Estimates for double Hilbert transforms

by

CHARLES FEFFERMAN* (Chicago Ill.)

Abstract. We prove that the truncated "singular integrals"

$$T_{\varepsilon, \delta} f(x, y) = \int \frac{\Omega_1(x)}{|x|^{n+\varepsilon}} \frac{\Omega_2(y)}{|y|^{n+\delta}} f(x - x', y - y') \, dx' \, dy'$$

converge almost everywhere as $\varepsilon, \delta \to 0$, for $f \in L^{\log^+ L}(R^{n+m})$. Here, $\Omega_1(x)$ and $\Omega_2(y)$ are smooth Calderón-Zygmund kernels on $R^n$ and $R^m$ respectively.

$I$. Introduction. Calderón-Zygmund theory deals with convolution operators whose kernels are singular at zero and infinity. To handle operators whose kernels have higher dimensional singular sets is an interesting and difficult problem about which little is known. The purpose of this paper is to extend the Calderón-Zygmund constructions for singular integrals (see [1]) to a (comparatively) simple case of an operator with a one-dimensional singular set. That operator is the double Hilbert transform, defined on functions of two variables by the equation

$$Tf(x, y) = \lim_{\eta_1, \eta_2 \to 0} \int \frac{f(x - x', y - y')}{x' y'} \, dx' \, dy'.$$

$T$ is often easy to deal with, because of the obvious factorization $T = T_{\bar{x}} T_{\bar{y}}$, where $T_{\bar{x}}$ is the Hilbert transform taken in the "$x$" variable and $T_{\bar{y}}$ is the Hilbert transform in the "$y$" variable. For example, we know that $T$ is bounded on $L^p (1 < p < \infty)$, since $\|Tf\|_p = \|T_{\bar{x}} (T_{\bar{y}} f)\|_p \leq C \|T_{\bar{x}} f\|_p \leq C \|f\|_p$.

Near $L^1$ we obtain from this argument that $Tf$ is in weak $L^1$ for $f \in L^{\log^+ L}$.

* Acknowledgement. It is a pleasure to acknowledge E. M. Stein, who pointed out to me the problem treated here, and was consistently helpful as I searched for a solution.
Actually, the factorization of $T$ works so beautifully that one has to think in order to find a non-trivial problem concerning it. One such problem concerns the “maximal operator”

$$ Af(x, y) = \sup_{\alpha, \beta > 0} \int_{|y - y'| > 2^\alpha} \left| \int_{|z - z'| > 2^\beta} f(x - z', y - y') \, dz' \, dy' \right|. $$

In one dimension, we know that the maximal Hilbert transform behaves just as well as the Hilbert transform, near $L^1$. The same is true in our situation:

**Theorem 1.** Let $f \in \operatorname{Llog}^+ L[0, 1]$, and say “support $(f) \subseteq [0, 1] \times [0, 1]$. Then $Af$ belongs to weak $L^1$ on $[0, 1] \times [0, 1]$ and

$$ \left\{ (x, y) \in [0, 1] \times [0, 1] : Af(x, y) > a \right\} \leq C \frac{1}{a} \left( \|f\|_{L^\infty} + C \right) $$

with $C$ independent of $a$ and $f$.

For a proof using complex methods, see Zygmund [4]. Our purpose here is to give a real-variable proof of Theorem 1. This forces us to extend the Calderón–Zygmund methods of [1] to the present, more singular context.

Our real-variable proof also establishes the analogue of Theorem 1 for operators of the form

$$ A^* f(x, y) = \sup_{\alpha, \beta > 0} \left| \int_{|y - y'| > 2^\alpha} \frac{\Omega_1(x')}{|x'|^m} \frac{\Omega_2(y')}{|y'|^n} f(x - x', y - y') \, dz' \, dy' \right| $$

on $E^m + m$, where $\Omega_1(x)' and $\Omega_2(y)'$ are Calderón–Zygmund kernels. It is enough to assume that $\Omega_i$ satisfies a Dini condition and $\Omega_i$ is smooth. This, of course, cannot be done by complex methods.

**II. Preliminaries.** Our proof of Theorem 1 is unfortunately rather complicated, and requires a good deal of notation and preliminary discussion. Here is some of it:

(a) Let $f$ be a function on $E^1$. For a given integer $k$, set $f_k$ equal to the average of $f$ over the dyadic interval of length $2^{-k}$ containing $x$. Define $f_k = f_k - f_{k-1}$; we obtain the Haar series of $f$. If $f = \sum f_k$, each $f_k$ is constant on dyadic intervals of length $2^{-k}$ and has average zero over dyadic intervals of length $2^{-k}$. The set of all functions with those two properties will be called $\mathcal{F}_k$. Obviously if $f_k \in \mathcal{F}_k$ and $f_{k+1} \in \mathcal{F}_{k+1}$, then $f_k$ and $f_{k+1}$ are orthogonal.

(b) Suppose that $(f_k)$ is a sequence of $L^2$ functions on $E^1$, and that the Fourier transform $f_k^*$ lives on $2^{-k-1} \leq |x| < 2^{-k}$ for each $k$. Then

$$ \left\| \sup_k \left\{ \sum_{k=0}^\infty |f_k^*| \right\} \right\| \leq C \left( \sum_{k=0}^\infty \|f_k^*\|_{L^2}^2 \right). $$

This is just a simple variant of the maximal theorem for $L^1$. To prove it, set $f = \sum_{k=0}^\infty f_k$, take a function $\varphi \in C^\infty_c (E^1)$ for which $\varphi = 1$ if $|x| \leq 1$, and $\varphi = 0$ if $|x| \geq 2$; and write $\varphi_k(x) = 2^k \varphi(2^k x)$. Then

$$ \sum_{k=0}^\infty f_k^* = \varphi \ast f^* (\varphi \ast f^* - \varphi \ast f_k^*) + \varphi \ast f_k^*. $$

The first term on the right-hand side is dominated by the maximal function of $f$, which has $L^1$-norm at most $C \|f\|_{L^1} \leq C \left( \sum_{k=0}^\infty \|f_k^*\|_{L^2}^2 \right)^{1/2}$. The second term on the right is dominated by $(\mathlarger{\sum}_{k=0}^\infty \|f_k^*\|_{L^2}^2)^{1/2} + (\mathlarger{\sum}_{k=0}^\infty \|f_k \ast f_k^*\|_{L^2}^2)^{1/2}$, which has $L^1$-norm $\leq (\mathlarger{\sum}_{k=0}^\infty \|f_k^*\|_{L^2}^2)^{1/2} + (\mathlarger{\sum}_{k=0}^\infty \|f_k \ast f_k^*\|_{L^2}^2)^{1/2} \leq C \left( \sum_{k=0}^\infty \|f_k^*\|_{L^2}^2 \right)^{1/2}$. This third term is handled similarly, completing the proof.

(c) For a function $f$ on $E^1$, we have defined $T_x f$ and $T_y f$ to be Hilbert transforms of $f$ in the $x$ and $y$ direction, respectively. Similarly, we can define $M_x f$, $M_y f$ the maximal functions in the $x$ and $y$ directions, and $A_x f$, $A_y f$, the maximal Hilbert transforms.

(d) If $J$ is any dyadic interval in $E^1$, $J^\perp$ denotes the dyadic interval containing $J$ and twice HilbertTransform.

**III. The basic construction.** As in the proof of the Calderón–Zygmund inequality, the main idea of our proof is to find an $L^1$ function $f^*$ corresponding to each $f \in \operatorname{Llog}^+ L$, so that $Af$ and $A^f$ are approximately the same. This section gives the construction of $f^*$.

Start, then, with a function $f \in \operatorname{Llog}^+ L(E^1)$ and a number $\alpha > 0$. For each fixed $y$, regard $f(x, y)$ as a function of $x$, and write $f = \sum f_k$, as defined in (a) above. Thus, for each fixed $y$, $f_k \in \mathcal{F}_k$. We abuse notation and write $f_k \in \mathcal{F}_k$. As we noted in (a), $\|f\|_{L^1} \leq C \left( \|f\|_{L^1} + C \right)$.
where \( S(f)(x, y) = \sum_{k=0}^{\infty} |f_k(x, y)|^2 \). Therefore, \( M_2(S(f)) \) belongs to weak \( L^1 \), and the set \( \Omega_4 = \{(x, y) \in \mathbb{R}^2 | M_2(S(f)) \geq a \} \) has measure

\[ |\Omega_4| \leq C \left( \|\log^* f\|_1 + C \right). \]

Using the set \( \Omega_4 \), we are going to replace each \( f_k \) by an \( \tilde{f}_k \in J_k \). Our replacement for \( f \) will be

\[ \tilde{f} = \sum_{k} f_k \tilde{f}_k \]

as follows: Let \( \alpha \) be a point of \( \mathbb{R}^2 \), and suppose that \( I \) is the dyadic interval of length \( 2^{1-k} \) containing \( \alpha \). Let \( \{J_k^I\}_{k=0}^{\infty} \) be the set of all the maximal dyadic intervals \( J \subseteq I \), for which the rectangle \( I \times J \subseteq \mathbb{R}^2 \) contains more than half its area contained in \( I \). For fixed \( x, k \), the intervals \( J_k^I \) are pairwise disjoint. The function \( f_k(x, \cdot) \) arises, simply by averaging \( f_k(x, \cdot) \) over the intervals \( J_k^I \). That is,

\[ f_k(x, y) = \frac{1}{|J_k^I|} \int_{J_k^I} f_k(x, y') dy' \quad \text{for} \quad y \in J_k^I; \]

\[ f_k(x, y) = f_k(x, y) \quad \text{for} \quad y \notin \bigcup_{I} J_k^I. \]

Clearly \( f_k \in J_k \). For future reference, define \( F_k(x, \cdot) \) by averaging \( f_k(x, \cdot) \) over the \( J_k^I \), i.e.,

\[ F_k(x, y) = \frac{1}{|J_k^I|} \int_{J_k^I} f_k(x, y') dy' \quad \text{if} \quad y \in J_k^I; \]

\[ F_k(x, y) = f_k(x, y) \quad \text{if} \quad y \notin \bigcup_{I} J_k^I. \]

Having defined replacements \( f_k \) for the \( f_k \), we can now set

\[ f = \sum_{k} f_k \]

**IV. Proof of Theorem 1.** Now that we have constructed a replacement \( f \) for \( f \), our first order of business is to show that \( f \in L^1(\mathbb{R}^2) \) and has suitable norm. Since \( f = \sum_{k} f_k \) with \( f_k \in J_k \), this amounts to showing that

\[ \sum_{k} \|f_k\|_{L^1} \leq C\left(\|\log^* f\|_1 + C\right) \]

We shall prove more:

\[ \sum_{k} \|F_k\|_{L^1} \leq C\left(\|\log^* f\|_1 + C\right), \]

Thus, \( f \in L^1 \) will follow immediately from (3).

Our first step in proving (3) is to replace \( F_k \) by a slightly smaller function \( F_k' \), defined as follows: \( F_k(x, y) = F_k(x, y) \) if

Case I. \( (a) \) \((x, y) \in \mathbb{R}^2, y \notin J_k^I, (x, y) \notin \Omega_4 \), \( J_k^I \) is then called a “case I interval” or

\( (b) \) \((x, y) \in \mathbb{R}^2, y \notin \bigcup_{I} J_k^I, (x, y) \notin \Omega_4 \), \( J_k^I \) occurs.

In the contrary case,

Case II. \( (a) \) \((x, y) \in \mathbb{R}^2, y \in J_k^I, (x, y) \notin \Omega_4 \), \( J_k^I \) is then called a “case II interval” or

\( (b) \) \((x, y) \in \mathbb{R}^2, y \in \bigcup_{I} J_k^I, (x, y) \notin \Omega_4 \),

we define \( F_k(x, y) = 0 \).

In other words, \( F_k(x, y) \) is defined to be \( F_k'(x, y) \) for some \((x, y)\), and zero for other \((x, y)\). Obviously, then, \( \|F_k\|_{L^1} \leq \|F_k'\|_{L^1} \). We claim that \( \|F_k\|_{L^1} \leq C\|F_k'\|_{L^1} \) as well. To see this, note that \( F_k \) is constant over each rectangle \( I \times J_k^I \) \((x, y, |I| = 2^{1-k})\) and over each interval \( I \times \{y\} \) with \( y \notin \bigcup_{I} J_k^I \) \((x, y, |I| = 2^{1-b})\). Thus, to prove our claim, we have only to observe that case II can occur on at most half the area (length) of any of those rectangles (intervals). This is obvious for the rectangles, since at most half the area of \( I \times J_k^I \) is contained in \( \Omega_4 \) by maximality of \( J_k^I \); and is also true for the intervals \( I \times \{y\} \) for almost all \( y \), since otherwise more than half of \( I \times J \) would lie in \( \Omega_4 \) for some small \( J \), contradicting \( y \notin \bigcup_{I} J_k^I \).

In any event, \( \|F_k\|_{L^1} \leq C\|F_k'\|_{L^1} \). Therefore, to prove (3) it will be enough to prove

\[ \sum_{k=0}^{\infty} \|F_k\|_{L^1} \leq C\left(\|\log^* f\|_1 + C\right). \]

For each \( x \in \mathbb{R}^1 \), let \( \{L_x\} \) be the collection of all the maximal dyadic intervals \( L \) with \((x, y) \in \Omega_4 \). Then set

\[ G_b(x, y) = \frac{1}{|L_x|} \int_{L_x} f_b(x, y') dy' \quad \text{for} \quad y \in L_x; \]

\[ G_b(x, y) = [f_b(x, y)] \quad \text{for} \quad y \notin \bigcup_{I} L_x. \]

We shall prove (4) by comparing \( F_k \) with \( G_b \). The comparison is not hard. For fixed \( x, F_k(x, \cdot) \) arises from \([f_k(x, \cdot)]\) by averaging over the case I intervals \( (J_k^I) \), and by setting \( F_k(x, \cdot) = 0 \) on the case II intervals and points. On the other hand, \( G_b(x, \cdot) \) arises from \([f_b(x, \cdot)]\) by averaging over certain other certain intervals \( (L_x) \). We have to understand how the \( (L_x) \) relate to the \( (J_k^I) \). It is not hard to see that for any \( x, k \), any \( L_x \), and any case I interval \( J_k^I \), we have \( L_x \subseteq J_k^I \) or else the intervals are disjoint. (This
is because \( L \) and \( J \) are dyadic intervals, so that \( L \cap J = \emptyset \), \( L \subseteq J \), or \( J \subseteq L \). \( J \subseteq L \) is impossible since by definition of case I and of \( L \), \( L' \subseteq \Omega_0 \) and \( J' \subseteq \Omega_0 \).

Therefore, for a fixed \( \alpha, \beta \), and \( y \),

\[
|F(x, y)| \leq |F^*(x, y)| = \frac{1}{|J'|} \int_{J'} G_\alpha(x, y') dy' \quad \text{if } y \in J'.
\]

If \( y \notin \bigcup J_{i+1} \), then either \( y \in \bigcup J_i \), which implies \( (x, y) \in \Omega_0 \) and case II occurs, so that \( F(x, y) = 0 \); or else \( y \notin \bigcup J_i \), in which case \( F(x, y) = G_\alpha(x, y) = |F(x, y)| \). In all cases then, \( |F(x, y)| \) is smaller than an “averaged-out” version of \( G_\alpha \). Consequently, \( \|F\|_{L^2} \leq \|G\|_{L^2} \), which is the desired comparison between \( F \) and \( G_\alpha \).

(4) now reduces to

\[
\left\| \sum_{k=0}^{m} \left| G_\alpha(x, y) \right|^2 \right\|_{L^2} \leq C a(\|f \|_{L^2} + 1).
\]

In other words

(5)

\[
\left\| \sum_{k=0}^{m} \left| G_\alpha(x, y) \right|^2 \right\|_{L^2} \leq C a(\|f \|_{L^2} + 1).
\]

So \( f \in L^2 \) reduces to (5). Mercifully, (5) can be proved directly, without further reductions, as follows: Regard \( (\theta_\alpha(x, y) = -\beta \in \mathbb{C} \) as a vector in \( F \). Then \( \theta_\alpha(x, y) \) is just an averaged-out version of \( (\theta_\alpha(x, y)) \), using the intervals \( L_i \) to average over. Thus

\[
\left\| \sum_{k=0}^{m} \left| \theta_\alpha(x, y) \right|^2 \right\|_{L^2} \leq C \left\| \theta_\alpha \right\|_{L^2} \leq C a(\|f \|_{L^2} + 1).
\]

On the other hand, for any \( x, y \), the interval \( L_i \) is defined to be big enough (i.e., \( (x, y) \times L_i \subseteq \Omega_0 \) that the average of \( \theta_\alpha(x, y) = \left( \frac{1}{|J'|} \int_{J'} \theta_\alpha(x, y) dy' \right) \) over \( L_i \) is at most \( a \). Thus, for \( y \in L_i \), \( \left\| \sum_{k=0}^{m} |G_\alpha(x, y)|^2 \right\|_{L^2} \) is at most \( a \). Similarly, \( \left\| \sum_{k=0}^{m} |G_\alpha(x, y)|^2 \right\|_{L^2} \) is at most \( a \), and in both cases \( \left\| \sum_{k=0}^{m} |G_\alpha(x, y)|^2 \right\|_{L^2} \leq C \), so that

\[
\left\| \sum_{k=0}^{m} \left| G_\alpha(x, y) \right|^2 \right\|_{L^2} \leq C a.
\]

Now we have strong control of both \( L^2 \) and \( L^\infty \) norms of

\[
\left\| \sum_{k=0}^{m} \left| G_\alpha(x, y) \right|^2 \right\|_{L^2} \leq C \left( \|f \|_{L^2} + 1 \right),
\]

which was the goal of Section IV.

V. More preliminaries. Starting with \( f \in \log^4 L(R^2) \), we now know how to find \( f' \), and we know that \( f' \in L^2(R^2) \). It remains to show that \( Af \) and \( A'f \) are approximately the same. Again, the proof is technical; we prefer to get some preliminaries out of the way before proceeding further.

(a) First of all, we shall reduce the double Hilbert transform to a more convenient form. Let \( H_k \) be a \( C^\infty \) function on \( R^2 \), equal to \( \frac{1}{a} \) for \( |x| > 1 \), and equal to zero for \( |x| < \frac{1}{2} \). For each integer \( k \), set \( H_k(x) = H_k(2^k x) \). Similarly, let \( v \) be a non-negative \( C^\infty \) function with compact support and total integral 1 on \( R^2 \), and set \( v_k(x) = v(x) \frac{1}{2^k} \).

For all practical purposes, we may replace

\[
Af(x, y) = \sup_{v_k} \left| \int \frac{f(x-a', y-y')}{a'y'} da'dy' \right|
\]

by the modified operator

\[
A_kf(x, y) = \sup_{v_k} \left| \int \frac{f(x-a', y-y')}{a'y'} H_k(x-a') \right|
\]

(For convenience, we write \( v_k \frac{1}{2^k} \) for the above integral.) To show that the two transforms are approximately the same, write

\[
\int \int \frac{f(x-a', y-y')}{a'y'} da'dy' = \left( \int \int f(x-a', y-y') da'dy' \right) \frac{1}{a'y'}
\]

and

\[
= \left[ \int \int v_k \frac{1}{2^k} \left( \frac{1}{y'} \right) f(x-a', y-y') da'dy' \right] + \left[ \int \int \left( \frac{1}{2^k} \right) f(x-a', y-y') da'dy' \right]
\]

Now, we have

\[
= \left[ \int \int v_k \frac{1}{2^k} \left( \frac{1}{y'} \right) f(x-a', y-y') da'dy' \right] + \left[ \int \int \left( \frac{1}{2^k} \right) f(x-a', y-y') da'dy' \right]
\]

for all \( k \).
The first term in brackets is dominated by $M_4(T_y e)(x, y)$ (see section II (0), which is certainly in weak $L^1$ if $f \in L\log L$). Similarly, if we take that $k$ for which $2^{-k+1} < r_k \leq 2^{-k}$, the second term in brackets is dominated by $M_4(T_y f)(x, y)$. So, in fact, $Af$ and $A_4 f$ are essentially the same.

(b) Let us analyze $H_2$. Writing $K_2 = H_2 - H_{2-1}$, we obtain $H_2 = \sum_{k=1}^{\infty} K_k$. Pick functions $\varphi_k$ on the real line, with the following properties: (i) $\varphi_k(x)$ is supported in $2^{2k-1} \leq |x| < 2^{2k+1}$, (ii) $\sum_{k=1}^{\infty} \varphi_k(x) = 1$, and (iii) $\left| \frac{\partial^{a_j} \varphi_k}{\partial x^{a_j}} (\xi) \right| = O(2^{-a_j})$ uniformly in $a_j$ for each $\alpha$. By property (ii), each $K_2$ is equal to $\sum_{k=1}^{\infty} K_{2k} \varphi_k = \sum_{k=1}^{\infty} K_{2k}$ where by definition $K_{2k} = K_{2k} * \varphi_k$. We shall need a little precise information on the size and behavior of the $K_{2k}$. It is not very hard to convince oneself of the following:

If $j > 0$, then $K_{2j}$ is a $C^\infty$ function with “thickness” $2^{-2}k$ and $L^\infty$-norm roughly $2^{-2k}$. In fact $\left| \frac{\partial^{a_j} \varphi_k}{\partial x^{a_j}} (\xi) \right| = O(2^{-a_j})$ uniformly in $a_j$ and $k$. If $j < 0$, then $K_{2j}$ is a $C^\infty$ function with “thickness” $2^{-k}$ and $L^\infty$-norm smaller than $2^{k}$. In fact $\left| \frac{\partial^{a_j} \varphi_k}{\partial x^{a_j}} (\xi) \right| = O(2^{-k})$ uniformly in $a_j$ and $k$.

From (6) we deduce easily that $\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \left| \frac{\partial^{a_j} \varphi_k}{\partial x^{a_j}} (\xi) \right| = O \left( \frac{2^{-j}}{(2^{-\infty} + |x|)^{1/2}} \right)$ uniformly in $x, m$, and $j$, provided $j \geq 0$. If $j < 0$, we obtain similarly from (7) that $\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \left| \frac{\partial^{a_j} \varphi_k}{\partial x^{a_j}} (\xi) \right| = O \left( \frac{2^{-j}}{(2^{-\infty} + |x|)^{1/2}} \right)$. As a consequence, we may apply the proof of the standard Calderón-Zygmund lemma (8), to prove inequality (9).

**Lemma**. Suppose that $f = \sum f_j$, where each $f_j$ lives on an interval $J$ with center $x_j$, and has average zero. Then outside the union of the doubles of the $J_j$'s, the maximal Hilbert transform of $f$ is dominated by the Marcinkiewicz integral

$$\sum_{j=1}^{\infty} \frac{\|f\|_{L^1} \chi_{J_j}}{(|J_j| + |x-x_j|)^{1/2}} \tag{8}$$

**Lemma.** Suppose that $f = \sum f_j$, where each $f_j$ lives on an interval $J_j$ and has average zero. Say that $J_j$ has length $2^{-k}$ and center $x_j$. Then for each integer $j$

$$\sum_{j=1}^{\infty} \frac{\|f_j\|_{L^1} \chi_{J_j}}{(|J_j| + |x-x_j|)^{1/2}} \leq C \cdot \sum_{j=1}^{\infty} \frac{\|f_j\|_{L^1} \chi_{J_j}}{(2^{-j} + |x-x_j|)^{1/2}} \tag{9}$$

**VI. Proof of Theorem 1, continued.** In this section we show that $Af$ and $A_4 f$ are roughly equal. By the results of V (a) this is equivalent to estimating $\left( \frac{1}{x} \right)^{1/2} \sum_{n \in \mathbb{N}} K_n^* * H_n^* (f - f^*) (x, y, z)$ for all $x$ and $y$. We have

$$H_n^* (f - f^*) = \sum_{n \in \mathbb{N}} K_n^* (f - f^*) \quad \text{for technical reasons, it will be expedient to split the sum into } \sum_{n \in \mathbb{N}} \quad \text{and } \sum_{n \in \mathbb{N}} \quad \text{and henceforth any sum on the index } n \text{ will be implicitly assumed to run only over } n \text{ of one fixed parity. At the end of the section, we shall then have proved estimates for } A_4^\text{even}(f-f^*), \quad \text{and } A_4^\text{odd}(f-f^*), \quad \text{defined similarly.}$$

Then we can write $A_4 (f-f^*) = A_4^\text{even}(f-f^*) + A_4^\text{odd}(f-f^*)$, to deduce the estimates we really want. With this minor embarrassment out of the way, we can proceed.

We have to estimate $\left( \frac{1}{x} \right)^{1/2} \sum_{n \in \mathbb{N}} K_n^* * H_n^* (f - f^*)$. Recall that $f = \sum f_k$ and $f^* = \sum f_k$ where $f_k \in f_k^* \in F_k$. We shall break up $H_n^* (f - f^*)$ into two parts. To do so, let $J_k = f_k - f_k^* = f_k - f_k \cdot \mathbb{1}_{J_k \cap J_k^*}$ with $x_k \in J_k$, $|J_k| \leq 2^{-k}$. Thus, for fixed $x_k$, $f_k(x, y, z)$ lives on the interval $J_k^*$. By the Marcinkiewicz integral, $f_k(x, y, z)$ has average zero, and satisfies $\|f_k(x, y, z)\|_{L^1} \leq |E_k(x, y, z)|^{1/2}$. Say also that $J_k^*$ has length $2^{-k/2}$ and center $y(x, z, k)$. Now

$$H_n^* (f_k - f_k^*) (x, y, z) = \sum_{x \in J_k^*} H_n^* f_k (x, y, z) = \sum_{x \in J_k^*} K_n^* f_k (x, y, z)$$

$$= \sum_{n \in \mathbb{N}} \sum_{x \in J_k \cap J_k^*} K_n^* f_k (x, y, z) + 2 \sum_{n \in \mathbb{N}} \sum_{x \in J_k \cap J_k^*} K_n^* f_k (x, y, z)$$

$$= H_n f_k (x, y, z) + H_n f_k (x, y, z)$$
Outside an exceptional set, we will be able to estimate both \( p \ast f \ast \mathcal{H}_{\alpha} f \) and \( f \ast p \ast \mathcal{H}_{\alpha} f \). Our estimates will be so sharp that we will be able to combine the results for different \( \alpha \) to estimate the maximal double Hilbert transform of \((f \ast f')\). 

(a) The term \( H_{\alpha} f \): We have

\[
H_{\alpha} f(x, \cdot) = \sum_{n=\infty}^{\infty} \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) = \sum_{n=\infty}^{\infty} K_{\alpha}^n \ast f(x, \cdot).
\]

Set \( H_{\alpha} f(x, \cdot) = \sum_{n=\infty}^{\infty} K_{\alpha}^n \ast f(x, \cdot) \). By (9),

\[
\|H_{\alpha} f(x, y)\| \leq C \cdot 2^{-m} \cdot \sum_{n=\infty}^{\infty} \|F_{\alpha} f(x, y)\|_{L^2} \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

By a well-known inequality for Marcinkiewicz integrals (see [2]) \( \|H_{\alpha} f\|_{L^1} \leq C \cdot 2^{-m} \|f\|_{L^1} \). On the other hand, \( H_{\alpha} f \in L^1 \), so that for different \( \alpha \), the \( H_{\alpha} f \) are orthogonal. Hence

\[
\left\| \sum_{n=\infty}^{\infty} H_{\alpha} f(x, \cdot) \right\|_{L^1} \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

by inequality (3). Since \( T_{\alpha} \), the Hilbert transform in the variable, is bounded on \( L^p \), it follows that

\[
\left\| \sum_{n=\infty}^{\infty} T_{\alpha}(H_{\alpha} f(x, \cdot)) \right\|_{L^1} \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

Recalling definitions, we have

\[
\sum_{n=\infty}^{\infty} T_{\alpha}(H_{\alpha} f(x, \cdot)) = \sum_{n=\infty}^{\infty} T_{\alpha} \left( \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) \right) = \sum_{n=\infty}^{\infty} \sum_{x \in a_n} T_{\alpha} K_{\alpha}^n \ast f(x, \cdot) = \sum_{n=\infty}^{\infty} R_n^\alpha.
\]

So we have proved that \( \sum_{n=\infty}^{\infty} R_n^\alpha \leq C \cdot 2^{-m} \|f\|_{L^1} + C \). On the other hand, the term \( K_{\alpha}^n \) in \( R_n^\alpha \) ensures that the Fourier transform, taken in the \( y \)-variable, of \( R_n^\alpha \) lives in the region \( 2^{j+1} \leq |y| \leq 2^{j+3} \). Because of our convention on \( \sum_{n=\infty}^{\infty} \) and \( \sum_{n=\infty}^{\infty} \), these regions are pairwise disjoint for fixed \( j \), and so the lemma of II (b) shows that

\[
\max_{n} \left\{ \sum_{n=\infty}^{\infty} R_n^\alpha \right\} \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

What does that mean? We know that

\[
\sum_{n=\infty}^{\infty} T_{\alpha} K_{\alpha}^n \ast f(x, \cdot) = \sum_{n=\infty}^{\infty} T_{\alpha} \left( \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) \right) = \sum_{n=\infty}^{\infty} R_n^\alpha.
\]

So our \( L^1 \)-inequality means simply that

\[
\left( \sum_{n=\infty}^{\infty} T_{\alpha} \left( \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) \right) \right) \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

Now applying the maximal theorem for \( L^1 \), we get

\[
\left\| \sum_{n=\infty}^{\infty} T_{\alpha} \left( \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) \right) \right\|_{L^1} \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

Summing over \( j \) yields

\[
\left\| \sum_{n=\infty}^{\infty} T_{\alpha} \left( \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) \right) \right\|_{L^1} \leq C \cdot 2^{-m} \|f\|_{L^1} + C,
\]

which implies that

\[
\left( \sum_{n=\infty}^{\infty} \left( \sum_{x \in a_n} K_{\alpha}^n \ast f(x, \cdot) \right) \right) \leq C \cdot 2^{-m} \|f\|_{L^1} + C.
\]

This is exactly the estimate we need for the contribution of the \( H_{\alpha} f \) to \( A_{\alpha}(f \ast f') \).
By the techniques we just used, we could also have proved that
\[ \| \sup_m \sum_{k=-\infty}^{\infty} H_m f_k \|_1 \leq C \alpha \| \log f \|_{1+\alpha} + C. \]

(b) The term \( H_m f_k \) will convince the reader that \( H_m f_k \) lives only on rectangles \( R \times J_{\alpha} \), where \( R \) is dyadic of length \( 2^{m+1} \), \( J_{\alpha} \) is an interval concentric with \( J \) and five times as large, and \( J = J_{\alpha}^{k,\alpha} \) for all \( k \). By definition of \( J_{\alpha}^{k,\alpha} \) such rectangles have at least one tenth of their area inside \( \Omega \). Consequently, \( H_m f_k \) lives entirely on the set \( \Omega \) at all points of \( B^k \) at which the maximal function of \( \chi_k \) exceeds \( 1/10 \). Furthermore, \( H_m f_k \) always has average zero over the component intervals of \( \Omega \) on \( B^k \times (y_k) \), for all \( y_k \in B^k \). So by inequality (8) above, \( \sup_{k \in \mathbb{Z}} \left( \frac{\bar{f}}{\alpha} \right) \sum_{k=-\infty}^{\infty} H_m f_k \) is dominated outside \( \Omega \) by a Marcinkiewicz integral (in the \( x \) variable) of the function \( \sum_{k=-\infty}^{\infty} H_m f_k \). Here, \( \Omega \) denotes the union of all the intervals \( I^k \times (y_k) \), where \( I^k \) is a component interval of \( x \in \mathbb{R} \) (of size \( y_k \)). Therefore, \( \sup_{k \in \mathbb{Z}} \left( \frac{\bar{f}}{\alpha} \right) \sum_{k=-\infty}^{\infty} H_m f_k \) is dominated outside \( \Omega \) by a Marcinkiewicz integral (again in the \( x \) variable) of \( \sum_{k=-\infty}^{\infty} H_m f_k \). It follows then from the \( L' \) boundedness of the Marcinkiewicz integrals (see [9]) that
\[ \left( \frac{\bar{f}}{\alpha} \right) \sum_{k=-\infty}^{\infty} H_m f_k > \frac{1}{2} \right) \sum_{k=-\infty}^{\infty} H_m f_k \right) \leq C \| \sup_{k \in \mathbb{Z}} \left( \frac{\bar{f}}{\alpha} \right) \sum_{k=-\infty}^{\infty} H_m f_k \|_1. \]

So to finish off our weak-type estimate for \( \sup_{k \in \mathbb{Z}} \left( \frac{\bar{f}}{\alpha} \right) \sum_{k=-\infty}^{\infty} H_m f_k \), we need only show that
\[ \| \sum_{k=-\infty}^{\infty} H_m f_k \|_1 \leq C \| \log f \|_{1+\alpha} + C. \]

To prove this inequality, write \( H_m f_k = H_m f_k \circ (f - f) - H_m f_k \), so that
\[ \sum_{k=-\infty}^{\infty} H_m f_k = H_m f_k - \sum_{k=-\infty}^{\infty} H_m f_k. \]

Now by standard results on the maximal Hilbert transform,
Since also $|Q| \leq C \log^+ f_1 + C$, we have

$$|[(x, y) \in \mathbb{R}^2: A_\alpha(f - f^*) < a]| \leq C \log^{+} f_1 + C.$$  

Thus $A_\alpha(f)$ and $A_\alpha(f^{-})$ are approximately equal.

**VII. Proof of Theorem 1: Mop-Up.** We have completed our program of constructing a $f^*$, showing that $f^* \in L^1$, and proving that $A(f) \approx A(f^*)$. The proof of Theorem 1 is now a triviality. Clearly

$$|[(x, y) \in \mathbb{R}^2: A(f)(x, y) > 2a]| \leq |[(x, y) \in \mathbb{R}^2: A(f^*)(x, y) > a]|$$

$$+ |[(x, y) \in \mathbb{R}^2: |A(f)(x, y) - A(f^*)|(x, y) > a]|.$$

The first term is at most $C \log^+ f_1$, by the Chebyshev inequality, the $L^2$-boundedness of $A$, and our estimate for $\|f^*\|_1$. The second term is at most $C (\log^+ f_1 + C)$ by inequality (17). Thus, $|[(x, y) \in \mathbb{R}^2: A(f)(x, y) > 2a]| \leq C \log^+ f_1 + C$. The proof of Theorem 1 is complete.

**VIII. Remarks.** (a) There must be a simpler way to do this.

(b) Note that in order to find a non-trivial proof, we have to ask for Theorem 1 in its full strength. For instance, suppose we merely want to know that $A_\alpha(f)$ is in weak $L^1$ for $f \in L^{1, \infty}$, where

$$A_\alpha(f)(x, y) = \sup_{\delta > 0} \int_{|x-y| > \delta} f(x-y, y-y') \, dy' \, dy.$$

This is a just semi-trivial observation, which we prove as follows: Take any function $\theta(x, y)$ on $\mathbb{R}^2$, with the properties

1. $\theta$ is homogeneous of degree zero on $\mathbb{R}^2$ and $C^\infty$ on $\mathbb{R}^2 - \{0\}$.

2. $\theta(x, y) = 1$ for $|x| \leq \frac{1}{3} |y|$, except at the origin.

3. $\theta(x, y) = 0$ for $|y| \leq \frac{1}{3} |x|$, except at the origin.

Then if $T$ is the double Hilbert transform, we have

$$[Tf(x, y)]^- = \text{sgn}(x) \text{sgn}(y) f^*(x, y)$$

and

$$[Tf(x, y)]^- = \text{sgn}(x) \cdot [\theta(x, y) \text{sgn}(y)] f^*(x, y) + \text{sgn}(x) \cdot [1 - \theta(x, y)] \sin \theta(x, y) \cdot f(x, y)$$

$$= \text{sgn}(x) \cdot K^+ f^*(x, y) + \text{sgn}(y) \cdot K^+ f^*(x, y).$$

where $K^+$ and $K^-$ are Calderón–Zygmund kernels on $\mathbb{R}^2$. In other words, $T_f = T_\alpha(K_1 f) + T_\alpha(K_2 f)$, where $T_\alpha$ and $T_\beta$ are as in (2). Had we taken $\delta$ into account in the above calculation, we would have found $A_\alpha f \leq A_\alpha(K_1 f) + A_\alpha(K_2 f)$ and trivial error terms, where $A_\alpha$ and $A_\beta$ are the maximal Hilbert transforms in the $x$ and $y$ directions, respectively. Therefore

$$\|A_\alpha f\|_{\text{weak} L^1} \leq C \|K_1 f\|_1 + C \|K_2 f\|_1 \leq C \|f\|_{L^{1, \infty}}. \quad \square$$

**References**


Received December 28, 1970