac{5}{4}, \frac{7}{4}]}(|\xi|) \leq \widehat{\lambda} \leq \mathbf{1}_{[1,2]}(|\xi|).$$

Let $\lambda_t(y) = t^2 \lambda(ty)$, where we choose the dilation so that t is the Fourier parameter, as opposed to the more common definition $t^{-2}\lambda(y/t)$, where t is the spatial parameter.

Let K be the distribution on $\mathbb{R}$

$$\sum_{j=-\infty}^{\infty} 2^j \psi(2^j y),$$

where ψ is a Schwartz function with $\widehat{\psi} \geq 0$ supported in a small neighborhood of -1. In particular, the distribution

$$\int_0^1 2^s K(2^s y) \, ds$$

is a non–zero multiple of the distribution associated with projection onto the negative frequencies on $\mathbb{R}$, which is itself a linear combination of the identity operator and the Hilbert transform.

To prove Theorem 1.1 it suffices to demonstrate the same norm estimates for the linear operators in which the distribution $1/y$ is replaced by $K(y)$, namely for

(4.2) $$\mathrm{T}_v f(x) = \int \lambda_{\mathsf{ann}} * f(x - yv(x)) K(y) \, dy$$

in which $\mathsf{ann} > 0$, and the measurable $v(x)$ are arbitrary.

For values of $2^j < \mathsf{ann}$, let

$$\widetilde{\mathrm{S}}_{j,\mathsf{ann}} f = \sum_{\substack{s \in \mathcal{AT}(\mathsf{ann}) \\ \mathrm{scl}(s)=2^j}} \langle f, \varphi_s \rangle \varphi_s,$$

$$\mathrm{T}_{j,v} f(x) = \int f(x - yv(x)) 2^j \psi(2^j y) \, dy.$$

Also, let Rot_τ be the operation of rotation by angle τ. The main point concerns the operator

(4.3)
$$\mathrm{S}_{j,\mathsf{ann}} := \lim_{Y \to \infty} -\!\!\!\!\!\!\int_{\mathrm{Box}(Y)} \mathrm{Dil}^2_{[0,2^{-s}) \times [0,2^{-s})} \mathrm{Rot}_{-\tau} \mathrm{Tran}_{-y}$$
$$\widetilde{\mathrm{S}}_{j,\mathsf{ann}} \mathrm{Tran}_y \mathrm{Rot}_\tau \mathrm{Dil}^2_{[0,2^s) \times [0,2^s)} \, dy \, d\tau \, ds,$$

where $\mathrm{Box}(Y) = [0, Y]^2 \times [0, 2\pi] \times [-1, 1]$.

Observe these points. (1) The integral is an average over dilations, rotations, and translations. Note that the averaging over dilations is done uniformly in all directions, and that we are implicitly averaging with respect to the Haar measure

on $\mathbb{R}_+$, with the rationale being that the dilations we are using are a representation of that group. (2) Due to the assumption (2.7), the limit in Y will exist. If ω_s were fixed we would only need to average over all translations indexed by a fixed rectangle. (3) The rectangles ω_s fill a fixed percentage of the annulus, due the assumption (2.9). (4) The limit, applied to a Schwartz function f, is seen to exist.

4.4. Lemma. *For each* $2^j < \mathsf{ann}$, *we have the identity*

$$\mathrm{S}_{j,\mathsf{ann}}\,\lambda_{\mathsf{ann}} * f = c(j, \mathsf{ann})\lambda_{\mathsf{ann}} * f,$$

where the constant $c(j, \mathsf{ann})$ *satisfies* $c^{-1} < |c(j, \mathsf{ann})| < c$, *for some absolute constant* c.

To deduce bounds for (4.2), observe that (2.20) concerns norm bounds for the sums

$$\sum_j c(j, \mathsf{ann})^{-1}\,\mathrm{T}_{j,v}\,\widetilde{\mathrm{S}}_{j,\mathsf{ann}}\lambda_{\mathsf{ann}} * f.$$

Of course the coefficients $c(j, \mathsf{ann})^{-1}$ do not appear in (2.20). Yet the placement of the absolute values in (2.21) and (2.22) demonstrates that the sum is unconditional over tiles, so that we can impose an arbitrary bounded sequence of coefficients, as we have done here.

The same norm bounds hold for averages of these sums, such as the averages that occur in (4.3). Hence, by the lemma just stated, we deduce the norm bounds for the operators in (4.2). It remains to prove our lemma.

Before proving Lemma 4.4, we record a simple lemma on convolutions.

4.5. Lemma. *Let* φ *and* ϕ *be real-valued Schwartz functions on* $\mathbb{R}^2$. *Then,*

$$\int_{[0,1]^2} \sum_{m \in \mathbb{Z}^2} \langle f, \mathrm{Tran}_{y+m}\,\varphi \rangle\,\mathrm{Tran}_{y+m}\,\phi\,dy = f * \Phi,$$

$$\text{where} \qquad \Phi(x) = \iint \overline{\varphi(u)}\phi(x+u)\,du.$$

In particular, $\widehat{\Phi} = \overline{\widehat{\varphi}}\widehat{\phi}$.

The proof is immediate. The sum and integral in question is in fact the integral

$$\int_{\mathbb{R}^2} f(z)\overline{\varphi(z-y)}\phi(x-y)\,dydz,$$

and one changes variables, $u = z - y$.

Proof of Lemma 4.4. Fix $\omega \in \Omega$, and let

$$\widetilde{\mathrm{S}}_\omega f := \sum_{\substack{s \in \mathcal{AT}(\mathsf{ann}) \\ \omega_s = \omega}} \langle f, \varphi_s \rangle \varphi_s.$$

Recall (2.7). Consider the limit

$$\mathrm{S}_\omega := \lim_{Y \to \infty} Y^{-2} \int_{[0,Y]^2} \mathrm{Tran}_{-y}\,\widetilde{\mathrm{S}}_\omega\,\mathrm{Tran}_y\,dy.$$

This limit can be explicitly written as an integral which is as in Lemma 4.5 with an additional dilation. Thus, this limit is convolution with respect to Φ_ω, where $\widehat{\Phi}_\omega = \widehat{\varphi}_\omega\overline{\widehat{\varphi}_\omega}$.
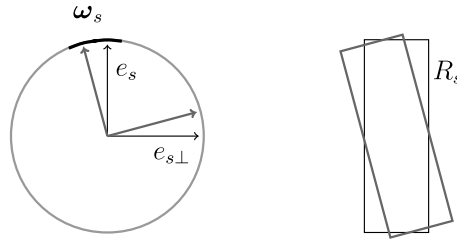
FIGURE 5. On the left, the interval $\boldsymbol{\omega}_s$ and the axes $(e_s, e_{s\perp})$. On the right, a rectangle R_s. In gray, the axes and rectangle R_s are rotated by an amount less than the length of $\boldsymbol{\omega}_s$.

By choice of φ, it has non–zero Fourier transform and is identically one on a small ball around the origin. Hence Φ_ω has Fourier transform with the same properties. The definition of $\mathrm{S}_{j,\mathsf{ann}}$ also incorporates an average over dilations and rotations. Concerning the average over rotations, recall (2.9), which states that we have an essentially maximal number of ω_s of a given scale. Concerning the dilations, observe that

$$\mathrm{Dil}^2_{[0,t)^2}\big[F * [\mathrm{Dil}^2_{[0,1/t)^2} G]\big] = [\mathrm{Dil}^1_{[0,t)^2} F] * G.$$

These observations lead to the conclusion that

$$\mathrm{S}_{j,\mathsf{ann}}\, f = c(j,\mathsf{ann})\Phi * f,$$

$$\Phi = \int_{-1}^{1}\int_{[0,2\pi]} \mathrm{Dil}^1_{[0,2^t)^2}\,\mathrm{Rot}_\tau\, \Phi\, d\tau\, dt.$$

The leading constant in the top line satisfies the claimed bound $c < |c(j,\mathsf{ann})| < c^{-1}$ for a positive constant c. The main point here is assumption (2.9). This function Φ has a Fourier transform that is identically constant for all $\mathsf{ann} < |\xi| < 2\mathsf{ann}$. The value of this constant depends only on the choice of φ. $\square$

5. THE MAIN LEMMAS

We need these notions associated to tiles and sets of tiles, all directly inspired by the proof of Carleson's theorem as given by Lacey and Thiele [12]. There are however some changes in the current context, due to the setting of the plane, and our frequency analysis in an annulus.

An important heuristic is that given a tile s, the interval $\boldsymbol{\omega}_s$ serves as an "angle of uncertainty" for the rectangle R_s. The geometric consequence of this can be described as follows. Write $R_s = R_{s1} \times R_{s2}$ in the coordinate axes $(e_s, e_{s\perp})$. There is an absolute constant K so that the following holds. Suppose that $(e, e_\perp)$ is some other choice of coordinate axes with

$$|e_s - e|,\ |e_{s\perp} - e_\perp| \le |\boldsymbol{\omega}_s|.$$

Write $\widetilde{R}_s = R_{s1} \times R_{s2}$, but the choice of coordinate axes is $(e, e_\perp)$. Then, we have $R_s \subset K\widetilde{R}_s$ and $\widetilde{R}_s \subset KR_s$. The implication is that R_s and $\widetilde{R}_s$ are essentially the same rectangles. See Figure 5.

There is a natural partial order on tiles, but it does take some care to define. For a fixed tile s, let $(e_s, e_{s\perp})$ be the coordinate system associated with R_s. While we have not done so to date, at this point we insist that the origin in each such
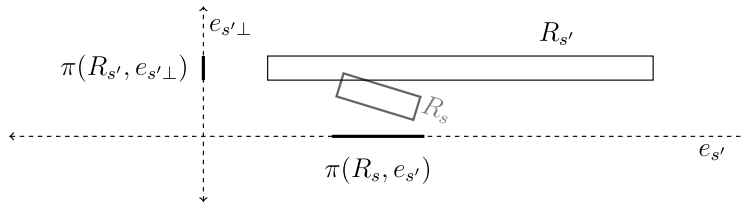
FIGURE 6. For tiles $s < s'$, only showing possible relative positions of R_s and R'_s, as well as the projections that define the partial order.

coordinate system specify the same point in the plane. Let $\mathcal{P}$ be the set of the projections of all R_s's onto e_s and let $\mathcal{P}_\perp$ be the set of the projections of all R_s's onto $e_{s\perp}$. We have assumed that $\mathcal{P}$ is a central grid in (2.6). Namely, recalling Definition 2.1, any two intervals $I \neq J$ in $\mathcal{P}$ which intersect, satisfy either $100I \subset J$, or $100J \subset I$. The reader can find the details on how to construct such a central grid structure in [7]. As all tiles lie in a single annulus, the collection of intervals $\mathcal{P}_\perp$ consist of disjoint intervals of equal length that partition $\mathbb{R}$.

Let $\pi(R, e_s)$ be the minimal interval in $\mathcal{P}$ containing the projection of R onto e_s, and let $\pi(R, e_{s\perp})$ be the minimal interval in $\mathcal{P}_\perp$ containing the projection of R onto $e_{s\perp}$.

For two tiles s, s', write $s < s'$ if and only if $\boldsymbol{\omega}_s \supset \boldsymbol{\omega}_{s'}$, $\pi(R_s, e_{s'}) \subset \pi(R_{s'}, e_{s'})$, and $\pi(R_{s'}, e_{s'\perp}) = \pi(R_{s'}, e_{s'\perp})$. This is illustrated in Figure 6.

We have been careful to define this partial order in such a way that it is transitive and that we have the conclusion

$$(5.1) \qquad \text{if } \boldsymbol{\omega}_s \times R_s \cap \boldsymbol{\omega}_{s'} \times R_{s'} \neq \emptyset, \text{ then } s \text{ and } s' \text{ are comparable under '<'.}$$

Proof. We can assume that $\boldsymbol{\omega}_s \supset \boldsymbol{\omega}_{s'}$, as these intervals of the circle come from a grid. So it suffices to check that $\pi(R_s, e_{s'}) \subset \pi(R_{s'}, e_{s'})$. By the assumption (2.6), in which κ is a small positive constant, we would have $\pi(R_s, e_{s'}) \subset 100\pi(R_s, e_s)$, and the last set is contained in $\pi(R_{s'}, e_{s'})$, by the central grid property. $\square$

A *tree* is a collection of tiles $\mathbf{T} \subset \mathcal{AT}(\mathsf{ann})$, for which there is a (non–unique) tile $\omega_{\mathbf{T}} \times R_{\mathbf{T}} \in \mathcal{AT}(\mathsf{ann})$ with $s < \omega_{\mathbf{T}} \times R_{\mathbf{T}}$ for all $s \in \mathbf{T}$. For $j = 1, 2$, call $\mathbf{T}$ a j–tree if the tiles $\{\boldsymbol{\omega}_{si} \times R_s : s \in \mathbf{T}\}$ are pairwise disjoint. 1–trees are especially important for us; Figure 7 presents a useful illustration of such a tree.

Observe that for any tree $\mathbf{T}$ we have, due to our comments about intervals of uncertainty,

$$R_s \subset KR_{\mathbf{T}}, \qquad s \in \mathbf{T}.$$

Thus, the spatial locations of the tiles are localized. In addition, as $\boldsymbol{\omega}_s \subset \boldsymbol{\omega}_{\mathbf{T}}$, we can regard all rectangles R_s as having a *fixed* set of coordinate axes, namely those of the top. This is a key point of the "tree lemma" below.

Fix a positive rapidly decreasing function χ, and for rectangle R set

$$(5.2) \qquad\qquad \chi_R = \mathrm{Tran}_{c(R)} \mathrm{Dil}_R^1 \chi,$$

$$\mathrm{dense}(s) := \sup_{s < s'} \int_{v^{-1}(\boldsymbol{\omega}_{s'})} \chi_{R_{s'}} \, dx,$$

$$\mathrm{dense}(\mathbf{S}) := \sup_{s \in \mathbf{S}} \mathrm{dense}(s),$$

$$\mathrm{sh}(\mathbf{S}) := \bigcup_{s \in \mathbf{S}} R_s \qquad (\text{the } \textit{shadow} \text{ of } \mathbf{S}),$$

$$\Delta(\mathbf{T})^2 := \sum_{s \in \mathbf{T}} \frac{|\langle f, \varphi_s \rangle|^2}{|R_s|} \mathbf{1}_{R_s}, \qquad \mathbf{T} \text{ is a 1--tree},$$

$$\mathrm{size}(\mathbf{S}) := \sup_{\substack{\mathbf{T} \subset \mathbf{S} \\ \mathbf{T} \text{ is a 1--tree}}} \big[|\mathrm{sh}(\mathbf{T})|^{-1} \sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle|^2 \big]^{1/2}.$$

Recall that we are to prove (2.21) and (2.22). In (2.21) $\|f\|_2 \simeq 1$, $|F| \leq 2$, and v is defined only on F.

The first two lemmas concern density and size, and are the same form. The third lemma, the 'tree lemma' relates size, density and trees.

5.3. Lemma. *Any* $\mathbf{S} \subset \mathcal{AT}(\mathsf{ann})$ *is the union of* $\mathbf{S}_{\mathrm{heavy}}$ *and* $\mathbf{S}_{\mathrm{light}}$ *satisfying these conditions:*

$$\mathrm{dense}(\mathbf{S}_{\mathrm{light}}) \leq \tfrac{1}{2} \mathrm{dense}(\mathbf{S}).$$

The collection $\mathbf{S}_{\mathrm{heavy}}$ *is a union of trees* $\mathbf{T} \in \mathcal{T}_{\mathrm{heavy}}$, *with*

$$(5.4) \qquad\qquad \sum_{\mathbf{T} \in \mathcal{T}_{\mathrm{heavy}}} |\mathrm{sh}(\mathbf{T})| \lesssim \mathrm{dense}(\mathbf{S})^{-1} |F|.$$

5.5. Lemma. *Any* $\mathbf{S} \subset \mathcal{AT}(\mathsf{ann})$ *is the union of* $\mathbf{S}_{\mathrm{big}}$ *and* $\mathbf{S}_{\mathrm{small}}$ *satisfying these conditions:*

$$\mathrm{size}(\mathbf{S}_{\mathrm{small}}) \leq \tfrac{1}{2} \mathrm{size}(\mathbf{S}).$$

The collection $\mathbf{S}_{\mathrm{big}}$ *is a union of trees* $\mathbf{T} \in \mathcal{T}_{\mathrm{big}}$, *with*

$$(5.6) \qquad\qquad \sum_{\mathbf{T} \in \mathcal{T}_{\mathrm{big}}} |\mathrm{sh}(\mathbf{T})| \lesssim \mathrm{size}(\mathbf{S})^{-2}.$$

5.7. Lemma. *For any tree* $\mathbf{T}$ *we have the estimate*

$$\sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle \langle \phi_s, \mathbf{1}_E \rangle| \lesssim \mathrm{dense}(\mathbf{T}) \, \mathrm{size}(\mathbf{T}) |R_{\mathbf{T}}|.$$

The first lemma has a proof which is essentially identical to the proof of the "mass lemma" in M. Lacey and C. Thiele [12]. We do not give the proof. The second lemma follows the well-established lines of proof, yet must introduce a new element or two to address the two-dimensional setting. The complete proof is given. The proof of the last lemma is quite close to that of M. Lacey and C. Thiele [12]. We shall give a proof.

The lemmas are combined in this way to prove (2.21), and hence (2.20) in the case of the weak L^2 estimate. Lemmas 5.3 and 5.5 should be applied so that their principal estimates (5.4) and (5.6) are approximately equal. The density of $\mathcal{AT}(\mathsf{ann})$ is at most a constant. The size of $\mathcal{AT}(\mathsf{ann})$ is at most a constant times the L^∞ norm of f. Thus, we can take the set of all tiles $\mathcal{AT}(\mathsf{ann})$ and decompose it

into sub-collections $\mathbf{S}_\sigma$, $\sigma \in \{2^n : n \in \mathbb{Z}\}$, so that $\mathbf{S}_\sigma$ is the union of trees $\mathbf{T} \in \mathcal{T}_\sigma$ such that

$$\sum_{\mathbf{T} \in \mathcal{T}_\sigma} |R_{\mathbf{T}}| \lesssim \sigma,$$

$$\mathrm{dense}(\mathbf{S}_\sigma) \lesssim \min(1, \sigma^{-1}),$$

$$\mathrm{size}(\mathbf{S}_\sigma) \lesssim \sigma^{-1/2}.$$

Hence, it follows that

$$\sum_{s \in \mathbf{S}_\sigma} |\langle f, \varphi_s \rangle \langle \phi_s, \mathbf{1}_E \rangle| \lesssim \sum_{\mathbf{T} \in \mathcal{T}} \min(1, \sigma^{-1}) \sigma^{-1/2} |R_{\mathbf{T}}|$$

$$\lesssim \min(\sigma^{1/2}, \sigma^{-1/2}).$$

This estimate is summable over $\sigma \in \{2^n : n \in \mathbb{Z}\}$ to an absolute constant. This completes the proof of (2.21).

A small variation on this argument proves (2.22). In this instance, the function $f = \mathbf{1}_E$, with $|E| \simeq 1$, so that the size of $\mathcal{AT}(\mathsf{ann})$ is $\lesssim 1$. Also, $\mathcal{AT}(\mathsf{ann})$ is a union of sub-collections $\mathbf{S}_\sigma$, $\sigma \in \{2^n : n \in \mathbb{N}\}$, so that $\mathbf{S}_\sigma$ is the union of trees $\mathbf{T} \in \mathcal{T}_\sigma$ satisfying

$$\sum_{\mathbf{T} \in \mathcal{T}_\sigma} |R_{\mathbf{T}}| \lesssim \sigma |F|,$$

$$\mathrm{dense}(\mathbf{S}_\sigma) \lesssim \sigma^{-1},$$

$$\mathrm{size}(\mathbf{S}_\sigma) \lesssim \min(1, (\sigma|F|)^{-1/2}).$$

Hence, it follows that

$$\sum_{s \in \mathbf{S}_\sigma} |\langle \mathbf{1}_E, \varphi_s \rangle \langle \phi_s, \mathbf{1}_F \rangle| \lesssim |F|^{1/2} \min(|F|^{1/2}, \sigma^{-1/2}).$$

Recall that in this instance, $|F| \leq \frac{1}{3}$. This estimate is summable over $\sigma \in \{2^n : n \in \mathbb{N}\}$ to $|F| \|\log|F|\|$. This completes the proof of (2.22). Our proof of (2.20) is complete, aside from the proofs of the three lemmas of this section.

6. The size lemma: Orthogonality

We give the proof of Lemma 5.5. We find that the proof involves for the most part a standard argument in the literature. See for instance the proof of the energy lemma in M. Lacey and C. Thiele [12]. Yet there is a point at which we will rely upon the strong maximal function, with respect to a choice of axes that is specified in a particular way by the set of tiles.

We can assume that all $\boldsymbol{\omega}_s$ lie in a fixed half circle. Note that in 1–trees, as the scale of tiles increase, the intervals $\boldsymbol{\omega}_{s1}$ move off in a clockwise direction, as depicted in Figure 7. We take advantage of this in a specific way below.

The initial step of the proof is to construct a collection of 1–trees $\mathbf{T} \in \mathcal{T}_+$ and use them to construct the collection of trees $\mathcal{T}_{\mathrm{big}}$. The process is inductive. Let $\mathrm{size}(\mathbf{S}) = \sigma$. Initialize

$$\mathcal{T}_+ := \emptyset, \qquad \mathcal{T}_{\mathrm{big}} = \emptyset, \qquad \mathbf{S}^{\mathrm{stock}} = \mathbf{S}.$$
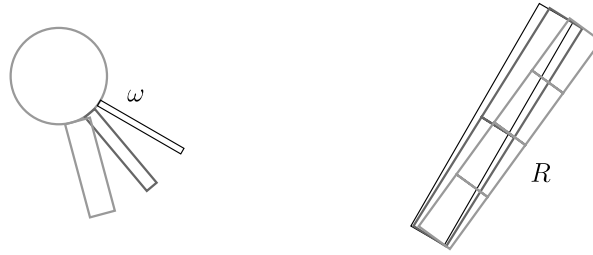
FIGURE 7. A few possible tiles in a 1–tree. Rectangles ω_s are on the left in different shades of gray. Possible locations of R_s are in the same shade of gray.

While size($\mathbf{S}^{\mathrm{stock}}$) $> \sigma/2$, select a 1–tree $\mathbf{T} \in \mathbf{S}^{\mathrm{stock}}$ such that

$$(6.1) \qquad \sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle|^2 \geq \frac{\sigma^2}{4} |R_{\mathbf{T}}|,$$

$|R_{\mathbf{T}}|$ is maximal, and $\boldsymbol{\omega}_{\mathbf{T}}$ is most clockwise. Then set $\tau(\mathbf{T}) \subset \mathbf{S}^{\mathrm{stock}}$ to be the maximal (with respect to inclusion) tree in $\mathbf{S}^{\mathrm{stock}}$ with top $\boldsymbol{\omega}_{\mathbf{T}} \times R_{\mathbf{T}}$. Update

$$\mathcal{T}_+ = \mathcal{T}_+ \cup \{\mathbf{T}\}, \qquad \mathcal{T}_{\mathrm{big}} = \mathcal{T}_{\mathrm{big}} \cup \{\tau(\mathbf{T})\}, \qquad \mathbf{S}^{\mathrm{stock}} = \mathbf{S}^{\mathrm{stock}} - \mathbf{T}.$$

When size($\mathbf{S}^{\mathrm{stock}}$) $< \sigma/2$, set $\mathbf{S}^{\mathrm{stock}} = \mathcal{T}_{\mathrm{small}}$, and the process stops.

To conclude the lemma, we need to show that

$$(6.2) \qquad \sum_{\mathbf{T} \in \mathcal{T}_+} |R_{\mathbf{T}}| \lesssim \sigma^{-2}.$$

We have constructed these trees so that they satisfy a property very useful to the verification of this inequality. Suppose $\mathbf{T} \neq \mathbf{T}' \in \mathcal{T}_+$, and $s \in \mathbf{T}$ and $s' \in \mathbf{T}'$ with $\boldsymbol{\omega}_s \subset \boldsymbol{\omega}_{s'1}$. Then $R_{\mathbf{T}} \cap R_{s'} = \emptyset$. We refer to this property as "strong disjointness". To see that it holds, we have $\boldsymbol{\omega}_{\mathbf{T}} \subset \boldsymbol{\omega}_{s'1}$, so that $\boldsymbol{\omega}_{\mathbf{T}}$ lies in the clockwise direction from $\boldsymbol{\omega}_{\mathbf{T}'}$. Hence $\mathbf{T}$ was constructed before $\mathbf{T}'$. Thus, if $R_{\mathbf{T}} \cap R_{s'} \neq \emptyset$, we would conclude from (5.1) that $s' < \boldsymbol{\omega}_{\mathbf{T}} \times R_{\mathbf{T}}$. Hence, $s' \in \tau(\mathbf{T})$, and so s' would have been removed from $\mathbf{S}^{\mathrm{stock}}$ and so could not be in $\mathbf{T}'$.

Adopt the notation $F(\mathbf{S}') = \sum_{s \in \mathbf{S}'} \langle f, \varphi_s \rangle \varphi_s$. Now observe that for $\mathbf{S}_+ = \bigcup_{\mathbf{T} \in \mathcal{T}_+} \mathbf{T}$, we have

$$\sigma^2 \sum_{\mathbf{T} \in \mathcal{T}_+} |R_{\mathbf{T}}| \lesssim \sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi \rangle_s, f$$
$$= \langle f, F(\mathbf{S}_+) \rangle$$
$$\leq \|f\|_2 \|F(\mathbf{S}_+)\|_2.$$

This follows by Cauchy–Schwarz. Recall that $\|f\|_2 = 1$. To conclude (6.2) we need only show that

$$(6.3) \qquad \|F(\mathbf{S}_+)\|_2 \lesssim \sigma \Big[\sum_{\mathbf{T} \in \mathcal{T}_+} |R_{\mathbf{T}}| \Big]^{1/2}.$$

For $s \in \mathbf{T}$, set

$$\mathbf{B}_=(s) = \{s' \in \mathbf{S}_+ - \mathbf{T} : \boldsymbol{\omega}_s = \boldsymbol{\omega}_{s'}\},$$
$$\mathbf{B}(s) = \{s' \in \mathbf{S}_+ - \mathbf{T} : \boldsymbol{\omega}_s \subset_{\neq} \boldsymbol{\omega}_{s'}\}.$$

Note that if $s' \in \mathbf{S}_+ - \mathbf{T}$ is such that $\langle \varphi_s, \varphi_{s'} \rangle \neq 0$, then $s' \in \mathbf{B}_=(s) \cup \mathbf{B}(s)$. Also, if $s' \notin \mathbf{B}_=(s) \cup \mathbf{B}(s)$, then $\langle \varphi_s, \varphi_{s'} \rangle = 0$. We estimate

$$(6.4) \qquad \|F(\mathbf{S}_+)\|_2^2 \lesssim \sum_{\mathbf{T} \in \mathcal{T}_+} \|F(\mathbf{T})\|_2^2$$

$$(6.5) \qquad + \sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi \rangle_s, F(\mathbf{B}_=(s)))$$

$$(6.6) \qquad + \sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{B}(s))) \rangle.$$

It is a routine matter to verify that for each 1–tree $\mathbf{T} \in \mathcal{T}_+$,

$$\|F(\mathbf{T})\|_2 \lesssim \sigma |R_{\mathbf{T}}|^{1/2}.$$

Therefore, the right-hand side of (6.4) is no more than

$$\sum_{\mathbf{T} \in \mathcal{T}_+} \|F(\mathbf{T})\|_2^2 \lesssim \sum_{\mathbf{T} \in \mathcal{T}_+} \sigma^2 |R_{\mathbf{T}}|.$$

For the term (6.5), we use this estimate for tiles s, s' with $\boldsymbol{\omega}_s = \boldsymbol{\omega}_{s'}$,

$$|\langle \varphi_s, \varphi_{s'} \rangle| \lesssim \operatorname{Tran}_{c(R_s)} \operatorname{Dil}_{R_s}^0 \chi(c(R_{s'})).$$

This with Cauchy–Schwarz gives

$$(6.5) \leq \Big[\sum_{s \in \mathbf{S}_+} |\langle f, \varphi_s \rangle|^2 \sum_{s \in \mathbf{S}_+} \Big| \langle \varphi_s, F(\mathbf{B}_=(s)) \rangle \Big|^2 \Big]^{1/2}$$

$$\lesssim \sum_{s \in \mathbf{S}_+} |\langle f, \varphi_s \rangle|^2$$

$$\lesssim \sigma^2 \sum_{\mathbf{T} \in \mathcal{T}_+} |R_{\mathbf{T}}|,$$

as required by (6.3).

As for (6.6), use Cauchy–Schwarz in s to bound it by

$$\sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{B}(s)) \rangle \leq \Big[\sum_{s \in \mathbf{S}_+} |\langle f, \varphi_s \rangle|^2 \sum_{s \in \mathbf{S}_+} |\langle \varphi_s, F(\mathbf{B}(s)) \rangle|^2 \Big]^{1/2}$$

$$\lesssim \Big[\sigma^2 \sum_{\mathbf{T} \in \mathcal{T}_+} |R_{\mathbf{T}}| \sum_{s \in \mathbf{S}_+} |\langle \varphi_s, F(\mathbf{B}(s)) \rangle|^2 \Big]^{1/2}.$$

We show that for each $\mathbf{T} \in \mathcal{T}_+$,

$$(6.7) \qquad \sum_{s \in \mathbf{T}} |\langle \varphi_s, F(\mathbf{B}(s)) \rangle|^2 \lesssim \sigma^2 |R_{\mathbf{T}}|.$$

This will complete the proof of (6.3).

Observe that if we set $\mathbf{B} = \bigcup_{s \in \mathbf{T}} \mathbf{B}(s)$, we have

$$\langle \varphi_s, F(\mathbf{B}(s)) \rangle = \langle \varphi_s, F(\mathbf{B}) \rangle.$$

This is a consequence of the grid structure on the intervals $\boldsymbol{\omega}_s$. Moreover, all tiles $s \in \mathbf{T} \cup \mathbf{B}$ have $\boldsymbol{\omega}_s \supset \boldsymbol{\omega}_{\mathbf{T}}$. This has two implications. The first is that all rectangles R_s can be regarded as rectangles with respect to a fixed set of coordinate axes, those for the top of the tree. Let M denote the strong maximal function computed in these coordinate directions.

The second is that the strong disjointness condition applies to each pair of tiles in the collection $\mathbf{B}$. This yields the essential observation that the rectangles $\{R_s : s \in \mathbf{B}\}$ are pairwise disjoint, and do not intersect $R_{\mathbf{T}}$.

Make a further diagonalization of the set $\mathbf{B}$. Set $\mathbf{B}_1 = \{s \in \mathbf{B} : R_s \subset 4R_{\mathbf{T}}\}$, and for $k > 1$, set $\mathbf{B}_k = \{s \in \mathbf{B} : R_s \subset 4^k R_{\mathbf{T}}, \ R_s \not\subset 4^{k-1} R_{\mathbf{T}}\}$. Let us point out that
$$\|F(\mathbf{B}_k)\|_2 \lesssim 4^k \sigma |R_{\mathbf{T}}|^{1/2}.$$
Indeed, recalling the notation (5.2),
$$|F(\mathbf{B}_k)| \lesssim \sigma \sum_{s \in \mathbf{B}_k} \chi_{R_s} * \mathbf{1}_R.$$
Therefore, for any function g,
$$\langle F(\mathbf{B}_k), g \rangle \lesssim \sigma \int \sum_{s \in \mathbf{B}_k} \mathbf{1}_{R_s} \chi_{R_s} * g \, dx$$
$$\lesssim \sigma \int \sum_{s \in \mathbf{B}_k} \mathbf{1}_{R_s} \, \mathrm{M} \, g \, dx$$
$$\lesssim \sigma 4^k |R_{\mathbf{T}}|^{1/2} \|g\|_2.$$
Clearly, this implies (6.7) for $\mathbf{B}_1$.

For $k > 1$, we can strengthen our inequality to the following:
$$\|F(\mathbf{B}_k)\|_{L^2(4^{k-1}R_{\mathbf{T}})} \lesssim 4^{-10k} |R_{\mathbf{T}}|^{1/2}.$$
We use this, together with the fact that the functions $\{\varphi_s : s \in \mathbf{T}\}$ are equivalent to an orthonormal basic sequence. Hence,
$$\sum_{s \in \mathbf{T}} |\langle \varphi_s, F(\mathbf{B}_k) \mathbf{1}_{4^{k-1}R_{\mathbf{T}}} \rangle|^2 \lesssim 4^{-k} |R_{\mathbf{T}}|.$$
But the functions $\{\varphi_s : s \in \mathbf{T}\}$ are highly concentrated in a neighborhood of $R_{\mathbf{T}}$. In particular, for any function g,
$$\sum_{s \in \mathbf{T}} |\langle \varphi_s \mathbf{1}_{\{\mathbb{R}^2 - 4^{k-1}R_{\mathbf{T}}\}}, g \rangle|^2 \lesssim 4^{-10k} \|g\|_2^2.$$
Clearly, this completes the proof of (6.7).

Proof of Lemma 5.7. We may fix the vector field, and assume that the standard basis $(e, e_\perp)$ is the basis for the rectangle $R_{\mathbf{T}}$. As a consequence, we can without loss of generality assume that this is basis for all the tiles $s \in \mathbf{T}$, and we write $R_s = R_{s,e} \times R_{s,e_\perp}$. Set $\delta = \mathrm{dense}(\mathbf{T})$ and $\sigma = \mathrm{size}(\mathbf{T})$. For this proof, we set
$$\chi_{R_s}^{(p)} = \mathrm{Tran}_{c(R_s)} \mathrm{Dil}_{R_s}^p, \qquad 0 < p < \infty.$$

Let $\mathcal{J}_e$ be those maximal dyadic intervals J in $\mathbb{R}$ for which $3J$ does not contain an interval $R_{s,e}$ for some $s \in \mathbf{T}$. This collection partitions $\mathbb{R}$. Let $\mathcal{J}_{e_\perp}$ be that partition of $\mathbb{R}$ into maximal dyadic intervals J such that $R_{\mathbf{T},e_\perp} \not\subset 3J$. Let $\mathcal{K} = \mathcal{J}_e \times \mathcal{J}_{e_\perp}$. For each rectangle $K \in \mathcal{K}$, set $\mathbf{T}(K, \pm)$ to be those $s \in \mathbf{T}$ for which $\pm \log |R_{s,e}|/|K_e| > 0$.

Choose signs $\varepsilon_s \in \{\pm\}$ such that

$$\sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle \langle \phi_s, \mathbf{1}_F \rangle| = \int_F \varepsilon_s \sum_{s \in \mathbf{T}} \langle f, \varphi_s \rangle \phi_s \, dx.$$

Let

$$A^{\mathbf{S}'} = \sum_{s \in \mathbf{S}'} \varepsilon_s \langle f, \varphi_s \rangle \phi_s.$$

Estimate, without appeal to cancellation,

$$\int_K |A^{\mathbf{T}(K,-)}| \, dx \lesssim \sigma \int_K \sum_{s \in \mathbf{T}(K,-)} \mathbf{1}_{\boldsymbol{\omega}_{s2}}(v(x)) \chi_{R_s}^{(\infty)}(x) \, dx.$$

The rectangles R_s are smaller than those of K, and do not intersect it. For each $x \in K$, the numbers $\chi_{R_s}^{(\infty)}(x) \lesssim 1$, and they decrease as $\mathrm{dist}(R_s, K) \cdot \mathrm{scl}(s)$ increases. The integral above is at most

$$\int_K |A^{\mathbf{T}(K,-)}| \, dx \lesssim \delta\sigma \min(|K|, |R_{\mathbf{T}}| \sup_{s \in \mathbf{T}} \sup_{x \in K} \chi_{R_s}^{(\infty)}(x)).$$

This is summable over $K \in \mathcal{K}$ to at most $\lesssim \delta\sigma |R_{\mathbf{T}}|$.

If $\mathbf{T}(K,+)$ is non-empty, then $K \subset 4R_{\mathbf{T}}$. Set

$$G_K := K \cap \bigcup_{s \in \mathbf{T}(K,+)} v^{-1}(\boldsymbol{\omega}_{s2}),$$

which contains the support of $\mathbf{1}_K A^{\mathbf{T}(K,+)}$. Our assertion is that $|G_K| \lesssim \delta |K|$. To see this, let $K = K_e \times K_{e\perp}$, and let K'_e be the dyadic interval that contains K_e and $|K'_e| = 2|K_e|$. This interval is smaller than that of $R_{\mathbf{T},e}$. Then, by maximality, $3K'_e$ contains some $R_{s,e}$, for $s \in \mathbf{T}$. Let s' be the tile $s < s' < \boldsymbol{\omega}_{\mathbf{T}} \times R_{\mathbf{T}}$ for which $R_{s'} = K'_e \times R_{\mathbf{T},e\perp}$. We have $G_K \subset v^{-1}(\boldsymbol{\omega}_{s'1})$. Also, since $\mathrm{dense}(\mathbf{T}) \leq \delta$, we deduce the important inequality that

$$|G_K| \leq |K \cap v^{-1}(\boldsymbol{\omega}_{s'})| \lesssim \delta |K|.$$

If $T(K,+)$ is a 2–tree, then the sets $\{\boldsymbol{\omega}_{s2} : s \in \mathbf{T}(K,+)\}$ are pairwise disjoint, so that the set $v^{-1}(\boldsymbol{\omega}_{s2})$ are either equal or disjoint. Hence

$$\mathbf{1}_K |A^{\mathbf{T}(K,+)}| \lesssim \sigma \mathbf{1}_{G_K}.$$

Our desired bound $\int_K |A^{\mathbf{T}(K,+)}| \, dx \lesssim \sigma\delta |K|$ is immediate.

If $\mathbf{T}(K,+)$ is a 1–tree, then all of the intervals $\boldsymbol{\omega}_{s2}$ contain $\boldsymbol{\omega}_{\mathbf{T}}$. Thus, for each x, there are $\varepsilon_{\pm}(x)$ so that $v(x) \in \boldsymbol{\omega}_{s2}$ iff $\varepsilon_-(x) \leq \mathrm{scl}(s) \leq \varepsilon_+(x)$. In particular, if $v(x) \in \boldsymbol{\omega}_{s2}$, then we have $|v(x) - e| \lesssim \varepsilon_-(x)/\mathsf{ann}$. This permits us to argue that the vector field $v(x)$ can be assumed to be constant. Specifically one sees the inequality

$$\mathbf{1}_{\boldsymbol{\omega}_{s2}}(v(x)) \int_{\mathbb{R}} |\varphi_s(x - yv(x)) - \varphi_s(x - ye)||\psi_s(y)| \, dy \lesssim \frac{\varepsilon_-(x)}{\mathrm{scl}(s)} \chi_{R_s}^{(2)}(x).$$

This inequality is in fact a straightforward calculation, but one that depends very much on the uniform assumptions (2.17).

We make more definitions, and in the first definition, replace the vector field by e:

$$\phi'_s(x) := \int \varphi_s(x - ye)\psi_s(y)\, dy,$$

$$A_K := \sum_{s \in \mathbf{T}(K,+)} \langle f, \varphi_s \rangle \mathbf{1}_{\boldsymbol{\omega}_{s2}}(v(x))\phi'_s, \qquad K \in \mathcal{K},$$

$$B := \sum_{s \in \mathbf{T}} \langle f, \varphi_s \rangle \phi'_s.$$

In the last line, it is a matter of convenience to assume that all trees $\mathbf{T}(K,+)$ are 1–trees, and that $\mathbf{T} = \bigcup_{K \in \mathcal{K}} \mathbf{T}(K,+)$. We have, by a straightforward estimate,

$$\left| A^{\mathbf{T}(K,+)} - A \right| \lesssim \sigma \mathbf{1}_{G_K},$$

so that it suffices to estimate the L^1 norm of A on the interval K.

There are two points. The first is that for each K, A_K is dominated by a maximal function applied to B. The second is that B obeys the inequality

$$(6.8) \qquad \|B\|_2 \lesssim \sigma |R_{\mathbf{T}}|^{1/2}.$$

Let us turn to the second claim. The claim will follow immediately if we can see that we have the universal estimate

$$(6.9) \qquad \Big\| \sum_{s \in \mathbf{T}} a_s \phi'_s \Big\|_2 \lesssim \Big[\sum_{s \in \mathbf{T}} |a_s|^2 \Big]^{1/2}.$$

This holds for all numerical sequences $\{a_s\}$.

Indeed, the assumptions that we have placed on the function φ, namely that it have Fourier support in a small ball around the origin, and the uniform assumption on the Fourier support of the functions ψ_s, as expressed in (2.16), imply that the functions ϕ'_s are supported in ω_s.[1] As these rectangles are either equal or disjoint, it suffices to prove the estimate (6.9) when the tree $\mathbf{T}$ has $\omega_s = \omega_{s'}$ for all s, $s' \in \mathbf{T}$. With this assumption, the assumption (2.17) (with the value of κ suitably small) implies the estimate

$$|\phi'_s| \lesssim \chi^{(2)}_{R_s}.$$

This is more than enough to prove (6.9) under this restrictive assumption on the tree.

Let us dominate A_K. For any choice of $\varepsilon_- < \varepsilon_+$, we have

$$\mathbf{1}_{\{\varepsilon_- < \mathrm{scl}(s) < \varepsilon_+\}} \phi'_s = (\zeta_{\varepsilon_-} - \zeta_{\varepsilon_+}) * \phi'_s$$

in which we take ζ to be a non–negative Schwartz function on the plane satisfying $\mathbf{1}_{[-1/2,1/2]^2} \leq \widehat{\zeta} \leq \mathbf{1}_{[-5/8,5/8]}$ and set $\zeta_\varepsilon(x_e, x_{e_\perp}) = \varepsilon\mathsf{ann}\zeta(\varepsilon x_e, \mathsf{ann} x_{e_\perp})$. The identity follows from the frequency properties of φ and of the class of functions ψ_t as described in (2.16). From this, we conclude that

$$|A_K| \lesssim \sup_{J \supset R} -\!\!\!\!\!\int_J |B(z)|\, dz,$$

---

[1]This calculation is available to us as we have replaced the vector field in the definition of ϕ'_s by a fixed vector e.

with the last quantity being a supremum over all rectangles J that contain K. This supremum is constant on K. Thus,

$$\int_{G_K} |A_K| \, dx \lesssim |G_K| \inf_{x \in K} \mathrm{M} \, B(x) \lesssim \delta |K| \inf_{x \in K} \mathrm{M} \, B(x),$$

where M denotes the strong maximal function with respect to the $(e, e_\perp)$ coordinates.

This last estimate is to be summed over $K \subset 4R_{\mathbf{T}}$:

$$\delta \sum_{K \subset 4R_{\mathbf{T}}} |K| \inf_{x \in K} \mathrm{M} \, B(x) \lesssim \delta \int_{4R_{\mathbf{T}}} \mathrm{M} \, B \, dx \lesssim \delta |R_{\mathbf{T}}|^{1/2} \|\mathrm{M} \, B\|_2 \lesssim \delta \sigma |R_{\mathbf{T}}|$$

This completes the the proof of the tree lemma.

References

[1] Barrionuevo, Jose, Lacey, Michael T., *A weak-type orthogonality principle*, Proc. Amer. Math. Soc., 131, 2003, 6, 1763–1769. MR1955263 (2004f:42025)

[2] Bourgain, J., *A remark on the maximal function associated to an analytic vector field*, Analysis at Urbana, Vol. I (Urbana, IL, 1986–1987), London Math. Soc. Lecture Note Ser., 137, 111–132, Cambridge Univ. Press, Cambridge, 1989. MR1009171 (90h:42028)

[3] Bourgain, J., *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. Funct. Anal., 1, 1991, 2, 147–187. MR1097257 (92g:42010)

[4] Carleson, Lennart, *On convergence and growth of partial sums of Fourier series*, Acta Math., 116, 1966, 135–157. MR0199631 (33:7774)

[5] Christ, Michael, *Personal communication.*

[6] Fefferman, Charles, *Pointwise convergence of Fourier series*, Ann. of Math. (2), 98, 1973, 551–571. MR0340926 (49:5676)

[7] Grafakos, Loukas, Li, Xiaochun, *Uniform bounds for the bilinear Hilbert transform, I*, Ann. of Math., to appear.

[8] Grafakos, Loukas, Tao, Terence, Terwilleger, Erin, L^p *bounds for a maximal dyadic sum operator*, Math. Z., 246, 2004, 1-2, 321–337. MR2031458 (2005a:42003)

[9] Katz, Nets Hawk, *Maximal operators over arbitrary sets of directions*, Duke Math. J., 97, 1999, 1, 67–79. MR1681088 (2000a:42036)

[10] Lacey, Michael, *Carleson's Theorem: Proof, Complements, Variation*, Pub. Mat., 48, 2004, no. 2, 251–307. MR2091007 (2005e:42020)

[11] Lacey, Michael, Li, Xiaochun, *On the Hilbert Transform on $C^{1+\epsilon}$ families of lines*, preprint, 2003.

[12] Lacey, Michael, Thiele, Christoph, *A proof of boundedness of the Carleson operator*, Math. Res. Lett., 7, 2000, 4, 361–370. MR1783613 (2001m:42009)

[13] Nagel, Alexander, Stein, Elias M., Wainger, Stephen, *Hilbert transforms and maximal functions related to variable curves*, Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, Proc. Sympos. Pure Math., XXXV, Part, 95–98, Amer. Math. Soc., Providence, R.I., 1979. MR0545242 (81a:42027)

[14] Stein, E. M., *Problems in harmonic analysis related to curvature and oscillatory integrals*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986), 196–221, Amer. Math. Soc., Providence, RI, 1987. MR0934224 (89d:42028)

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332
E-mail address: `lacey@math.gatech.edu`

Department of Mathematics, University of California–Los Angeles, Los Angeles, California 90055-1555
E-mail address: `xcli@math.ucla.edu`
Current address: School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540