Maximal Theorems
for the Directional Hilbert Transform
on the Plane

Michael T. Lacey
Xiaochun Li

Vienna, Preprint ESI 1400 (2003) October 24, 2003

Supported by the Austrian Federal Ministry of Education, Science and Culture
Available via anonymous ftp from FTP.ESI.AC.AT
or via WWW, URL: http://www.esi.ac.at
Maximal Theorems for the Directional Hilbert Transform on the Plane

Michael T. Lacey*  
Georgia Institute of Technology

Xiaochun Li†  
University of California, Los Angeles

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

\[
H_v f(x) = \text{p.v.} \int_{\mathbb{R}} f(x - vy) \frac{dy}{y}
\]

Let \( \zeta \) be a Schwartz function with frequency support in the annulus \( 1 \leq |\xi| \leq 2 \), and \( \zeta f = \zeta \ast f \). We prove that the maximal operator \( \sup_{|v|=1}|H_v \zeta f| \) 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.

1 Introduction, Principal Theorem

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

\[
H_v f(x) = \text{p.v.} \int_{\mathbb{R}} f(x - vy) \frac{dy}{y}
\]

This definition is independent of the length of \( v \), and below we shall only concern ourselves with \( |v| = 1 \). Let \( \zeta \) be a Schwartz function with frequency support in the annulus \( 1 \leq |\xi| \leq 2 \), and \( \zeta f = \zeta \ast f \). Our Theorem is

1.1 Theorem. The maximal operator \( H_v^* f := \sup_{|v|=1}|H_v \zeta f| \) 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 for the directional maximal function, namely

\[
M^* f(x) := \sup_{|v|=1} \sup_{t>0} (2t)^{-1} \int_{-t}^{t} |\zeta f(x - yv)| \, dy
\]

*Research supported in part by an NSF grant. Parts of this research were completed while on research stays at the Universite d’Paris-Sud, Orsay, Erwin Schrödinger Institute, Vienna, Austria.

†Research supported in part by an NSF Grant.
For both this operator, and $H^*$, the estimate of weak square integrability is sharp, as was pointed out to us by M. Christ, [3]. This argument may be summarized as follows. Begin with 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 $|x_2| < \frac{1}{100} |x_1|$, we have $H^* f(x) \simeq M^* f(x) \simeq |x|^{-1}$. That is, there will be no cancellation in computing either maximal function. And $|x|^{-1}$ is just in weak $L^2$.

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, [2], but also see [4]. As such, we use a method which is adopted from the proof of Carleson’s Theorem given by M. Lacey and C. Thiele [7].

In comparison to the proof of Lacey and Thiele, we find that the notion of a tile requires some care to define, and a new ingredient is needed in the proof of the “size Lemma,” an orthogonality statement. In all other aspects, the style of proof is very similar to that of [7].

Finally, there is an outstanding question, attributable to E. M. Stein [8], 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

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

is a bounded operator on say $L^2$. For positive results for analytic and real analytic vector fields, see [1, 8]. In a subsequent paper, the authors [6] will prove that the operator above is bounded on $L^2$ if $v \in C^{1+\epsilon}$, for any positive $\epsilon$. The results of this paper are a crucial aspect of the proof of this result.

## 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 $K$. And $A \simeq B$ iff $A \lesssim B$ and $B \lesssim A$. We use the notation $\mathbb{1}_A$ to denote the indicator function of the set $A$. And the Fourier transform on $\mathbb{R}^2$ is denoted by $\hat{f}(\xi) = \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} f(x) \, dx$, with a similar definition on the real line. We use the notation

$$\int_A f \, dx := |A|^{-1} \int_A f \, dx.$$

### 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.
Let $\rho$ 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 $\omega$ 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 $\omega$ are said to be $2^l \times a$. Notice 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 $\omega$ is $\text{scl}(\omega) := 2^l$ and that the annular parameter of $\omega$ is $\text{ann}(\omega) := a$, $\text{ann}(\omega) = 2^l$. In referring to the coordinate axes of an annular rectangle, we shall always mean $(e, e_\perp)$ as above.

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 $\omega$ 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 $\omega = r_1 \times r_2$ for intervals $r_i, r_i, i = 1, 2$. Moreover, $1 \leq |r_i| \cdot |r_i| \leq 2$ 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 $AT$ which satisfy these conditions. For $s, s' \in AT$ we write $s = R_s \times \omega_s$, and require that

1. $\omega_s$ is an annular rectangle,
2. $R_s$ and $\omega_s$ are dual rectangles,
3. $\{ R : R \times \omega_s \in AT \}$ partitions $\mathbb{R}^2$, for all $\omega_s$.
4. $\text{ann}(\omega_s) = 2^l$ for some fixed $a > 0$,
5. $\# \{ \omega_s : \text{scl}(s) = 2^l, \text{ann}(s) = \text{ann} \} \geq c \text{ann} 2^{-l}, \quad l \in \mathbb{Z}$.
Figure 2: An annular rectangle $\omega_s$. 

We assume that there are auxiliary sets $\omega_s, \omega_{s1}, \omega_{s2} \subset \mathbb{T}$ associated to $s$—or more specifically $\omega_s$—which satisfy these properties.

(2.9) \[ \Omega := \{ \omega_s, \omega_{s1}, \omega_{s2} : s \in \mathcal{A}\mathcal{T} \} \text{ is a grid in } \mathbb{T}, \]
(2.10) \[ \omega_{s1} \cap \omega_{s2} = \emptyset, \quad \omega_s := \text{hull}(\omega_{s1}, \omega_{s2}), \]
(2.11) \[ \omega_{s1} \text{ lies clockwise from } \omega_{s2} \text{ on } \mathbb{T}, \]
(2.12) \[ |\omega_s| \leq K \frac{\text{scl}(\omega_s)}{\text{ann}(\omega_s)}, \]
(2.13) \[ \{ \frac{\xi}{|\mathbb{T}|} : \xi \in \omega_s \} \subset \rho \omega_{s1}, \]
(2.14) \[ \text{to each } \omega_s \text{ there is a } \overline{\omega}_s \in \Omega \text{ such that } 2\omega_s \subset \overline{\omega}_s \subset 4\omega_s. \]

Recall that $\rho$ is the rotation that takes $e$ into $e_{\perp}$. Thus, $e_{\omega_s} \in \omega_{s1}$. See figure 1 which depicts the two rectangles that make up a phase rectangle $s \in \mathcal{A}\mathcal{T}$.

Note that $|\omega_s| \geq |\omega_{s1}| \geq \text{scl}(\omega_s)/\text{ann}(\omega_s)$. Thus, $e_{\omega_s}$ is in $\omega_{s1}$, and $\omega_s$ 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 $e' \in \omega_s$ 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 $\mathcal{A}\mathcal{T} (\text{ann})$ satisfying the conditions (2.4)—(2.13) above. The constant $a > 0$ appears in (2.7) and (2.8). We extend the definition
of $e_\omega$, $e_{\omega\perp}$, $\text{ann}(\omega)$ and $\text{scl}(\omega)$ to annular tiles in the obvious way, using the notation $e_s$, $e_{s\perp}$, $\text{ann}(s)$ and $\text{scl}(s)$.

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

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

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

$$\text{Dil}^p_{R_1 \times R_2} f(x_1, x_2) := \frac{1}{(|R_1||R_2|)^{1/p}} f\left(\frac{x_1}{|R_1|}, \frac{x_2}{|R_2|}\right), \quad 0 < p < \infty, \quad \text{(Dilation)}.$$ 

In the last display, $R_1 \times R_2$ 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 $L^p$ norm.

For a function $\varphi$ and tile $s \in \mathcal{A}T$ set

$$\varphi_s := \text{Mod}_{c(J)} \text{Dil}^2_{R_s} \text{Tr}_{c(R_s)} \varphi,$$

where $c(J)$ is the center of $J$. Below, we shall consider $\varphi$ 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 $\psi : \mathbb{R}^2 \to \{|x| = 1\}$ that achieves the maximum in our operator. Fix a Schwartz function $\psi$ on $\mathbb{R}$ with frequency support in a small neighborhood of 1. Then, define

$$(2.15) \quad \phi_s(x) := \int_{\mathbb{R}} \varphi_s(x - y\psi(x)) \text{scl}(s) \psi(\text{scl}(s)y) \, dy$$

$$= 1_{\omega_{s2}}(v(x)) \int_{\mathbb{R}} \varphi_s(x - y\psi(x)) \text{scl}(s) \psi(\text{scl}(s)y) \, dy.$$

An essential feature of this definition is that the support of the integral is contained in the set $\{v(x) \in \omega_{s2}\}$, a fact which can be routinely verified. That is, we can insert the indicator $1_{\omega_{s2}}(v(x))$ without loss of generality. The set $\omega_{s2}$ serves to localize the vector field, while $\omega_{s1}$ serves to identify the location of $\varphi_s$ in the frequency coordinate.

Remark: It seems likely that we could prove our theorem using the simpler definition $\phi_s = (1_{\omega_{s2}} \circ v) \varphi_s$. This would be the natural analog of the decomposition used by Lacey and Thiele [7]. But, in our subsequent paper on this subject [6], we will need to consider a truncation of the Hilbert transform, which suggests the definition of $\phi_s$ 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

$$C_{\text{ann},v} f := \sum_{s \in \mathcal{A}T(\text{ann})} \langle f, \varphi_s \rangle \phi_s, \quad j \in \mathbb{Z}. $$
In this display, $\mathcal{AT}(\text{ann}) := \{s \in \mathcal{AT} : \text{ann}(s) = \text{ann}\}$.

### 2.16 Lemma

Assume that the vector field is only measurable. The operator $\mathcal{C}_{\text{ann},v}$ 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 $\text{ann}$.

We shall prove that $\mathcal{C}_{\text{ann},v}$ maps $L^2$ into weak $L^2$, with constant independent of $\text{ann}$ 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

\begin{equation}
|\langle \mathcal{C}_{\text{ann},v}f, 1_F \rangle| \leq \sum_{s \in \mathcal{AT}(\text{ann})} |\langle f, \varphi_s \rangle \langle \phi_s, 1_F \rangle| \lesssim |F|^{1/2}.
\end{equation}

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

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

\begin{equation}
|\langle \mathcal{C}_{\text{ann},v}1_E, 1_F \rangle| \lesssim \lambda^{-p}|E|, \quad 2 < p < \infty.
\end{equation}

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

\begin{equation}
|\langle \mathcal{C}_{\text{ann},v}1_E, 1_F \rangle| \leq \sum_{s \in \mathcal{AT}(\text{ann})} |\langle \varphi_s, 1_E \rangle \langle \phi_s, 1_F \rangle| \\
\lesssim |F|\{\log(|E|/|F|)\}, \quad 9|F| < |E|.
\end{equation}

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

The proofs of (2.17) and (2.18) are given in Section 4.

### 3 Proof of Theorem 1.1

We show that Lemma 2.16 implies our main result, Theorem 1.1. Let $\kappa$ be a small but absolute constant, which depends upon the exact choice of $\varphi$. $\lambda$ be a smooth radial function satisfying

\begin{equation}
1_{\left[\frac{3}{2} - \kappa, \frac{3}{2} + \kappa\right]}(|\xi|) \leq \hat{\lambda} \leq 1_{\left[\frac{3}{2} - 2\kappa, \frac{3}{2} + 2\kappa\right]}(|\xi|)
\end{equation}

Let $\lambda_t(y) = t^{-2}\lambda(y/t)$.

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

\[ \sum_{j=-\infty}^{\infty} 2^j \psi(2^j y) \]
Recall that $\psi \geq 0$ is 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 positive 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
\[
(3.2) \quad T_v f(x) = \int \lambda_{\text{ann}} \ast f(x - yv(x)) K(y) \, dy
\]
in which $\text{ann} > 0$, and the measurable $v(x)$ are arbitrary. We shall do so by arguing that these operators are appropriate limiting averages of the operators that occur in Lemma 2.16.

For values of $2^j < \text{ann}$, let
\[
\tilde{S}_{j, \text{ann}} f = \sum_{s \in A T(\text{ann}) \atop \text{sc}(s) = 2^j} (f, \varphi_s) \varphi_s
\]
\[
T_{j,v} f(x) = \int f(x - yv(x)) 2^j \psi(2^j y) \, dy
\]
Also, let $\text{Rot}_\tau$ be the operation of rotation by angle $\tau$. The main point concerns the operator
\[
(3.3) \quad S_{j, \text{ann}} := \lim_{Y \to \infty} \int_{\text{Box}(Y)} \text{Rot}_{-\tau} \text{Tran}_{-y} \tilde{S}_{j, \text{ann}} \text{Tran}_y \text{Rot}_{\tau} \, dy d\tau
\]
where $\text{Box}(Y) = [0, Y]^2 \times [0, 2\pi]$. The limit, applied to a Schwartz function $f$ is seen to exist.

3.4 Lemma. For each $2^j < \text{ann}$, we have the identity
\[
S_{j, \text{ann}} \lambda_{\text{ann}} \ast f = c(j, \text{ann}) \lambda_{\text{ann}} \ast f
\]
where the constant $c(j, \text{ann})$ satisfies $c^{-1} < |c(j, \text{ann})| < c$, for some absolute constant $c$.

To deduce bounds for (3.2), observe that Lemma 2.16 concerns norm bounds for the sums
\[
\sum_j c(j, \text{ann})^{-1} T_{j,v} \tilde{S}_{j, \text{ann}} \lambda_{\text{ann}} \ast f
\]
Of course the coefficients $c(j, \text{ann})^{-1}$ do not appear in Lemma 2.16. Yet the placement of the absolute values in (2.17) and (2.18) 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 (3.3). Hence, by the Lemma just stated, we deduce the norm bounds for the operators in (3.2). It remains to prove our Lemma.

We record a simple lemma on convolutions.

3.5 Lemma. Let $\varphi$ and $\phi$ be real valued Schwartz functions on $\mathbb{R}^2$. Then,

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

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

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

The proof is immediate. The integral in question is

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

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

Proof of Lemma 3.4. Fix $\omega \in \Omega$, and let

$$\tilde{S}_\omega f := \sum_{s \in \mathcal{A}T(\text{ann})} \langle f, \varphi_s \rangle \varphi_s$$

Recall (2.6). It then follows from Lemma 3.5 that

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

is convolution with respect to $\Phi_\omega$, where $\hat{\Phi}_\omega = \hat{\varphi}_\omega \overline{\hat{\varphi}_\omega}$.

By choice of $\varphi$, it has non-zero Fourier transform, and is identically one on a small ball around the origin. Hence $\Phi_\omega$ has Fourier transform with the same properties. Recall (2.8), which states that we have an essentially maximal number of $\omega_s$ of a given scale. We see that

$$S_{j,\text{ann}} f = c(j, \text{ann}) \Phi \ast f$$

where $\Phi = \int_{[0,2\pi]} \text{Rot}_\tau \Phi_\omega \, d\tau$. And the absolute lower bound on $|c(j, \text{ann})|$ follows from the bound in (2.8). This function $\Phi$ has Fourier transform that is identically 1 in a small neighborhood of $|\xi| = \frac{3}{2}$. This permits us an absolute choice of $\kappa$, as in (3.1), and so completes our proof.
4 The Main Lemmas

We need these notions associated to tiles and sets of tiles. There is a natural partial order on tiles given by \( s < s' \) iff \( \omega_s \supset \omega_{s'} \), \( R_{s1} \subset R_{s'} \), and \( R_{s2} = R_{s'2} \). We are free to restrict attention to a set of tiles for which we have the conclusion

\[
\text{(4.1)} \quad \text{If } \omega_s \times R_s \cap \omega_{s'} \times R_{s'} \neq \emptyset, \text{ then } s \text{ and } s' \text{ are comparable under } \overset{\sim}{<}.
\]

A tree is a collection of tiles \( T \subset \mathcal{AT}(\text{ann}) \), for which there is a (non–unique) tile \( \omega_T \times R_T \in \mathcal{AT}(\text{ann}) \) with \( s < \omega_T \times R_T \) for all \( s \in T \). For \( j = 1, 2 \), call \( T \) a \( j \)-tree if the tiles \( \{ \omega_{s_i} \times R_s : s \in T \} \) are pairwise disjoint.

Fix a positive rapidly decreasing function \( \chi \), and for rectangle \( R \) set

\[
\text{(4.2)} \quad \chi_R = D_1 \chi_{c(R)}
\]

\[
\text{dense}(s) := \sup_{s < s'} \int v^{-1}(\omega_{s'}) \chi_{R_{s'}} \, dx,
\]

\[
\text{dense}(S) := \sup_{s \in S} \text{dense}(s),
\]

\[
\text{sh}(S) := \bigcup_{s \in S} R_s \quad \text{(the shadow of } S \text{)}
\]

\[
\Delta(T)^2 := \sum_{s \in T} \frac{|(f, \varphi_s)|^2}{|R_s|} 1_{R_s}, \quad T \text{ is a 1–tree},
\]

\[
\text{size}(S) := \sup_{T \text{ is a 1–tree}} \left[ ||\text{sh}(T)||^{-1} \sum_{s \in T} \left| (f, \varphi_s) \right|^2 \right]^{1/2}
\]

Recall that \( ||f||_2 \simeq 1, |F| \leq 2 \), and \( v \) is defined only on \( F \).

4.3 Lemma. Any \( S \subset \mathcal{AT}(\text{ann}) \) is the union of \( S_{\text{heavy}} \) and \( S_{\text{light}} \) satisfying these conditions.

\[
\text{dense}(S_{\text{light}}) \leq \frac{1}{2} \text{dense}(S).
\]

The collection \( S_{\text{heavy}} \) is a union of trees \( T \in T_{\text{heavy}} \), with

\[
\text{(4.4)} \quad \sum_{T \in T_{\text{heavy}}} \text{sh}(T) \lesssim \text{dense}(S)^{-1} |F|.
\]

4.5 Lemma. Any \( S \subset \mathcal{AT}(\text{ann}) \) is the union of \( S_{\text{big}} \) and \( S_{\text{small}} \) satisfying these conditions.

\[
\text{size}(S_{\text{small}}) \leq \frac{1}{2} \text{size}(S).
\]

The collection \( S_{\text{big}} \) is a union of trees \( T \in T_{\text{big}} \), with

\[
\text{(4.6)} \quad \sum_{T \in T_{\text{big}}} \text{sh}(T) \lesssim \text{size}(S)^{-2}.
\]

We need a lemma that relates the notions of density, size and tree.
4.7 Lemma. For any tree $T$ we have the estimate

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

The first lemma has a proof which is essentially identical to the proof of the “mass lemma” in M. Lacey and C. Thiele [7]. 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 [7]. We shall give a proof.

The lemmas are combined in this way to prove (2.17), and hence Lemma 2.16 in the case of the weak $L^2$ estimate. Lemma 4.3 and Lemma 4.5 should be applied so that their principal estimates (4.4) and (4.6) are approximately equal. The density of $A\mathcal{T}(\text{ann})$ is at most a constant. The size of $A\mathcal{T}(\text{ann})$ is at most a constant times the $L^\infty$ norm of $f$. Thus, we can take the set of all tiles $A\mathcal{T}(\text{ann})$ and decompose it into subcollections $S_\sigma$, $\sigma \in \{2^n : n \in \mathbb{Z}\}$, so that $S_\sigma$ is the union of trees $T \in T_\sigma$ such that

$$\sum_{T \in T_\sigma} |R_T| \lesssim \sigma,$$

$$\text{dense}(S_\sigma) \lesssim \min(1, \sigma^{-1}),$$

$$\text{size}(S_\sigma) \lesssim \sigma^{-1/2}.$$ 

Hence, it follows that

$$\sum_{s \in S_\sigma} |\langle f, \varphi_s \rangle \langle \phi_s, 1_E \rangle| \lesssim \sum_{T \in T} \min(1, \sigma^{-1}) \sigma^{-1/2} |R_T|$$

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

This estimate is summable over $\sigma \in \{2^n : n \in \mathbb{Z}\}$ to an absolute constant. This completes the proof of (2.17).

A small variation on this argument proves (2.18). In this instance, the function $f = 1_E$, with $|E| \simeq 1$, so that the size of $A\mathcal{T}(\text{ann})$ is $\lesssim 1$. And $A\mathcal{T}(\text{ann})$ is a union of subcollections $S_\sigma$, $\sigma \in \{2^n : n \in \mathbb{N}\}$, so that $S_\sigma$ is the union of trees $T \in T_\sigma$ satisfying

$$\sum_{T \in T_\sigma} |R_T| \lesssim \sigma |F|,$$

$$\text{dense}(S_\sigma) \lesssim \sigma^{-1},$$

$$\text{size}(S_\sigma) \lesssim \min(1, (\sigma |F|)^{-1/2}).$$

Hence, it follows that

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

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 (2.18). Our proof of Lemma 2.16 is complete, aside from the proofs of the three Lemmas of this section.
The Size Lemma: Orthogonality

We give the proof of Lemma 4.5. 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 [7]. 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.

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

$$T_+ := \emptyset, \quad T_{big} = \emptyset, \quad S^{stock} = S.$$  

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

$$\sum_{s \in T} \langle f, \varphi_s \rangle^2 \geq \frac{\sigma^2}{4} |R_T|,$$

$|R_T|$ is maximal, and $\omega_T$ is most anti–clockwise. Then set $\tau(T) \subset S^{stock}$ to be the maximal (with respect to inclusion) tree in $S^{stock}$ with top $\omega_T \times R_T$. Update

$$T_+ = T_+ \cup \{T\}, \quad T_{big} = T_{big} \cup \{\tau(T)\}, \quad S^{stock} = S^{stock} - T.$$  

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

To conclude the Lemma, we need to show that

$$\sum_{T \in T_+} |R_T| \lesssim \sigma^{-2}.$$  

We have constructed these trees so that they satisfy a property very useful to the verification of this inequality. Suppose $T \neq T' \in T_+$ and $s \in T$ and $s' \in T'$ with $\omega_s \subsetneq \omega_{s'}$. Then $R_T \cap R_{s'} = \emptyset$. We refer to this property as "strong disjointness." To see that it holds, we have $\omega_T \subset \omega_{s'}$, so that $\omega_T$ lies counterclockwise with respect to $\omega_{T'}$. Hence $T$ was constructed before $T'$. Thus, if $R_T \cap R_{s'} \neq \emptyset$, then $s' \in \tau(T)$, and so $s'$ would have been removed from $S^{stock}$ and so could not be in $T'$.

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

$$\sigma^2 \sum_{T \in T_+} |R_T| \lesssim \sum_{s \in S_+} \langle f, \varphi_s \rangle \langle \varphi_s, f \rangle$$

$$= \langle f, F(S_+) \rangle$$

$$\leq \|f\|_2 \|F(S_+)\|_2$$

This follows by Cauchy–Schwarz. Recall that $\|f\|_2 = 1$. To conclude (5.2) we need only show that

$$\|F(S_+)\|_2 \lesssim \sigma \left( \sum_{T \in T_+} |R_T| \right)^{1/2}.$$  

11
For \( s \in T \), set
\[
\begin{align*}
\mathbf{B}_e(s) &= \{ s' \in \mathbf{S}_+ - T : \omega_s = \omega_{s'} \}, \\
\mathbf{B}(s) &= \{ s' \in \mathbf{S}_+ - T : \omega_s \not\subset \omega_{s'} \}.
\end{align*}
\]

Note that if \( s' \in \mathbf{S}_+ - T \) is such that \( \langle \varphi_s, \varphi_{s'} \rangle \neq 0 \), then \( s' \in \mathbf{B}_e(s) \cup \mathbf{B}(s) \). And if \( s' \not\in \mathbf{B}(s) \), then \( \langle \varphi_s, \varphi_{s'} \rangle = 0 \). We estimate
\[
\begin{align*}
\| F(\mathbf{S}_+) \|_2^2 &\lesssim \sum_{T \in \mathcal{I}_+} \| F(T) \|_2^2 \\
&+ \sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{B}_e(s)) \rangle \\
&+ \sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{B}(s)) \rangle.
\end{align*}
\]

It is a routine matter to verify that for each 1–tree \( T \in \mathcal{I}_+ \),
\[
\| F(T) \|_2 \lesssim \sigma |R_T|^{1/2}.
\]

Therefore, (5.4) is no more than
\[
\sum_{T \in \mathcal{I}_+} \| F(T) \|_2^2 \lesssim \sum_{T \in \mathcal{I}_+} \sigma^2 |R_T|.
\]

For the term (5.5), we use this estimate for tiles \( s, s' \) with \( \omega_s = \omega_{s'} \),
\[
| \langle \varphi_s, \varphi_{s'} \rangle | \lesssim \text{Dil}^0_{R_s} \text{Tran}_{c(R_s)} \chi(c(R_s)).
\]

This with Cauchy–Schwarz gives
\[
\begin{align*}
(5.5) \leq & \left[ \sum_{s \in \mathbf{S}_+} | \langle f, \varphi_s \rangle |^2 \sum_{s \in \mathbf{S}_+} | \langle \varphi_s, F(\mathbf{B}_e(s)) \rangle |^2 \right]^{1/2} \\
\lesssim & \sum_{s \in \mathbf{S}_+} | \langle f, \varphi_s \rangle |^2 \\
\lesssim & \sigma^2 \sum_{T \in \mathcal{I}_+} |R_T|,
\end{align*}
\]

as required by (5.3).

As for (5.6), use Cauchy–Schwarz in \( s \),
\[
\begin{align*}
\sum_{s \in \mathbf{S}_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{B}(s)) \rangle \leq & \left[ \sum_{s \in \mathbf{S}_+} | \langle f, \varphi_s \rangle |^2 \sum_{s \in \mathbf{S}_+} | \langle \varphi_s, F(\mathbf{B}(s)) \rangle |^2 \right]^{1/2} \\
\lesssim & \left[ \sigma^2 \sum_{T \in \mathcal{I}_+} |R_T| \sum_{s \in \mathbf{S}_+} | \langle \varphi_s, F(\mathbf{B}(s)) \rangle |^2 \right]^{1/2}.
\end{align*}
\]
We show that for each $T \in T_+$,

\[(5.7) \quad \sum_{s \in T} |\langle \varphi_s, F(B(s)) \rangle|^2 \lesssim \sigma^2 |R_T| \]

This will complete the proof of (5.3).

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

\[\langle \varphi_s, F(B(s)) \rangle = \langle \varphi_s, F(B) \rangle\]

Moreover, all tiles $s \in T \cup B$ have $\omega_s \supset \omega_T$. 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. 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 $B$. This yields the essential observation that the rectangles $\{R_s : s \in B\}$ are pairwise disjoint, and do not intersect $R_T$.

Make a further diagonalization of the set $B$. Set $B_1 = \{s \in B : R_s \subset 4R_T\}$, and for $k > 1$, set $B_k = \{s \in B : R_s \subset 4^k R_T, R_s \not\subset 4^{k-1} R_T\}$. Let us point out that

\[\|F(B_k)\|_2 \lesssim 4^k \sigma |R_T|^{1/2}.\]

Indeed, recalling the notation (4.2)

\[|F(B_k)| \lesssim \sigma \sum_{s \in B_k} \chi_{R_s} \ast 1_R.\]

Therefore, for any function $g$,

\[\langle F(B_k), g \rangle \lesssim \sigma \int \sum_{s \in B_k} 1_{R_s} \chi_{R_s} \ast g \, dx \]

\[\lesssim \sigma \int \sum_{s \in B_k} 1_{R_s} Mg \, dx \]

\[\lesssim \sigma 4^k |R_T|^{1/2} \|g\|_2\]

Clearly, this implies (5.7) for $B_1$.

For $k > 1$, we can strengthen our inequality to the following.

\[\|F(B_k)\|_{L^2(4^{k-1} R_T)} \lesssim 4^{-10k} |R_T|^{1/2}.\]

Then, certainly,

\[\sum_{s \in T} |\langle \varphi_s, F(B_k) 1_{4^{k-1} R_T} \rangle|^2 \lesssim 4^{-k} |R_T| |.\]

But the functions $\{\varphi_s : s \in T\}$ are highly concentrated in a neighborhood of $R_T$. In particular, for any function $g$,

\[\sum_{s \in T} |\langle \varphi_s 1_{(\mathbb{R}^2 - 4^{k-1} R_T)}, g \rangle|^2 \lesssim 4^{-k} \|g\|_2^2.\]

Clearly, this completes the proof of (5.7).
Proof of Lemma 4.7

We may fix the vector field, and assume that the standard basis \((e, e_\perp)\) are the basis for the rectangle \(R_T\). As a consequence, we can without loss of generality assume that this is the basis for all the tiles \(s \in T\), and we write \(R_s = R_{s,e} \times R_{s,e_\perp}\). Set \(\delta = \text{dense}(T)\) and \(\sigma = \text{size}(T)\).

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

Choose signs \(\varepsilon_s \in \{\pm\}\) such that

\[
\sum_{s \in T} |(f, \varphi_s)\langle \phi_s, 1_F \rangle| = \int_F \varepsilon_s \sum_{s \in T} (f, \varphi_s) \phi_s \, dx
\]

Let

\[
A^{S'} = \sum_{s \in S'} \varepsilon_s (f, \varphi_s) \phi_s.
\]

Estimate, without appeal to cancellation

\[
\int_K |A^{T(K, \pm)}| \, dx \leq \frac{1}{\sigma} \int_K \sum_{s \in T(K, \pm)} 1_{\omega_{s2}}(v(x)) \, dx
\]

The rectangles \(R_s\) are smaller than those of \(K\), and do not intersect it. Thus, the \(\chi^{(\infty)}_{R_s}(x) \lesssim 1\) decrease as \(\text{scl}(s)\) increases for \(x \in K\). The integral above is at most

\[
\lesssim \delta \min(|K|, |R_T|) \sup_{x \in K} \chi^{(\infty)}_{R_s}(x).
\]

This is summable over \(K \in K\) to at most \(\lesssim \delta |R_T|\).

If \(T(K, \pm)\) is non-empty, then \(K \subset 4R_T\). Set

\[
G_K := K \cap \bigcup_{s \in T(K, \pm)} v^{-1}(\omega_{s2})
\]

which contains the support of \(1_K A^{T(K, \pm)}\). Our assertion is that \(|G_K| \lesssim \delta |K|\). To see this, let \(K = K_e \times K_{e_\perp}\), and \(K'\) be the dyadic interval that contains \(K_e\) and \(|K'_e| = 2|K_e|\). This interval is small than that of \(R_{T,e}\). Then, by maximality, 3\(K'_e\) contains some \(R_{s,e}\) for \(s \in T\). Let \(s'\) be the tile \(s < s' < \omega_T \times R_T\) for which \(R_{s'} = K'_e \times R_{T,e_\perp}\). We have \(G_K \subset v^{-1}(\omega_{s'1})\), and since \(\text{dense}(T) \leq \delta\), we have

\[
|G_K| \leq |K \cap v^{-1}(\omega_{s'})| \lesssim \delta |K|.
\]

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

\[
1_K A^{T(K, \pm)} \lesssim \sigma 1_G.
\]

14
Our desired bound \( \int_{K} |A^{T(K, +)}| \, dx \lesssim \sigma \delta |K| \) is immediate.

If \( T(K, +) \) is a 1–tree, then all of the interval \( \omega_s \) contain \( \omega_T \). Thus, for each \( x \), there are \( \varepsilon_{\pm}(x) \) so that \( v(x) \in \omega_s \) iff \( \varepsilon_{\pm}(x) \leq \text{scl}(s) \leq \varepsilon_{\pm}(x) \). In particular, if \( v(x) \in \omega_s \), then we have \( |v(x) - e| \lesssim \varepsilon_{\pm}(x)/\text{ann} \). This permits us to argue that the vector field \( v(x) \) can be assumed to be constant. By way of a straightforward calculation, one sees that

\[
1_{\omega_s}(v(x)) \int_{\mathbb{R}} |a_s(x - yv(x)) - \varphi_s(x - ye)||\psi_s(y)| \, dy \lesssim \frac{\varepsilon_{\pm}(x)}{\text{scl}(s)} \lambda_R(x).
\]

So we set

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

\[
A = \sum_{s \in T(K, +)} \langle f, \alpha_s \rangle 1_{\omega_s}(v(x))\phi'_s,
\]

\[
B = \sum_{s \in T(K, +)} \langle f, \alpha_s \rangle \phi'_s.
\]

We have, by a straightforward estimate,

\[
|A^{T(K, +)} - A| \lesssim \sigma 1_{G_K}
\]

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

The main point is that \( A \) is dominated by a maximal function applied to \( B \). For any choice of \( \varepsilon < \varepsilon_+ \), we have

\[
1_{\varepsilon_+ < \text{scl}(s) < \varepsilon_+} \phi'_s = (\varepsilon_+ - \varepsilon_+) * \phi'_s
\]

in which we take \( \zeta \) to be non-negative Schwartz function on the plane satisfying \( 1_{[-1/2, 1/2]} \leq \zeta \leq 1_{[-5/8, 5/8]} \) and set \( \psi(x, e_{\perp}) = \varepsilon_{\text{ann}}(\varepsilon xe_{e}, \text{ann}xe_{\perp}) \). The identity follows from the frequency properties of \( \psi \) and \( \varphi \). From this, we conclude that

\[
|A| \lesssim \sup_{K \subset R} \int_{G_K} |B(z)| \, dz
\]

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

\[
\int_{G_K} |A| \, dx \lesssim |G_K| \inf_{x \in K} MB(x) \lesssim \delta |K| \inf_{x \in K} MB(x)
\]

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

This last estimate is to be summed over \( K \subset 4R_T \).

\[
\delta \sum_{K \subset 4R_T} |K| \inf_{x \in K} MB(x) \lesssim \delta \int_{4R_T} MB \, dx
\]

\[
\lesssim \delta |R_T|^{1/2} \|MB\|_2
\]

\[
\lesssim \delta \sigma |R_T|
\]

This completes our proof.
References

[1] J. Bourgain, *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., vol. 137, Cambridge Univ. Press, Cambridge, 1989, pp. 111–132. MR 90h:42028

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

[3] Michael Christ, *Personal Communication*.

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

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

[6] Michael Lacey and XiaoChun Li, *On the Hilbert Transform on \( C^{1+\varepsilon} \) families of lines* (2003).

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

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

Michael T. Lacey | XiaoChun Li
---|---
School of Mathematics | Department of Mathematics
Georgia Institute of Technology | University of California, Los Angeles
Atlanta GA 30332 | Los Angeles CA
lacey@math.gatech.edu | xcli@math.ucla.edu