A NEW WAY OF LOOKING AT DISTRIBUTIONAL ESTIMATES; APPLICATIONS FOR THE BILINEAR HILBERT TRANSFORM.

DMITRIY BILYK AND LOUKAS GRAFAKOS

Abstract. Distributional estimates for the Carleson operator acting on characteristic functions of measurable sets of finite measure were obtained by Hunt [12]. In this article we describe a simple method that yields such estimates for general operators acting on one or more functions. As an application we discuss how distributional estimates are obtained for the linear and bilinear Hilbert transform. These distributional estimates show that the square root of the bilinear Hilbert transform is exponentially integrable over compact sets. They also provide restricted type endpoint results on products of Lebesgue spaces where one exponent is 1 or the sum of the reciprocal of the exponents is 3/2. The proof of the distributional estimates for the bilinear Hilbert transform rely on an improved energy estimate for characteristic functions with respect to sets of tiles from which appropriate exceptional subsets have been removed.

1. Introduction

Let T be a linear or sublinear operator defined on a subspace of measurable functions on a measure space (X, μ) and taking values in the space of measurable functions on a measure space (Y, ν). We are interested in estimates of the form

$$\nu\big(\{x : \ |T(\chi_F)(x)| > \lambda\}\big) \le \mu(F)\,\varphi(\lambda) \tag{1}$$

where the function φ is decreasing. An important example is given when $\varphi(\lambda) = C\,\lambda^{-p}$ for some $0 < C, p < \infty$. Then T is said to be of restricted weak type (p, p). This means that T restricted to characteristic functions maps $L^p(\mu)$ to $L^{p,\infty}(\nu)$ (with norm at most $C^{1/p}$).

A single restricted weak type (p, p) estimate does not provide boundedness information beyond L^p. Knowledge of restricted weak type (p, p) estimates for two values of $p = p_1$ and $p = p_2$ yields a distributional estimate

$$\nu\big(\{x : \ |T(\chi_F)(x)| > \lambda\}\big) \le \mu(F)\ \min\left(C_1\,\lambda^{-p_1}, C_2\,\lambda^{-p_2}\right), \tag{2}$$

which captures the L^p boundedness of T on L^p for p between p_1 and p_2.

Suppose that T is of restricted weak type $(1, 1)$. Then we expect $\varphi(\lambda) = C\lambda^{-1}$ as $\lambda \to 0$. If T also happens to be bounded on L^p for all $1 < p < \infty$, then we expect that the function $\varphi(\lambda)$ has decay faster than any negative power of λ as $\lambda \to \infty$. We may therefore guess that $\varphi(\lambda) = Ce^{-c\lambda}$ as $\lambda \to \infty$ for some fixed $C, c > 0$. This is indeed the case with the classical Hilbert transform defined for functions g on the line by

$$H(g)(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|t|\ge\varepsilon} g(x-t)\frac{dt}{t}\,, \qquad x \in \mathbf{R},$$

2000 *Mathematics Subject Classification.* 46B70, 42B99.

Key words and phrases. multilinear operators, distributional estimates.

Both authors have been supported by the National Science Foundation under grant DMS 0400387.

which satisfies the distributional estimate

$$\big|\{x\in\mathbf{R}:\ |H(\chi_F)(x)|>\lambda\}\big|\le C\,|F|\begin{cases}\frac{1}{\lambda} & \text{for } \lambda<1,\\ \\ e^{-c\lambda} & \text{for } \lambda\ge 1,\end{cases} \tag{3}$$

for some constants $C, c > 0$. Here $|\ |$ denotes Lebesgue measure on the line. A proof of estimate will be given in section 3.

Good distributional estimates are also known for the Carleson-Hunt operator:

$$\mathcal{C}(f)(x)=\sup_{N>0}\bigg|\int_{-N}^{+N}\widehat{f}(\xi)e^{2\pi ix\xi}d\xi\bigg|$$

defined for functions f on the line. Then again the decreasing function φ is explicit. We have for some constants C_o, c_o:

$$\big|\{x\in\mathbf{R}:\ |\mathcal{C}(\chi_F)(x)|>\lambda\}\big|\le C_o\,|F|\begin{cases}\frac{1}{\lambda}\big(1+\log\frac{1}{\lambda}\big) & \text{for } \lambda<1\\ \\ e^{-c_o\lambda} & \text{for } \lambda\ge 1.\end{cases} \tag{4}$$

Estimate (4) was first obtained by Hunt [12] for the analogous operator on the circle using a variation of the method developed by Carleson [4] in his proof of the boundedness of the maximal partial sums of square integrable functions on the circle. Approximately 35 years later, estimate (4) was reproved by Grafakos, Tao, and Terwilleger [10] using time-frequency analysis via a refinement of the L^2 argument of Lacey and Thiele [16]. The latter was influenced by the work of Fefferman [5].

It is worth mentioning that extrapolation techniques by Antonov [1] show that estimate (4) implies the boundedness of $\mathcal{C}$ on $L\log L\log\log\log L$ of every compact set; this implies the almost everywhere convergence of the partial Fourier integrals of functions locally in this class. On this, the reader may also want to consult the work by Sjölin and Soria [18].

Our purpose in this article is to discuss a way of studying distributional estimates that does not require the estimation of the measure of a set. In the next section we formulate a proposition that shows the equivalence of the two approaches for linear and multilinear operators. As an application we derive distributional estimates for the linear and the bilinear Hilbert transforms.

The second author of this paper would like to thank T. Tao for discussing with him many of the ideas presented in this work. Both authors would like to thank the organizers of the 7th lnternational Conference on Harmonic Analysis and Partial Differential Equations at El Escorial for the wonderful research atmosphere and their warm hospitality during the meeting.

2. New look at distributional estimates

As in the previous section suppose that T is a complex-valued linear or sublinear operator defined on a subspace of measurable functions on a measure space (X,μ) and taking values in the space of measurable functions on a measure space (Y,ν). In the sequel φ^{-1} denotes the inverse of a strictly decreasing continuous function φ which vanishes at infinity.

Proposition 2.1. *Let T be as described previously, let φ be a (strictly) decreasing function on $(0,\infty)$, and let $A>0$.*

September 4, 2005.

(I) Suppose that for all pairs of finite measure subsets (E, F) of $Y \times X$, there is a subset E' of E such that

$$\nu(E') \geq \frac{1}{2}\nu(E) \qquad \textit{and} \qquad \left| \int_{E'} T(\chi_F)\, d\nu \right| \leq A\, \nu(E)\, \varphi^{-1}\Big(\frac{\nu(E)}{\mu(F)}\Big)\,. \tag{5}$$

Then for all $\lambda > 0$ we have

$$\nu\big(\{y \in Y:\ |T(\chi_F)(y)| > \lambda\}\big) \leq\ 4\, \mu(F)\, \varphi\Big(\frac{\lambda}{2\sqrt{2}A}\Big)\,. \tag{6}$$

(II) Conversely, suppose that for all finite measure subsets F of X the distributional estimate

$$\nu\big(\{y \in Y:\ |T(\chi_F)(y)| > \lambda\}\big) \leq\ \mu(F)\, \varphi\big(\lambda\big) \tag{7}$$

holds for all $\lambda > 0$. Then for all pairs of finite measure subsets (E, F) of $Y \times X$, there is a subset E' of E such that

$$\nu(E') \geq \frac{1}{2}\nu(E) \qquad \textit{and} \qquad \left| \int_{E'} T(\chi_F)\, d\nu \right| \leq \nu(E)\, \varphi^{-1}\Big(\frac{\nu(E)}{2\mu(F)}\Big)\,. \tag{8}$$

Proof. For simplicity, in the proof below we use the Lebesgue notation $|\ |$ to denote the measures μ and ν.

Let us first prove assertion (I). Take E to be one of the four sets $\{\mathrm{Re}\, T(\chi_F) > \frac{\lambda}{\sqrt{2}}\}$, $\{\mathrm{Re}\, T(\chi_F) < -\frac{\lambda}{\sqrt{2}}\}$, $E = \{\mathrm{Im}\, T(\chi_F) > \frac{\lambda}{\sqrt{2}}\}$, or $E = \{\mathrm{Im}\, T(\chi_F) < -\frac{\lambda}{\sqrt{2}}\}$. Then

$$\frac{\lambda|E|}{2\sqrt{2}} \leq \frac{\lambda|E'|}{\sqrt{2}} \leq \left| \int_{E'} T(\chi_F)\, dx \right| \leq A\, |E| \varphi^{-1}\Big(\frac{|E|}{|F|}\Big)$$

and solve for $|E|$ to obtain

$$|E| \leq |F|\, \varphi\Big(\frac{\lambda}{2\sqrt{2}A}\Big)\,.$$

The claimed distributional estimate (6) follows by summing over the four choices of E.

We now prove assertion (II). Given E and F define

$$E' = \{y \in E:\ |T(\chi_F)(y)| \leq c_1\}$$

for some $c_1 > 0$. Then $T(\chi_F)$ is bounded and therefore integrable on E' and it is a consequence of (7) that

$$|E \setminus E'| = |\{x \in E:\ |T(\chi_F)(x)| > c_1\}| \leq |F|\, \varphi(c_1).$$

Picking c_1 so that $\varphi(c_1) = \frac{|E|}{2|F|}$ we conclude that $|E'| \geq \frac{1}{2}|E|$ while the second inequality in (8) follows by passing the absolute value inside the integral and using the definition of c_1. □

We note that if T is a real-valued, then both constants $2\sqrt{2}$ and 4 in in (6) may be replaced by 2. We also notice that nowhere in the proof we used that T is a linear or sublinear operator. The only thing we used is that T is well-defined on characteristic functions.

Next we give a multivariable version of the previous proposition. Here we assume that T is a defined on a subspace of measurable functions on the product of measure spaces $(X_1, \mu_1) \times \cdots \times (X_m, \mu_m)$ that contains all m-tuples of characteristic functions of sets of sets of finite measure. We also assume that T takes values in the set of measurable functions of another measure space (Y, ν) and that it is complex-valued.

Proposition 2.2. *Let T be as above, φ be a decreasing function, and let $A, \gamma_1, \dots, \gamma_m > 0$. (I) Suppose that for all $m+1$-tuples of subsets $(E, F_1, \dots, F_m)$ of $Y \times X_1 \times \cdots \times X_m$ with finite measure, there is a subset E' of E such that*

$$\nu(E') \geq \frac{\nu(E)}{2} \text{ and } \left| \int_{E'} T(\chi_{F_1}, \dots, \chi_{F_m})\, d\nu \right| \leq A\nu(E)\, \varphi^{-1}\Big(\frac{\nu(E)}{\mu_1(F_1)^{\gamma_1} \dots \mu_m(F_m)^{\gamma_m}} \Big). \tag{9}$$

Then for all $\lambda > 0$ we have

$$\nu\big(\{y \in Y : |T(\chi_{F_1}, \dots, \chi_{F_m})(y)| > \lambda\}\big) \leq 4\, \mu_1(F_1)^{\gamma_1} \dots \mu_m(F_m)^{\gamma_m}\, \varphi\Big(\frac{\lambda}{2\sqrt{2}A} \Big). \tag{10}$$

(II) Conversely, suppose that for all finite measure subsets F of X the distributional estimate

$$\nu\big(\{y \in Y : |T(\chi_{F_1}, \dots, \chi_{F_m})(y)| > \lambda\}\big) \leq \mu_1(F_1)^{\gamma_1} \dots \mu_m(F_m)^{\gamma_m}\, \varphi\big(\lambda\big) \tag{11}$$

holds for all $\lambda > 0$. Then for all $m+1$-tuples of finite measure subsets $(E, F_1, \dots, F_m)$ of $Y \times X_1 \times \cdots \times X_m$, there is a subset E' of E such that

$$\nu(E') \geq \frac{\nu(E)}{2} \text{ and } \left| \int_{E'} T(\chi_{F_1}, \dots, \chi_{F_m})\, d\nu \right| \leq \nu(E)\, \varphi^{-1}\Big(\frac{\nu(E)}{2\mu_1(F_1)^{\gamma_1} \dots \mu_m(F_m)^{\gamma_m}} \Big). \tag{12}$$

The proof is the same as in the case $m = 1$ and is omitted.

We note that if T is a multilinear singular integral operator, then the exponents γ_j satisfy $\gamma_1 + \cdots + \gamma_m = 1$ by homogeneity.

3. Distributional estimates for the (linear) Hilbert transform

In this section, we prove the distributional estimate (3). In this case we take both spaces X and Y to be $\mathbf{R}$ and both measures μ and ν to be Lebesgue measure. We begin by noting that the inverse function of

$$\varphi(\lambda) = \begin{cases} \frac{1}{\lambda} & \text{for } \lambda < 1 \\ \\ e^{-c\lambda} & \text{for } \lambda \geq 1 \end{cases}$$

is the function

$$\varphi^{-1}(t) = c \min\Big(1, \frac{1}{t}\Big)\Big(1 + \log^+ \frac{1}{t}\Big), \qquad t > 0$$

for some constant $c > 0$. Using Proposition 2.1 we have that (3) is equivalent to the following:

There is a constant $C > 0$ such that given any pair (E, F) there is $E' \subsetneqq E$ with $|E'| \geq \frac{1}{2}|E|$ and

$$\left| \int_{E'} H(\chi_F)\, dx \right| \leq C|E| \min\Big(1, \frac{|F|}{|E|}\Big)\Big(1 + \log^+ \frac{|F|}{|E|}\Big). \tag{13}$$

We will obtain (13) as a consequence of the more general statement below:

Proposition 3.1. *Suppose that T is bounded from $L^2(X)$ to $L^2(Y)$ and T and T^* are of restricted weak type $(1, 1)$. Then the following is valid: For any pair (E, F) of sets of finite measure there is $E' \subsetneqq E$ with $|E'| \geq \frac{1}{2}|E|$ such that*

$$\left| \int_{E'} T(\chi_F)\, dx \right| \leq C|E| \min\Big(1, \frac{|F|}{|E|}\Big)\Big(1 + \log_2^+ \frac{|F|}{|E|}\Big). \tag{14}$$

Proof. Let c_0 be the norm of T from $L^1 \to L^{1,\infty}$ and c_0^* be the norm of T^* from $L^1 \to L^{1,\infty}$. The proof uses an iteration argument.

Case 1: $|E| \geq |F|$. Set $E' = E \setminus \Big\{|T(\chi_F)| > 2c_0 \dfrac{|F|}{|E|}\Big\}$. Then

$$\left| \int_{E'} T(\chi_F)\, dx \right| \leq 2\, c_0 |E'| \frac{|F|}{|E|} \leq 2\, c_0\, |F| = 2\, c_0\, |E| \min\Big(1, \frac{|F|}{|E|}\Big). \tag{15}$$

Note that $|E'| \geq \frac{1}{2}|E|$ since

$$|E \setminus E'| \leq \left| \Big\{|T(\chi_F)| > 2c_0 \frac{|F|}{|E|}\Big\} \right| \leq c_0 \Big(2c_0 \frac{|F|}{|E|}\Big)^{-1} |F| \leq \frac{1}{2}|E|\,.$$

Case 2: Suppose that $|F| > |E|$. Then by the previous case there is an $F' \subsetneqq F$ such that $|F'| \geq \frac{1}{2}|F|$ and

$$\left| \int_{F'} T^*(\chi_E)\, dx \right| = \left| \int_{E} T(\chi_{F'})\, dx \right| \leq 2\, c_0^*\, |E|.$$

Find a subset F'' of $F \setminus F'$ with $|F''| \geq \frac{1}{2}|F \setminus F'|$ such that

$$\left| \int_{F''} T^*(\chi_E)\, dx \right| = \left| \int_{E} T(\chi_{F''})\, dx \right| \leq 2\, c_0^*\, |E|\,.$$

Find by induction sets $F^j \subsetneqq F \setminus (F' \cup \cdots \cup F^{j-1})$ such that $|F^j| \geq \frac{1}{2}|F \setminus (F' \cup \cdots \cup F^{j-1})|$ and

$$\left| \int_{F^j} T^*(\chi_E)\, dx \right| = \left| \int_{E} T(\chi_{F^j})\, dx \right| \leq 2\, c_0^*\, |E|, \quad j = 1, 2, \ldots, m.$$

Stop when $F^m = F \setminus (F' \cup \cdots \cup F^{m-1})$ has size at most $|E|$. Note that $m \leq 1 + \log_2 \frac{|F|}{|E|}$.

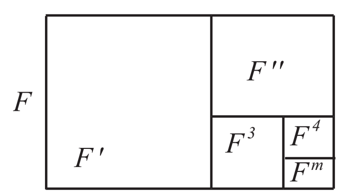

FIGURE 1. Choosing the subsets F^j of F.

Sum the estimates

$$\left| \int_{E} T(\chi_{F^j})\, dx \right| \leq 2\, c_0^*\, |E|, \quad j = 1, 2, \ldots, m \leq 1 + \log_2 \frac{|F|}{|E|}$$

and the easy estimate

$$\left| \int_{E} T(\chi_{F^m})\, dx \right| \leq \|T\|_{L^2 \to L^2}\, |E|$$

to obtain

$$\left| \int_{E'} T(\chi_F)\, dx \right| \leq (2\, c_0^* + \|T\|_{L^2 \to L^2}) |E| \Big(1 + \log_2^+ \frac{|F|}{|E|}\Big) \tag{16}$$

with $E' = E$.

Combining estimates (15) and (16) obtained in cases 1 and 2, respectively, we obtain (14) with constant $C = \max\big(2\, c_0, 2\, c_0^* + \|T\|_{L^2 \to L^2}\big)$. □

September 4, 2005.

We note that if T happens to have a bounded kernel or T can be written as a limit of operators with bounded kernel, and T and T^* are weak type $(1,1)$, then T must necessarily be L^2 bounded (with $\|T\|_{L^2\to L^2} \le 10\left(\|T\|_{L^1\to L^{1,\infty}}\|T^*\|_{L^1\to L^{1,\infty}}\right)^{\frac{1}{2}}$) by the interpolation theorem in [11]. Therefore in this case the L^2 boundedness assumption on T in Proposition 3.1 can be dropped. It is unknown, as of this writing, whether the conditions on the kernel in [11] can be dropped.

A multilinear version of Proposition 3.1 is proved in [2].

Before we end this section we point out that there exist analogous statements to Proposition 3.1 for bilinear operators.

These will not be discussed in detail but the main idea is that for all (E, F_1, F_2) there is $E' \subsetneqq E$ such that $|E'| \ge \frac{1}{2}|E|$ and

$$\left|\int_{E'} T(\chi_{F_1}, \chi_{F_2})\, dx\right| \le C|E|\Phi\left(\frac{|F_1|}{|E|}, \frac{|F_2|}{|E|}\right).$$

Here Φ is a suitable function of two variables whose precise form will be investigated for the bilinear Hilbert transform, see for instance (38), (44), (45).

We note that the equivalent facts discussed up to this point may also be stated without specific references to operators. One could, for instance, consider the set $\mathcal{P}$ of all pairs (f,α) (where f is a measurable function and $\alpha > 0$) satisfying (5) with $\mu(F)$ replaced by α and $T(\chi_F)$ replaced by f. Then elements of $\mathcal{P}$ are exactly those that satisfy a version of (6) with $\mu(F)$ replaced by α and $T(\chi_F)$ replaced by f. Taking $\mathcal{P}$ to be the set of all pairs $(f,\alpha) = (T(\chi_F), \mu(F))$ for any measurable subset F of X we obtain Proposition 2.1. For Proposition 2.2 we may take $\mathcal{P}$ to be the set of all pairs $(f,\alpha) = (T(\chi_{F_1}, \dots, \chi_{F_m}), \mu_1(F_1)^{\gamma_1} \dots \mu_m(F_m)^{\gamma_m})$.

4. Distributional estimates for the bilinear Hilbert transform

The family of bilinear Hilbert transforms was introduced in the early sixties by A. Calderón in his study of the first commutator, an operator arising in a series decomposition of the Cauchy integral along Lipschitz curves. Properties of the bilinear Hilbert transforms remained elusive until the appearance of the fundamental work of Lacey and Thiele [14], [15] in the late nineties who established their boundedness on certain products of Lebesgue spaces.

The bilinear Hilbert transforms also arise in a variety of other problems in bilinear Fourier analysis in a way analogous to that which the linear Hilbert transform arises in linear Fourier analysis. For instance, the study the convergence of the mixed Fourier series:

$$\lim_{N\to\infty} \sum_{\substack{|m-\alpha n|\le N \\ |m-n|\le N}} \sum \widehat{F}(m)\widehat{G}(n)e^{2\pi i(m+n)x} \tag{17}$$

for functions F, G on the circle is related to boundedness properties of the bilinear Hilbert transform, see [6], [8].

The bilinear Hilbert transform operator is defined for a parameter $\alpha \in \mathbf{R}$ by

$$H_\alpha(f,g)(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|t|\ge\varepsilon} f(x-t)g(x+\alpha t)\frac{dt}{t}, \qquad x \in \mathbf{R}$$

for functions f, g on the line. Lacey and Thiele [14], [15] proved that H_α maps $L^{p_1} \times L^{p_2}$ to L^p, whenever $1 < p_1, p_2 \le \infty$, $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p}$, and $\frac{2}{3} < p < \infty$, whenever $\alpha \ne -1$.

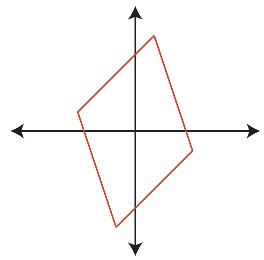

FIGURE 2. The summation in (17) is taken over all lattice points (m,n) inside an N-dilate of a fixed quadrilateral in $\mathbf{R}^2$.

In this and the later sections we indicate the ideas of the proof of estimates analogous to (4) for the bilinear Hilbert transform H_α. These estimates are contained in the following.

Theorem 4.1. *Let $2 \le p_2 < \infty$ and $\alpha \in \mathbf{R} \setminus \{0,-1\}$. Then there exist constants $C = C(\alpha, p_2), c = c(\alpha, p_2)$ such that for all measurable sets F_1, F_2 of finite measure we have*

$$|\{|H_\alpha(\chi_{F_1}, \chi_{F_2})| > \lambda\}| \le C\left(|F_1|\,|F_2|^{\frac{1}{p_2}}\right)^{\frac{p_2}{p_2+1}} \begin{cases} \lambda^{-\frac{p_2}{p_2+1}}(1+\log\frac{1}{\lambda})^{\frac{2p_2}{p_2+1}} & \textit{when } \lambda < 1, \\ e^{-c\sqrt{\lambda}} & \textit{when } \lambda \ge 1. \end{cases} \tag{18}$$

Analogously, the following estimate is valid for $2 \le p_1 < \infty$:

$$|\{|H_\alpha(\chi_{F_1}, \chi_{F_2})| > \lambda\}| \le C\left(|F_1|^{\frac{1}{p_1}}\,|F_2|\right)^{\frac{p_1}{p_1+1}} \begin{cases} \lambda^{-\frac{p_1}{p_1+1}}(1+\log\frac{1}{\lambda})^{\frac{2p_1}{p_1+1}} & \textit{when } \lambda < 1, \\ e^{-c\sqrt{\lambda}} & \textit{when } \lambda \ge 1. \end{cases} \tag{19}$$

These estimates correspond to the line segments $\{(\frac{1}{p_1}, \frac{1}{p_2}) : \ p_1 = 1,\ 2 \le p_2 < \infty\}$ and $\{(\frac{1}{p_1}, \frac{1}{p_2}) : \ 2 \le p_1 < \infty,\ p_2 = 1\}$. As a corollary we obtain the following distributional estimate corresponding to the line segment $\{(\frac{1}{p_1}, \frac{1}{p_2}) : \ 1 \le p_1, p_2 \le 2,\ \frac{1}{p_1} + \frac{1}{p_2} = \frac{3}{2}\}$.

Corollary 4.2. *For any $\alpha \in \mathbf{R} \setminus \{0,-1\}$ there exist constants $C = C(\alpha), c = c(\alpha)$ such that for all measurable sets F_1, F_2 of finite measure we have*
(20)

$$|\{|H_\alpha(\chi_{F_1}, \chi_{F_2})| > \lambda\}| \le C\left(|F_1|^{\frac12}|F_2|^{\frac12}\min(|F_1|, |F_2|)^{\frac12}\right)^{\frac23} \begin{cases} \lambda^{-\frac23}(1+\log\frac{1}{\lambda})^{\frac43} & \textit{for } \lambda < 1 \\ e^{-c\sqrt{\lambda}} & \textit{for } \lambda \ge 1. \end{cases}$$

Remark. In the distributional estimate (20), the expression $|F_1|^{\frac12}|F_2|^{\frac12}\min(|F_1|^{\frac12}, |F_1|^{\frac12})$ is dominated by $|F_1|^{\frac{1}{p_1}}|F_2|^{\frac{1}{p_2}}$, where $1 \le p_j \le 2$ and $\frac{1}{p_1} + \frac{1}{p_2} = \frac{3}{2}$. Thus this estimate (up to a logarithmic term) is similar to restricted weak type estimate for such exponents.

We note that the exponential decay of the distribution function of H_α as $\lambda \to \infty$ is not as strong as in the case of the Carleson-Hunt operator. At the moment we don't know if this decay is sharp. Estimates (18), (19), and (20) not only capture the boundedness of H_α on products of Lebesgue spaces but also yield other crucial quantitative information such as local exponential integrability and boundedness on other rearrangement invariant spaces even at the endpoint cases. Along these lines we have the following corollary concerning the exponential integrability of H_α.

Corollary 4.3. *Let $\alpha \in \mathbf{R} \setminus \{0,-1\}$ and $c = c(\alpha)$ be as in Corollary 4.2. Then there is a constant $C' = C'(\alpha)$ such athat for any bounded measurable set K and for all measurable* *sets F_1, F_2 of finite measure the following holds:*

$$\int_K e^{c'|H_\alpha(\chi_{F_1},\chi_{F_2})(x)|^{\frac{1}{2}}}dx \le C'\Big(|K| + \big(|F_1|^{\frac{1}{2}}|F_2|^{\frac{1}{2}}\min(|F_1|,|F_2|)^{\frac{1}{2}}\big)^{\frac{2}{3}}\Big)$$

for any $0 < c' < c$.

We obtain Theorem 4.1 using the model sum reduction of Lacey and Thiele [14],[15], a tree analysis based on a selection inspired by Lacey [13], and an "improved energy estimate" borrowed from the proof of (4) by Grafakos, Tao, and Terwilleger [10]. A variant of this energy estimate also appeared in the related work of Muscalu, Thiele, and Tao [17]. A detailed proof of Theorem 4.1 is given in [3]. An outline of the proof in this article is presented below.

5. Decomposition of the bilinear Hilbert transform

We fix $\alpha \in \mathbf{R} \setminus \{0, -1\}$. Until the end of this paper we will drop the dependence of H_α on α and will simply denote it by H. We will use the notation $\langle f, g\rangle$ for the complex inner product $\int f(x)\overline{g(x)}dx$. We will also use the notation $A \lesssim B$ to express that a the quantity A is at most a constant multiple of the quantity B.

The main object of study will the trilinear form

$$(f_1, f_2, f_3) \to \int H(f_1, f_2)(x)f_3(x)dx$$

for three functions f_1, f_2, f_3 which will be characteristic functions of sets of finite measure, i.e. $f_1 = \chi_{F_1}$, $f_2 = \chi_{F_2}$, and $f_3 = \chi_{E'}$.

We fix L to be the smallest integer greater than $2^{10}\max\{|\alpha|, \frac{1}{|\alpha|}, \frac{1}{|1+\alpha|}\}^3$. The dependence of the bounds on α will enter the proof through polynomial dependence on L. The distribution $p.v.\frac{1}{t}$ that appears in the definition of H can be written as $c_1\delta_0 + c_2\gamma$ for some constants c_1, c_2, where δ_0 is the Dirac mass at the origin and γ is another distribution that satisfies $\widehat{\gamma} = \chi_{(0,\infty)}$. Since all the estimates that we are going to be proving in this paper are trivial for δ_0, we may restrict our attention to γ. Let θ be a smooth function which is equal to 1 on $(-\infty, 2L)$ and 0 on $(3L, \infty)$. Define

$$\widehat{\psi}(\xi) = \theta(\xi) - \theta(2\xi).$$

Observe that $\widehat{\psi}$ is nonzero and is supported in $[L, 3L]$. For each integer k we define

$$\psi_k(x) = 2^{-\frac{k}{2}}\psi(2^{-k}x).$$

Then a quick examination of Fourier transforms yields that

$$\gamma = \sum_{k\in\mathbf{Z}} 2^{-\frac{k}{2}}\psi_k\,.$$

Matters therefore reduce to the study of the trilinear form

$$\Lambda(f_1, f_2, f_3) := \sum_{k\in\mathbf{Z}} 2^{-\frac{k}{2}} \int\int f_1(x-t)f_2(x+\alpha t)f_3(x)\psi_k(t)\,dt\,dx. \tag{21}$$

We further decompose the function ψ into a sum of at most $2L$ functions $\psi^{(M)}$ such that $\widehat{\psi^{(M)}}$ is supported in the interval $[M - \frac{1}{2}, M + \frac{1}{2}]$ for $L \le M \le 2L$. It suffices to study each such function separately. For notational convenience, we will omit the dependence on M and will just write ψ.

September 4, 2005.

Next we fix a Schwartz function ϕ of L^2 norm 1, with Fourier transform supported in $[-\frac{1}{2}, \frac{1}{2}]$, which also has the property that for all $\xi \in \mathbf{R}$ we have

$$\sum_{l\in\mathbf{Z}} \left|\widehat{\phi}(\xi - l/2)\right|^2 \equiv C_0$$

for some constant $C_0 > 0$. Let $u = I_u \times \omega_u$ be a rectangle in $\mathbf{R}^2$ and set

$$\phi_u(x) = |I_u|^{-\frac{1}{2}} \phi\Big(\frac{x - c(I_u)}{|I_u|}\Big) e^{2\pi i c(\omega_u) x},$$

where $c(J)$ denotes the center of the interval J. For each $k \in \mathbf{Z}$ we consider the set of dyadic rectangles of scale k:

$$\mathbf{S}_k = \{(2^k n, 2^k(n+1)) \times (2^{-k}m/2, 2^{-k}(m/2+1)) \,|\, m, n \in \mathbf{Z}\}.$$

Then $\mathbf{S} = \bigcup_k \mathbf{S}_k$ is the set of all dyadic rectangles of area 1 in $\mathbf{R}^2$.

Using the result in [9] one has that

$$f = \frac{1}{C_0} \sum_{u\in\mathbf{S}_k} \langle f, \phi_u\rangle \phi_u$$

where the series converges a.e. for all $f \in L^p$, $1 < p < \infty$.

Inserting this decomposition of the identity in the k^{th} term of (21), as in [15], we obtain

$$\Lambda(f_1, f_2, f_3) := \sum_{k\in\mathbf{Z}} \sum_{u_1,u_2,u_3\in\mathbf{S}_k} C_{k,u_1,u_2,u_3} \Lambda_{k,u_1,u_2,u_3}(f_1, f_2, f_3), \tag{22}$$

where

$$C_{k,u_1,u_2,u_3} = C_0^{-3} \int_{\mathbf{R}} \int_{\mathbf{R}} \phi_{u_1}(x - t)\phi_{u_2}(x + \alpha t)\phi_{u_3}(x)\psi_k(t)\, dt\, dx$$

and

$$\Lambda_{k,u_1,u_2,u_3}(f_1, f_2, f_3) = 2^{-\frac{k}{2}} \langle f_1, \phi_{u_1}\rangle\langle f_2, \phi_{u_2}\rangle\langle f_3, \phi_{u_3}\rangle.$$

A quick examination of the coefficients C_{k,u_1,u_2,u_3} yields that, with $A_i = \frac{c(I_{u_i})}{|I_{u_i}|}$ for $i = 1, 2, 3$, (note $A_i = n - \frac{1}{2}$ for some $n \in \mathbf{Z}$) for any $m \geq 10$, there exists a constant C_m such that

$$|C_{k,u_1,u_2,u_3}| \leq C_m \left(1 + \frac{\operatorname{diam}\{A_i\}}{4L}\right)^{-m} = C_m \left(1 + \frac{\max_{i,j} |c(I_{u_i}) - c(I_{u_j})|}{2^k\, 4L}\right)^{-m}. \tag{23}$$

Setting $F_1(x, t) = \phi_{u_1}(x - t)\phi_{u_2}(x + \alpha t)$, $F_2(x, t) = \phi_{u_3}(x)\psi_k(t)$ we have

$$\begin{aligned} \widehat{F_1}(\xi, \tau) &= \frac{1}{1+\alpha} \widehat{\phi_{u_1}}\Big(\frac{\alpha\xi - \tau}{1+\alpha}\Big)\widehat{\phi_{u_2}}\Big(\frac{\xi + \tau}{1+\alpha}\Big), \\ \widehat{F_2}(\xi, \tau) &= \frac{1}{1+\alpha} \widehat{\phi_{u_3}}(\xi)\widehat{\psi_k}(\tau). \end{aligned}$$

Applying the two-dimensional Plancherel formula, we deduce

$$|C_{k,u_1,u_2,u_3}| \leq \frac{C}{|1+\alpha|} \int\int \left|\widehat{\phi}\Big(\frac{\alpha\xi - \tau}{1+\alpha} - B_1\Big)\widehat{\phi}\Big(\frac{\xi + \tau}{1+\alpha} - B_2\Big)\widehat{\phi}(\xi - B_3)\psi(\tau)\right| d\xi d\tau, \tag{24}$$

where $B_i = \frac{c(\omega_{u_i})}{|\omega_{u_i}|} = 2^k c(\omega_{u_i})$. Note that each B_i is an integer or a half-integer.

Assuming that the integral above is not zero, an easy calculation shows that the triple of parameters B_1, B_2, B_3 depends only on the parameter B_3. More precisely, for each value of B_3, the quantities B_1 and B_2 only take a finite number of values depending on α.

Next we introduce parameters $\nu_1, \nu_2, \mu_1, \mu_2$ by setting

$$A_1 = A_3 + \nu_1, \qquad A_2 = A_3 + \nu_2, \qquad B_1 = \frac{\alpha}{\alpha+1} B_3 + \mu_1, \qquad B_2 = \frac{1}{\alpha+1} B_3 + \mu_2.$$

We also set $\nu = \max|\nu_i|$. We aim to reduce the sum over $u_1, u_2, u_3 \in \mathbf{S}_k$ as the rapidly converging sum over $\nu_1, \nu_2, \mu_1, \mu_2$ of the sum over the tiles u_3.

For N sufficiently large we have

$$|\Lambda(f_1, f_2, f_3)| \le \sum_{\nu=0}^{\infty} C_N (1 + \tfrac{\nu}{4L})^{-N} \sum_{\substack{(\nu_1,\nu_2):\\ \max|\nu_i|=\nu}} \sum_{\mu_1} \sum_{\mu_2} \Big| \sum_{k\in\mathbf{Z}} \sum_{u_3 \in \mathbf{S}_k} \varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,u_3} \Lambda_{k,u_1,u_2,u_3}(f_1, f_2, f_3) \Big| \tag{25}$$

where $u_1 = u_1(u_3)$ and $u_2 = u_2(u_3)$ are uniquely determined by u_3 in terms of $\nu_1, \nu_2, \mu_1, \mu_2$, $\varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,u_3}$ is a constant of modulus at most 1, and μ_1 and μ_2 take only a finite number of values depending on α.

It will clearly suffice to study the boundedness of the expression inside the absolute values in (25) and to obtain bounds independent of μ_i and polynomial in ν, since for each ν , there are of the order of ν pairs (ν_1, ν_2) with $\max|\nu_i| = \nu$.

To facilitate the study of the sums above, we introduce *tri-tiles*. A tri-tile is a rectangle $s = I_s \times \omega_s$ and three subrectangles s_1, s_2, s_3 built in the following way:

Let (u_1, u_2, u_3) be a triple of rectangles participating in the sum in (25). Define $I_s = I_{s_i} = I_{u_3}$. Defining the frequency projections requires a little bit more work, we cannot just use the dyadic grid. We want these projections to satisfy the following properties:

$$\mathcal{J} = \bigcup_{s\in S} \left(\omega_s \cup \omega_{s_1} \cup \omega_{s_2} \cup \omega_{s_3}\right) \text{ is a grid.} \tag{26}$$

$$\text{If } \omega_{s_i} \subsetneq J \text{ for some } J \in \mathcal{J}, \text{ then } \omega_{s_j} \subsetneq J \text{ for some } J \in \mathcal{J} \text{ for all } j = 1, 2, 3. \tag{27}$$

$$\omega_{s_i} \ne \omega_{s_j} \text{ for } i \ne j \tag{28}$$

We build these intervals by induction on the cardinality of the set of triples of rectangles but we omit here the precise construction.

We define the functions adapted to the tri-tile s with parameters $\nu_1, \nu_2, \mu_1, \mu_2$ as follows:

$$\begin{aligned}
\varphi_{s_1}^{\nu_1,\mu_1,\alpha}(x) &= |I_s|^{-\frac{1}{2}} \phi\left(\frac{x - c(I_s)}{|I_s|} - \nu_1\right) e^{2\pi i \left(\frac{\alpha}{\alpha+1} c(\omega_{s_1}) + \theta_{s_1}|\omega_{s_1}|\right)x} = \phi_{u_1}(x), \\
\varphi_{s_2}^{\nu_2,\mu_2,\alpha}(x) &= |I_s|^{-\frac{1}{2}} \phi\left(\frac{x - c(I_s)}{|I_s|} - \nu_2\right) e^{2\pi i \left(\frac{1}{\alpha+1} c(\omega_{s_2}) + \theta_{s_2}|\omega_{s_2}|\right)x} = \phi_{u_2}(x), \\
\varphi_{s_3}^{\alpha}(x) &= |I_s|^{-\frac{1}{2}} \phi\left(\frac{x - c(I_s)}{|I_s|}\right) e^{2\pi i \left(c(\omega_{s_3}) + \theta_{s_3}|\omega_{s_3}|\right)x} = \phi_{u_3}(x),
\end{aligned}$$

where the error terms θ_{s_i} in the modulations are chosen so that $\frac{\alpha}{\alpha+1} c(\omega_{s_1}) + \theta_{s_1}|\omega_{s_1}| = c(\omega_{u_1})$, $\frac{1}{\alpha+1} c(\omega_{s_2}) + \theta_{s_2}|\omega_{s_2}| = c(\omega_{u_2})$, and $c(\omega_{s_2}) + \theta_{s_3}|\omega_{s_3}| = c(\omega_{u_3})$. Obviously, $|\theta_{s_i}| \le CL$.

Then the expression inside the absolute values in (25) becomes exactly

$$\sum_{s_3 \in \bigcup_{k\in\mathbf{Z}} \mathbf{S}_k} |I_s|^{-\frac{1}{2}} \varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,s} \langle f_1, \varphi_{s_1}^{\nu_1,\mu_1,\alpha} \rangle \langle f_2, \varphi_{s_2}^{\nu_2,\mu_2,\alpha} \rangle \langle f_3, \varphi_{s_3}^{\alpha} \rangle.$$

This expression needs to be controlled with bounds that grow polynomially in the parameters ν_1, ν_2, and are independent of μ_1, μ_2. We will work with sums over finite sets of tri-tiles and get bounds independent of the choice of the finite set, which is clearly sufficient by a limiting argument.

For notational convenience, in the sequel we will suppress the dependence of the functions φ_{s_j} on the parameters $\nu_1, \nu_2, \mu_1, \mu_2$. Notice that

$$|\varphi_{s_k}(x)| \leq C\left(1 + \left|\frac{x - c(I_s)}{|I_s|} - \nu_k\right|\right)^{-10} \leq C\left(1 + \left|\frac{x - c(I_s)}{|I_s|}\right|\right)^{-10}(1+\nu)^{10}.$$

6. Estimates for the model sums. The case $I_s \subseteqq \Omega$.

Let S be a finite set of tri-tiles with fixed data ν_1, ν_2, μ_1, and μ_2. Then we define the "model sum" associated with S as follows:

$$H_S(f_1, f_2)(x) = \sum_{s \in S} |I_s|^{-\frac{1}{2}} \varepsilon_s \langle f_1, \varphi_{s_1}\rangle \langle f_2, \varphi_{s_2}\rangle \varphi_{s_3}(x).$$

We set

$$\Omega = \left\{x : \, M(\chi_{F_1})(x) > 8 \, \min\left(1, \frac{|F_1|}{|E|}\right)\right\} \bigcup \left\{x : \, M(\chi_{F_2})(x) > 8 \, \min\left(1, \frac{|F_2|}{|E|}\right)\right\},$$

where M is the Hardy-Littlewood maximal function. Since M if of weak type $(1,1)$ with constant at most 2, it is easy to see that $|\Omega| < \frac{1}{2}|E|$. We now set $E' = E \setminus \Omega$. Obviously, then $|E'| \geq \frac{1}{2}|E|$.

The main task is to obtain a good estimate for the expression

$$\int_{E'} H_S(\chi_{F_1}, \chi_{F_2})(x) dx = \langle H_S(\chi_{F_1}, \chi_{F_2}), \chi_{E'}\rangle.$$

To do so we will break the model sum into two parts: the sum over those $s \in S$ for which $I_s \subseteqq \Omega$ (easier case) and the sum over tiles with $I_s \nsubseteqq \Omega$.

In this article, we briefly discuss (without proof) the easier case. Given a set of tiles S we set

$$S_J = \{s \in S : I_s = J\}.$$

Then for $A > 1$ and F_1, F_2 sets of finite measure one can show (see [3], [15])

$$\|H_{S_J}(\chi_{F_1}, \chi_{F_2})\|_{L^1((AJ)^c)} \leq (1+\nu)^{20} C_M A^{-M} \, |J| \, \inf_{x \in J} M(\chi_{F_1})(x) \, \inf_{x\in J} M_2(\chi_{F_2})(x), \tag{29}$$

where $M_2(g) = M(g^2)^{1/2}$. Using this estimate, it is not difficult to obtain the following:

$$\left|\int_{E'} \sum_{s: I_s \subseteqq \Omega} |I_s|^{-\frac{1}{2}} \langle \chi_{F_1}, \phi_{s_1}\rangle \langle \chi_{F_2}, \phi_{s_2}\rangle \phi_{s_3}(x) \, dx\right| \leq C_\nu (\min|F_1|, |F_2|)^{\frac{1}{2}} |F_1|^{\frac{1}{2}} |F_2|^{\frac{1}{2}} |E|^{-\frac{1}{2}}. \tag{30}$$

where C_ν has at mot polynomial growth in ν.

The proof of estimate (30) follows by grouping all dyadic intervals $J \subseteqq \Omega$ into sets $\mathcal{F}_k$ $(k \geq 0)$ in the following way:

$$\mathcal{F}_k = \{J : \, 2^k J \subseteqq \Omega, \, 2^{k+1} J \nsubseteqq \Omega\}.$$

We note that

$$\sum_{J \in \mathcal{F}_k} |J| \leq 4|\Omega| \leq 2|E|.$$

Indeed, assume J_{max} is a maximal element of $\mathcal{F}_k$ with respect to inclusion. If $J \subseteqq J_{max}$ and $|J| < |J_{max}|$, then J must have a common endpoint with J_{max} (otherwise, we would have

$2^{k+1}J = 2^k(2J) \subseteqq 2^k J_{max} \subseteqq \Omega$, thus $J \notin \mathcal{F}_k$). Thus, for each particular scale, J_{max} may contain at most 2 intervals belonging to $\mathcal{F}_k$. Therefore

$$\sum_{J\in\mathcal{F}_k, J\subseteqq J_{max}} |J| \le \sum_{k=0}^{\infty} 2^{-k+1}|J_{max}| \le 4|J_{max}|.$$

Since the maximal elements of $\mathcal{F}_k$ are disjoint, summing over them we obtain the required conclusion.

Also, for any $J \in \mathcal{F}_k$ we have $E' \subseteqq (\Omega)^c \subseteqq (2^k J)^c$. Thus we have:

$$\begin{aligned}
& \left| \int_{E'} H_{\{I_s \subseteqq \Omega\}}(\chi_{F_1}, \chi_{F_2}) dx \right| \\
\le\ & \sum_{J\subseteqq\Omega} \left| \int_{E'} H_{S_J}(\chi_{F_1}, \chi_{F_2}) dx \right| \\
=\ & \sum_{k=0}^{\infty} \sum_{J\in\mathcal{F}_k} \left| \int_{E'} H_{S_J}(\chi_{F_1}, \chi_{F_2}) dx \right| \\
\le\ & \sum_{k=0}^{\infty} \sum_{J\in\mathcal{F}_k} \|H_{S_J}\|_{L^1((2^k J)^c)} \\
\le\ & C_M(1+\nu)^{20} \sum_{k=0}^{\infty} \sum_{J\in\mathcal{F}_k} |J| 2^{-kM} \inf_{x\in J} M(\chi_{F_1}) \inf_{x\in J} M_2(\chi_{F_2}) \\
\le\ & C_M(1+\nu)^{20} \sum_{k=0}^{\infty} 2^{-kM} C_0^{2k+2} \sum_{J\in\mathcal{F}_k} |J| \inf_{2^{k+1}J} M(\chi_{F_1}) \inf_{2^{k+1}J} M_2(\chi_{F_2}) \\
\le\ & C'(1+\nu)^{20} \sum_{k=0}^{\infty} 2^{-kM} C_0^{2k+2} \sum_{J\in\mathcal{F}_k} |J| \frac{|F_1|}{|E|} \left(\frac{|F_2|}{|E|}\right)^{\frac{1}{2}} \\
\le\ & C(1+\nu)^{20} |F_1|\, |F_2|^{\frac{1}{2}} |E|^{-\frac{1}{2}}.
\end{aligned}$$

7. Estimates for model sums. The case $I_s \nsubseteq \Omega$.

We will now deal with the harder case $I_s \nsubseteq \Omega$. This part of the proof is based on an adaptation of the $L^2 \times L^2 \to L^{1,\infty}$ estimate in [13].

We denote by P the set of all tri-tiles $s \in S$, for which $I_s \nsubseteq \Omega$. Tri-tiles admit a partial order. We say that $s < s'$ if $I_s \subseteqq I_{s'}$ and $\omega_{s'} \subseteqq \omega_s$. We note that s and s' intersect as rectangles if and only if they are comparable under " $<$ ".

The construction of tri-tiles has as a consequence that if $s < s'$, then $\omega_{s'} \subseteqq \omega_{s_i}$ for some $i = 1, 2, 3$ or it is disjoint with all ω_{s_i}'s.

We say that a collection of tri-tiles T is a tree with top t if for all $s \in T$, $s < t$. Every finite collection of tri-tiles S is a union of trees. Indeed, if we denote by S^* the set of all elements in S which are maximal under " $<$ ", and, for each $t \in S^*$, T_t is the maximal tree in S with top t, then $S = \cup_{t\in S^*} T_t$. We refine the notion of the tree by saying that T is a *j-tree* ($j = 1, 2, 3$) if T is a tree with top T and for every $s \in T$, $\omega_{s_j} \cap \omega_t = \emptyset$.

For a tree T, $s \in T$, $s \neq t$, at most one of the intervals ω_{s_i} can intersect ω_t. Thus if we denote $T_k = \{s \in T : \omega_{s_k} \cap \omega_t \neq \emptyset\}$, $k = 1, 2, 3$, then T_k is a j-tree for $j \neq k$ (there are also elements such that $\omega_{s_i} \cap \omega_t = \emptyset$ for all $i = 1, 2, 3$, but those may be added to any of the

T_k's). Then $T = \cup_{k=1}^{3} T_k$, i.e. any tree is a union of at most three subtrees which are j-trees for at least two choices of j.

We define the k-energy of a finite set of tiles S by

$$\mathcal{E}_k(S) = \sup \frac{1}{\|f_k\|_2} \left(|I_t|^{-1} \sum_{s \in T} |\langle f_k, \varphi_{s_k} \rangle|^2 \right)^{\frac{1}{2}}, \tag{31}$$

where the supremum is taken over all k-trees $T \subseteqq S$. Note that a singleton $\{s\}$ is a k-tree for all k, so for all $s \in S$,

$$|I_s|^{-\frac{1}{2}} |\langle f_k, \varphi_{s_k} \rangle| \leq \mathcal{E}_k(S) \, \|f_k\|_2.$$

Now fix some $j = 1, 2, 3$ and let T be a k-tree for $k \neq j$. Applying the above estimate and the Cauchy-Schwarz inequality, we deduce

$$\begin{aligned} |\langle H_T(f_1, f_2), f_3 \rangle| &\leq \sum_{s \in T} \frac{|\langle f_j, \varphi_{s_j} \rangle|}{|I_s|^{\frac{1}{2}}} \prod_{k \neq j} |\langle f_k, \varphi_{s_k} \rangle| \\ &\leq \mathcal{E}_j(S) \, \|f_j\|_2 \sum_{s \in T} \prod_{k \neq j} |\langle f_k, \varphi_{s_k} \rangle| \\ &\leq \mathcal{E}_j(S) \, \|f_j\|_2 \, |I_t| \prod_{k \neq j} \frac{1}{\|f_k\|_2} \left(|I_t|^{-1} \sum_{s \in T} |\langle f_k, \varphi_{s_k} \rangle|^2 \right)^{\frac{1}{2}} \|f_k\|_2 \\ &\leq |I_t| \prod_{j=1}^{3} \mathcal{E}_j(S) \, \|f_j\|_2. \end{aligned} \tag{32}$$

This is crucial estimate on a single tree that will be used in conjunction with the idea that any tree can be written as a union of three trees of the above type.

Next, we state the main lemma which will allow us to obtain the estimates for the model sums (cf. [13]).

Lemma 7.1. *Let S be a finite set of tri-tiles. Then S can be written as a union of two sets $S = S_1 \cup S_2$, which have the following properties. Let S_1^* be the set of elements which are maximal in S_1 under "$<$" (i.e. S_1 is a union of trees with tops in S_1^*). We then have*

$$\sum_{t \in S_1^*} |I_t| \leq C_1 (1 + \nu)^{20} \mathcal{E}_k(S)^{-2}, \tag{33}$$

$$\mathcal{E}_k(S_2) \leq \frac{1}{2} \, \mathcal{E}_k(S). \tag{34}$$

This lemma only yields weak-type estimates from $L^2 \times L^2$ into $L^{1,\infty}$. But the fact that we are now working with the set of tiles $P = \{s \in S : I_s \nsubseteq \Omega\}$ and all functions are characteristic of some sets gives us an advantage quantified by the following energy estimate which appeared in [10], [7], and is essentially contained in [17]:

Lemma 7.2. *For $k = 1, 2$ and $f_k = \chi_{F_k}$, there exists a constant $C > 0$, such that the following estimate is valid:*

$$\mathcal{E}_k(P) \leq C |E|^{-\frac{1}{2}} \min \left[\left(\frac{|F_k|}{|E|} \right)^{\frac{1}{2}}, \left(\frac{|F_k|}{|E|} \right)^{-\frac{1}{2}} \right] \tag{35}$$

Using these lemmata we can derive an estimate of the model sum for the case $I_s \nsubseteq \Omega$ in the following way. We construct inductively the sequence of pairwise disjoint sets P_j such that

$$P = \bigcup_{j=-\infty}^{n_0} P_j$$

and the following properties are satisfied:
(1) $\mathcal{E}_k(P_j) \le 2^{j+1}$ for $k = 1, 2, 3$.
(2) P_j is a union of trees T_{jk} such that $\sum_k |I_{top(T_{jk})}| \le C_0(1+\nu)^{20} 2^{-2j}$ for all $j \le n_0$.
(3) $\mathcal{E}_k(P \setminus (P_{n_0} \cup \dots \cup P_j)) \le 2^j$ for $k = 1, 2, 3$.
Using the families P_j we obtain the following:

$$\begin{aligned}
& |\langle H_P(\chi_{F_1}, \chi_{F_2}), \chi_{E'} \rangle| \\
\le \;& \sum_{j=-\infty}^{\infty} \sum_k |\langle H_{T_{jk}}(\chi_{F_1}, \chi_{F_2}), \chi_{E'} \rangle| \\
\le \;& C \sum_{j=-\infty}^{\infty} \big(\sum_k |I_{top_{T_{jk}}}| \big) \mathcal{E}_1(\chi_{F_1}, S_j) \mathcal{E}_2(\chi_{F_2}, S_j) \mathcal{E}_3(\chi_{E'}, S_j) |F_1|^{\frac12} |F_2|^{\frac12} |E|^{\frac12} \\
\le \;& C' \sum_{j=-\infty}^{\infty} 2^{-2j} \min\big(|F_1|^{-\frac12}, \frac{|F_1|^{\frac12}}{|E|}, 2^j\big) \min\big(|F_2|^{-\frac12}, \frac{|F_2|^{\frac12}}{|E|}, 2^j\big) 2^j \, |F_1|^{\frac12} |F_2|^{\frac12} |E|^{\frac12} \\
= \;& C' \sum_{j=-\infty}^{\infty} 2^{-j} \min\big(|F_1|^{-\frac12}, \frac{|F_1|^{\frac12}}{|E|}, 2^j\big) \min\big(|F_2|^{-\frac12}, \frac{|F_2|^{\frac12}}{|E|}, 2^j\big) \, |F_1|^{\frac12} |F_2|^{\frac12} |E|^{\frac12}, \qquad (36)
\end{aligned}$$

where we used the estimate on a single tree (32) and the improved energy estimate (35).
It takes some work but one can show that the last expression above is at most

$$C_1 \min(|F_1|^{\frac12}, |F_2|^{\frac12}) \min\left(\frac{|F_1|^{\frac12} |F_2|^{\frac12}}{|E|^{\frac12}}, |E|^{\frac12} \right) \left(1 + \left| \left| \log \frac{|F_1|}{|E|} \right| - \left| \log \frac{|F_2|}{|E|} \right| \right| \right), \qquad (37)$$

for some $C_1 > 0$. We summarize the results so far in a proposition.

Proposition 7.3. *There exists a constant C_1 such that, for any sets E,F_1,F_2 with the property that $|E|^2 \ge |F_1|\,|F_2|$ there exists a set $E' \subsetneqq E$ with $|E'| \ge \frac12 |E|$ such that for any set of tri-tiles S we have the following estimate:*

$$\left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})(x)dx \right| \le C_1 |E| \min\left[\frac{|F_1|}{|E|}, \frac{|F_2|}{|E|} \right]^{\frac12} \left[\frac{|F_1|}{|E|} \frac{|F_2|}{|E|} \right]^{\frac12} \left(1 + \log \frac{|E|^2}{|F_1|\,|F_2|} \right). \qquad (38)$$

This estimate is also valid for the bilinear Hilbert transform H.

Proof. The result for H_S follows from the estimates (30) and (37). Note that the construction of E' did not depend on the choice of the set of tri-tiles, so E' is the same for any S, and by an averaging argument this estimate is also valid for H. □

It is clear that, since both adjoints of H_S, are "essentially" the same operators, the same estimate (with different constants) also holds for them.

September 4, 2005.

8. $L^{r_1} \times L^{r_2} \to L^r$ boundedness of the bilinear Hilbert transform.

We briefly mention why estimates (30) and (37) imply boundedness of the model sum operator H_S from $L^{r_1} \times L^{r_2}$ to L^r for $\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r}$, $r_1, r_2 > 1$, $r > \frac{2}{3}$. This section can be skipped by citing [14], [15].

Take some p_1, p_2 such that $\frac{1}{p_1} + \frac{1}{p_2} = \frac{3}{2}$ and $p_1, p_2 > 1$. We will show that H_S is of restricted weak type (r_1, r_2, r) where $\frac{1}{r_1} = \frac{1}{p_1} - \varepsilon$, $\frac{1}{r_2} = \frac{1}{p_2} - \varepsilon$ and $\frac{1}{r} = \frac{3}{2} - 2\varepsilon$. Then the strong boundedness for the claimed range of exponents follows by the interpolation theorem of Grafakos and Tao [11] (the operator H_S has bounded kernel whenever S is a finite set).

We recall that a bilinear operator T is of restricted weak type (r_1, r_2, r) if and only if the following is valid: For any sets E, F_1, F_2 of finite measure there exists a set $E' \subset E$ with $|E'| \geq \frac{1}{2}|E|$, such that

$$\left| \int_{E'} T(\chi_{F_1}, \chi_{F_2})(x) dx \right| \lesssim \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}}. \tag{39}$$

Take arbitrary sets E, F_1, F_2 of finite positive measure. It follows from (30) and (37) that

$$\left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})\, dx \right| \lesssim \frac{|F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{2}}} \left(1 + \left| \log \frac{|F_1|}{|E|} \right| \right) \left(1 + \left| \log \frac{|F_2|}{|E|} \right| \right). \tag{40}$$

We will use the fact that $1 + \log x \lesssim x^{\varepsilon}$ for $x \geq 1$. In the case when $|E| \geq \max(|F_1|, |F_2|)$ we can estimate the righthand side of (40) by the expression

$$\frac{|F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{2}}} \left(1 + \log \frac{|E|}{|F_1|}\right) \left(1 + \log \frac{|E|}{|F_2|}\right) \lesssim \frac{|F_1|^{\frac{1}{p_1}-\varepsilon} |F_2|^{\frac{1}{p_2}-\varepsilon}}{|E|^{\frac{1}{2}-2\varepsilon}} = \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}}.$$

Now consider the case $|F_1| \leq |E| \leq |F_2|$ (as the case $|F_2| \leq |E| \leq |F_1|$ is symmetric). Fix some $\varepsilon_1 > 2\varepsilon$. Put $\alpha = \frac{1}{p_1} - \varepsilon + \varepsilon_1$ (ε and ε_1 have to be chosen small enough, so that $\alpha \leq 1$) and $\beta = \frac{1}{p_2} - \varepsilon_1 + \varepsilon$ (thus $\beta \leq 1$ also). We have $\alpha + \beta = \frac{3}{2}$. Thus, similarly to (40), we obtain:

$$\begin{aligned} \left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})\, dx \right| &\lesssim \frac{|F_1|^{\alpha} |F_2|^{\beta}}{|E|^{\frac{1}{2}}} \left(1 + \log \frac{|E|}{|F_1|}\right) \left(1 + \log \frac{|F_2|}{|E|}\right) \\ &\lesssim \frac{|F_1|^{\alpha} |F_2|^{\beta}}{|E|^{\frac{1}{2}}} \left(\frac{|E|}{|F_1|}\right)^{\varepsilon_1} \left(\frac{|F_2|}{|E|}\right)^{\varepsilon_1 - 2\varepsilon} \\ &= \frac{|F_1|^{\frac{1}{p_1}-\varepsilon} |F_2|^{\frac{1}{p_2}-\varepsilon}}{|E|^{\frac{1}{2}-2\varepsilon}} = \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}}. \end{aligned}$$

The remaining case is $|E| \leq \min(|F_1|, |F_2|)$. We observe that in this case the set Ω is empty, since $M(\chi_{F_i}) \leq 1$. We therefore only need to use (36) which for $|E|$ small yields:

$$\begin{aligned} \left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})\, dx \right| &\lesssim \min(|F_1|, |F_2|)^{\frac{1}{2}} |E|^{\frac{1}{2}} \left(1 + \log \frac{|F_1|}{|E|}\right) \left(1 + \log \frac{|F_2|}{|E|}\right) \\ &\lesssim |F_1|^{\frac{1}{p_1}-\frac{1}{2}} |F_2|^{\frac{1}{p_2}-\frac{1}{2}} |E|^{\frac{1}{2}} \left(\frac{|F_1|}{|E|}\right)^{\frac{1}{2}-\varepsilon} \left(\frac{|F_2|}{|E|}\right)^{\frac{1}{2}-\varepsilon} \\ &= \frac{|F_1|^{\frac{1}{p_1}-\varepsilon} |F_2|^{\frac{1}{p_2}-\varepsilon}}{|E|^{\frac{1}{2}-2\varepsilon}} = \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}}. \end{aligned}$$

Thus, for any measurable sets E and F_1, F_2, H_S satisfies (39) and this implies that it is of restricted weak type (r_1, r_2, r). The strong type estimates for the same range of exponents can now be obtained by varying r_1 and r_2 and using the result on interpolation between adjoint operators (cf. [11]).

9. Distributional estimates corresponding to the case $p_1 = 1$, $2 \le p_2 < \infty$.

In this section we fix $2 \le p_2 < \infty$. In the case $p_2 = 2$ for the moment we shall assume that $|F_1| \le |F_2|$. We consider four cases:

CASE: $p_2 = 2$, $|E|^{\frac{3}{2}} \ge |F_1|\,|F_2|^{\frac{1}{2}}$, $|F_1| \le |F_2| \le |E|$.

Since $|E|^{\frac{3}{2}} \ge |F_1|\,|F_2|^{\frac{1}{2}}$ and $|F_2| \ge |F_1|$, we have $|E|^2 \ge |F_1|\,|F_2|$. Using estimate (38) we obtain

$$\left|\int_{E'} H(\chi_{F_1}, \chi_{F_2})dx\right| \le C_1 \frac{|F_1|\,|F_2|^{\frac{1}{2}}}{|E|^{\frac{1}{2}}}\left(1 + \log \frac{|E|^{\frac{3}{2}}}{|F_1|\,|F_2|^{\frac{1}{2}}}\right). \tag{41}$$

We note that this estimate is also valid if $|E| \ge \max |F_i|$, even when $|F_1| \ge |F_2|$. We will use this estimate in the inductive procedures below.

CASE: $p_2 > 2$, $|E|^{1+\frac{1}{p_2}} \ge |F_1|\,|F_2|^{\frac{1}{p_2}}$, $|E| \ge |F_2|$.

Let $\alpha = \frac{1}{2} - \frac{1}{p_2} > 0$, $\beta = 1 - \frac{1}{p_2} > 0$. Since $|E| \ge |F_2|$ we must have $|E|^2 \ge |F_1|\,|F_2|$. Using (38) we obtain

$$\begin{aligned}\left|\int_{E'} H(\chi_{F_1}, \chi_{F_2})(x)\,dx\right| &\le C_1 \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\left(\frac{|F_2|}{|E|}\right)^{\alpha}\left(1 + \log \frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}} + \log \frac{|E|^{\beta}}{|F_2|^{\beta}}\right) \\ &\lesssim \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\left(1 + \log \frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right), \end{aligned} \tag{42}$$

since the function $f(x) = x^{\alpha}(1 + \log \frac{1}{x^{\beta}})$ is bounded on $[0, 1]$ when $\alpha > 0$ (here $x = \frac{|F_2|}{|E|}$).

CASE: $p_2 \ge 2$, $|E|^{1+\frac{1}{p_2}} \ge |F_1|\,|F_2|^{\frac{1}{p_2}}$, $|E| \le |F_2|$ (which implies $|E| \ge |F_1|$).

In this case we will obtain an estimate via an iterative procedure. The iteration procedure will consist of two parts. At first, we set $F_2^0 = F_2$. We will continue this part of the iteration until the first integer n such that $|F_2^n| \le |E|$. Let H^{*2} be the second dual of H. At the j^{th} step, according to the estimates above, we choose a subset S^j of F_2^j with $|S^j| \ge \frac{1}{2}|F_2^j|$, such that:

$$\left|\int_{S^j} H^{*2}(\chi_{F_1}, \chi_E)(x)dx\right| \lesssim \frac{|F_1|\,|E|^{\frac{1}{p_2}}}{|F_2^j|^{\frac{1}{p_2}}}\left(1 + \log \frac{|F_2^j|^{1+\frac{1}{p_2}}}{|F_1|\,|E|^{\frac{1}{p_2}}}\right) \le |F_1|\left(1 + \log \frac{|F_2|^{1+\frac{1}{p_2}}}{|F_1|\,|E|^{\frac{1}{p_2}}}\right).$$

Then we set $F_2^{j+1} = F_2^j \setminus S^j$. Obviously, for the number of steps n we have $n \lesssim 1 + \log \frac{|F_2|}{|E|}$. Thus, we have

$$\begin{aligned}\Big|\int_E H(\chi_{F_1},\chi_{F_2})\,dx\Big| &\lesssim |F_1|\Big(1+\log\frac{|F_2|^{1+\frac{1}{p_2}}}{|F_1|\,|E|^{\frac{1}{p_2}}}\Big)\Big(1+\log\frac{|F_2|}{|E|}\Big)+\Big|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\Big|\\ &\lesssim \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\Big(1+\log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\Big)^2+\Big|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\Big|.\end{aligned}$$

In the last line we have used the following simple inequality

$$\Big(1+\log\big(ab^{1+\frac{1}{p_2}}\big)\Big)\Big(1+\log b\Big)\lesssim\Big(1+\log\frac{a}{b^{\frac{1}{p_2}}}\Big)^2 b^{\frac{1}{p_2}} \tag{43}$$

for $a\geq 1$, $b\geq 1$, such that $ab^{-\frac{1}{p_2}}\geq 1$ (with $a=\frac{|E|}{|F_1|}$, $b=\frac{|F_2|}{|E|}$).

It remains to estimate the term

$$\Big|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\Big|.$$

In the second part of the iteration process we proceed in a similar manner, only now we will be splitting either F_2 or E, depending on which one is larger in size. We set $E^n = E$. At the j^{th} step, if $|E^j| \geq |F_2^j|$, we choose $S^j \subset E^j$ such that $|S^j| \geq \frac{1}{2}|E^j|$ and

$$\begin{aligned}\Big|\int_{S^j} H(\chi_{F_1},\chi_{F_2^j})\,dx\Big| &\lesssim \frac{|F_1|\,|F_2^j|^{\frac{1}{p_2}}}{|E^j|^{\frac{1}{p_2}}}\Big(1+\log\frac{|E^j|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2^j|^{\frac{1}{p_2}}}\Big)\\ &\leq |F_1|\frac{|F_2^j|^{\frac{1}{p_2}}}{|E^j|^{\frac{1}{p_2}}}\Big(1+\log\frac{|E^j|^{\frac{1}{p_2}}}{|F_2^j|^{\frac{1}{p_2}}}+\log\frac{|E|}{|F_1|}\Big)\\ &\lesssim |F_1|\Big(1+\log\frac{|E|}{|F_1|}\Big),\end{aligned}$$

where we have once again made use of the fact that $f(x) = x\cdot\log\frac{1}{x}$ is bounded on $[0,1]$ ($x=\frac{|F_2^j|^{\frac{1}{p}}}{|E^j|^{\frac{1}{p}}}\leq 1$).

In the other case, when $|F_2^j| \geq |E^j|$, we choose $S^j \subset F_2^j$ with $|S^j| \geq \frac{1}{2}|F_2^j|$ such that

$$\Big|\int_{S^j} H^{*2}(\chi_{F_1},\chi_{E^j})\,dx\Big|\lesssim\frac{|F_1|\,|E^j|^{\frac{1}{p_2}}}{|F_2^j|^{\frac{1}{p_2}}}\Big(1+\log\frac{|F_2^j|^{1+\frac{1}{p_2}}}{|F_1|\,|E^j|^{\frac{1}{p_2}}}\Big).$$

An identical calculation and the fact that $|F_2^j| \leq |E|$ show that this can also be dominated by $|F_1|\big(1+\log\frac{|E|}{|F_1|}\big)$.

In the first case we set $F_2^{j+1} = F_2^j$, $E^{j+1} = E^j \setminus S^j$. In the second case we set $F_2^{j+1} = F_2^j \setminus S^j$, $E^{j+1} = E^j$. We proceed until the first integer m such that both $|E^m|, |F_2^m| \leq |F_1|$.

Obviously, the number of steps in the second part $m \lesssim \big(1+\log\frac{|E|}{|F_1|}\big)$. We now have

$$\begin{aligned}
\Big|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\Big| &= \Big|\int_{E^{n+1}\cup S^n} H(\chi_{F_1},\chi_{F_2^n})\,dx\Big| \\
&\le \Big|\int_{S^n} H(\chi_{F_1},\chi_{F_2^n})(x)\,dx\Big| + \Big|\int_{E^{n+1}} H(\chi_{F_1},\chi_{F_2^{n+1}})\,dx\Big| \\
&\lesssim |F_1|\Big(1+\log\frac{|E|}{|F_1|}\Big) + \Big|\int_{E^{n+1}} H(\chi_{F_1},\chi_{F_2^{n+1}})\,dx\Big| \\
&\lesssim \dots \\
&\lesssim m\,|F_1|\Big(1+\log\frac{|E|}{|F_1|}\Big) + \Big|\int_{E^m} H(\chi_{F_1},\chi_{F_2^m})\,dx\Big| \\
&\lesssim |F_1|\Big(1+\log\frac{|E|}{|F_1|}\Big)^2 + |E^m|^{\frac12}|F_1|^{\frac14}|F_2^m|^{\frac14} \\
&\lesssim |F_1|\frac{|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\Big(1+\log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\Big)^2,
\end{aligned}$$

where we made use of the boundedness of H on $L^4\times L^4\to L^2$ and the following inequality: For any $a\ge 1$,$b\ge 1$, such that $ba^{-\frac{1}{p_2}}\ge 1$ we have

$$(1+\log b)^2 \lesssim (1+\log(ba^{-\frac{1}{p_2}}))^2 a^{\frac{1}{p_2}},$$

with $a=\frac{|F_2|}{|E|}$, $b=\frac{|E|}{|F_1|}$.

CASE: $p_2\ge 2$, $|E|^{1+\frac{1}{p_2}}\le |F_1|\,|F_2|^{\frac{1}{p_2}}$ (still assuming that $|F_1|\le|F_2|$ when $p_2=2$).

Here we will need the following lemma which can be proved by an inductive procedure similar to the one described above and whose proof is omitted (see [3]).

Lemma 9.1. *Let $2\le p_2<\infty$. For all measurable sets E, F_1, F_2 of finite measure satisfying $|E|^{1+\frac{1}{p_2}}\le |F_1|\,|F_2|^{\frac{1}{p_2}}$ (and also $|F_1|\le|F_2|$ when $p_2=2$) we have*

$$\Big|\int_E H(\chi_{F_1},\chi_{F_2})(x)dx\Big| \lesssim |E|\Big(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\Big)^2.$$

We conclude this section by reviewing what we have deduced so far.

In the case $p_2>2$ we obtain the following estimate: For any sets F_1, F_2, and E of finite measure we can find $E'\subset E$ with $|E'|\ge\frac12|E|$ such that

$$\Big|\int_{E'} H(\chi_{F_1},\chi_{F_2})\,dx\Big| \lesssim |E|\min\Big[1,\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\Big]\Big[1+\Big|\log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\Big|\Big]^2. \tag{44}$$

We now remove the assumption that $|F_1|\le|F_2|$ when $p_2=2$. For $p_2=2$, we can consider the (symmetric) case when $|F_1|\ge|F_2|$, proceed as above with the roles of F_1 and F_2 interchanged and putting together the two estimates we obtain: For any sets F_1, F_2, and E of finite measure we can find a set $E'\subset E$ with $|E'|\ge\frac12|E|$ such that

$$\Big|\int_{E'} H(\chi_{F_1},\chi_{F_2})\,dx\Big| \lesssim |E|\min\Big[1,\frac{\min(|F_i|)^{\frac12}|F_1|^{\frac12}\,|F_2|^{\frac12}}{|E|^{\frac32}}\Big]\Big[1+\Big|\log\frac{|E|^{\frac32}}{\min(|F_i|)^{\frac12}|F_1|^{\frac12}\,|F_2|^{\frac12}}\Big|\Big]^2. \tag{45}$$

10. DISTRIBUTIONAL ESTIMATES FOR THE BILINEAR HILBERT TRANSFORM

We now have all the tools we need to prove Theorem 4.1.

For a given $\lambda > 0$, we set $E_\lambda^+ = \{H(\chi_{F_1}, \chi_{F_2}) > \lambda\}$ and $E_\lambda^- = \{H(\chi_{F_1}, \chi_{F_2}) < -\lambda\}$. Suppose that $|E_\lambda^+|^{1+\frac{1}{p_2}} > |F_1|\,|F_2|^{\frac{1}{p_2}}$. Then by (44) there is a subset S_λ^+ of E_λ^+ of at least half its measure so that

$$\frac{\lambda}{2}|E_\lambda^+| \le \left|\int_{S_\lambda^+} H(\chi_{F_1}, \chi_{F_2})\,dx\right| \le C_3 \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E_\lambda^+|^{\frac{1}{p_2}}}\left(1+\log\frac{|E_\lambda^+|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right)^2.$$

This implies that

$$|E_\lambda^+| \le C_4 \left(|F_1|\,|F_2|^{\frac{1}{p_2}}\right)^{\frac{p_2}{p_2+1}} \cdot \lambda^{-\frac{p_2}{p_2+1}}\left(1+\log\frac{1}{\lambda}\right)^{\frac{2p_2}{p_2+1}}. \tag{46}$$

But then this implies that there is a $\lambda_0 > 0$ such that for $\lambda > \lambda_0$ we have $|E_\lambda^+|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$. Thus for $\lambda > \lambda_0$, $|E_\lambda^+|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$ holds and estimate (44) gives

$$\lambda|E_\lambda^+| \le C_5\,|E_\lambda^+|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E_\lambda^+|^{1+\frac{1}{p_2}}}\right)^2,$$

from which one easily deduces that

$$|E_\lambda^+| \le \frac{1}{2}C\,e^{-c\sqrt{\lambda}}\left(|F_1|\,|F_2|^{\frac{1}{p_2}}\right)^{\frac{p_2}{p_2+1}}. \tag{47}$$

Suppose now that $\lambda \le \lambda_0$. As shown, if $|E_\lambda^+|^{1+\frac{1}{p_2}} > |F_1|\,|F_2|^{\frac{1}{p_2}}$, then (46) holds. If $|E_\lambda^+|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$ then (47) holds which is even stronger.

The same argument applies for the set E_λ^- with the same λ_0.

For $p_2 = 2$ we run the same argument for estimate (45) and at the end dominate the expression $\min(|F_1|^{\frac{1}{2}}, |F_2|^{\frac{1}{2}})|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}$ by $|F_1|\,|F_2|^{\frac{1}{2}}$.

By a simple rescaling argument (replacing the constants C, c by different ones) we may take $\lambda_0 = 1$. Therefore estimate (18) is now proved. The companion estimate (19) is proved likewise.

Finally, we note that Corollaries 4.2 and 4.3 are easy consequences of (18) and (19).

The authors would like to thank the referee for his/her useful suggestions and comments.

References

[1] N. Y. Antonov, *Convergence of Fourier series*, Proceedings of the XXth Workshop on Function Theory (Moscow 1995) East J. Approx. **2** (1996), 187–196.

[2] D. Bilyk and L. Grafakos, *Interplay between distributional estimates and boundedness in harmonic analysis*, Bull. Lon. Math. Soc. to appear.

[3] D. Bilyk and L. Grafakos, *Distributional estimates for the bilinear Hilbert transforms*, submitted.

[4] L. Carleson, *On convergence and growth of partials sums of Fourier series*, Acta Math. **116** (1966), 135–157.

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

[6] D. Fan and S. Sato, *Transference of certain multilinear multiplier operators*, J. Austral. Math. Soc. **70** (2001), 37–55.

[7] L. Grafakos, *Classical and Modern Fourier Analysis*, Prentice Hall, Pearson Education, Upper Saddle River NJ, 2003.

[8] L. Grafakos and P. Honzík, *Maximal transference and summability of multilinear Fourier series*, J. Austral. Math. Soc., to appear.

[9] L. Grafakos and C. Lennard, *Characterization of $L^p(R^n)$ using Gabor frames*, J. Fourier Anal. and Appl. **7** (2001), 101–126.
[10] L. Grafakos, T. Tao, and E. Terwilleger, *L^p bounds for a maximal dyadic sum operator*, Math. Zeit. **246** (2004), 321–337.
[11] L. Grafakos and T. Tao, *Multilinear interpolation between adjoint operators*, J. Funct. Anal. **199** (2003), 379–385.
[12] R. A. Hunt, *On the convergence of Fourier Series,* Orthogonal Expansions and their Continuous Analogues (Proc. Conf. Edwardsville, IL 1967), D. T. Haimo (ed), Southern Illinois Univ. Press, Carbondale IL, 235–255.
[13] M. T. Lacey *On the bilinear Hilbert transform*, Doc. Math. **1998**, Extra Vol. II, 647–656.
[14] M. T. Lacey and C. M. Thiele, *L^p bounds for the bilinear Hilbert transform,* $2 < p < \infty$, Ann. Math. **146** (1997), 693–724.
[15] M. T. Lacey and C. M. Thiele, *On Calderón's conjecture*, Ann. of Math. **149** (1999), 475–496.
[16] M. T. Lacey and C. M. Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), 361–370.
[17] C. Muscalu, T. Tao, and C. Thiele, *Multi-linear operators given by singular multipliers*, J. Amer. Math. Soc. **15** (2002), 469–496.
[18] P. Sjölin and F. Soria, *Remarks on a theorem by N.Yu. Antonov*, Studia Math. **158** (2003), 79–97.

Dmitriy Bilyk, Department of Mathematics, University of Missouri, Columbia, MO 65211, USA

E-mail address: bilykd@math.missouri.edu

Loukas Grafakos, Department of Mathematics, University of Missouri, Columbia, MO 65211, USA

E-mail address: loukas@math.missouri.edu

September 4, 2005.