STUDIA MATHEMATICA, T. XLIV. (1972)

Estimates for double Hilbert transforms

by

CHARLES FEFFERMAN* (Chicago Ill.)

Abstract. We prove that the truncated "singular integrals"

$$T_{\varepsilon\delta}f(x,y) = \iint\limits_{\substack{|x'|>\varepsilon\\ |y'|>\delta}} \frac{\Omega_1(x')}{|x'|^n}\,\frac{\Omega_2(y')}{|y'|^m}\,f(x-x',\,y-y')\,dx'\,dy'$$

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

$Tf = \lim\limits_{\varepsilon,\delta\to 0} T_{\varepsilon\delta}f$ is the simplest convolution operator whose kernel has a continuum of singular points.

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

$$Tf(x,y) = \lim_{\varepsilon_1,\varepsilon_2\to 0} \iint\limits_{\substack{|x'|>\varepsilon_1\\ |y'|>\varepsilon_2}} \frac{f(x-x',\,y-y')}{x'y'}\,dx'\,dy'. \tag{1}$$

T is often easy to deal with, because of the obvious factorization $T = T_xT_y$, where T_x is the Hilbert transform taken in the "x" variable and T_y is the Hilbert transform in the "y" variable. For example, we know that T is bounded on L^p $(1<p<\infty)$, since $\|Tf\|_p = \|T_x(T_yf)\|_p \leqslant C\|T_yf\|_p \leqslant C'\|f\|_p$. 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"

$$(2) \qquad Af(x,y) = \sup_{\varepsilon_1,\varepsilon_2>0} \Big| \iint_{\substack{|x'|>\varepsilon_1 \\ |y'|>\varepsilon_2}} \frac{f(x-x', y-y')}{x'y'}\, dx'\,dy' \Big|.$$

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 $f \in L\log^+ L(R^2)$, and say support $(f) \leqslant [0,1]\times[0,1]$. Then Af belongs to weak L^1 on $[0,1]\times[0,1]$ and*

$$\big|\{(x,y)\in[0,1]\times[0,1] \mid Af(x,y) > \alpha\}\big| \leqslant \frac{C}{\alpha}\,(\|f\log^+ f\|_1 + C)$$

with C independent of α 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

$$A^*f(x,y) = \sup_{\varepsilon_1,\varepsilon_2>0} \Big| \iint_{\substack{|x'|>\varepsilon_1 \\ |y'|>\varepsilon_2}} \frac{\Omega_1(x')}{|x'|^m}\,\frac{\Omega_2(y')}{|y'|^n} f(x-x', y-y')\,dx'\,dy' \Big|$$

on R^{n+m}, where $\dfrac{\Omega_1(x)}{|x|^m}$ and $\dfrac{\Omega_2(y)}{|y|^n}$ are Calderón–Zygmund kernels. It is enough to assume that Ω_1 satisfies a Dini condition and Ω_2 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 R^1. For a given integer k, set $\bar{f}^k$ equal to the average of f over the dyadic interval of length 2^{-k} containing x. Defining $f_k = \bar{f}^k - \bar{f}^{k-1}$, we obtain the Haar series of f, $f = \sum_{k=-\infty}^{\infty} f_k$. Each f_k is constant on dyadic intervals of length 2^{-k} and has average zero over dyadic intervals of length 2^{1-k}. The set of all functions with those two properties will be called $\mathscr{I}_k$. Obviously if $f_1 \in \mathscr{I}_{k_1}$ and $f_2 \in \mathscr{I}_{k_2}$ $(k_1 \neq k_2)$, then f_1 and f_2 are orthogonal.

If $f \in L\log^+ L[0,1]$, then we have the "Littlewood–Paley" inequality

$$\|Sf\|_1 \leqslant C(\|f\log^+ f\|_1 + C), \qquad \text{where } Sf(x) \equiv \Big(\sum^{\infty} |f_k(x)|^2\Big)^{1/2}.$$

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

$$\Big\|\sup_m \Big|\sum_{k=-\infty}^{m} f^k(\cdot)\Big|\Big\|_2 \leqslant C\cdot\Big(\sum_{k=-\infty}^{\infty} \|f^k\|_2^2\Big)^{1/2}.$$

This is just a simple variant of the maximal theorem for L^2. To prove it, set $f = \sum_{k=-\infty}^{\infty} f^k$, take a function $\varphi_0 \in C^\infty(R^1)$ for which $\hat{\varphi_0}(x) = 1$ if $|x| \leqslant 1$, and $\hat{\varphi_0}(x) = 0$ if $|x| \geqslant 2$; and write $\varphi_k(x) \equiv 2^k\varphi_0(2^k x)$. Then

$$\sum_{k=-\infty}^{m} f^k(x) = \varphi_m * f(x) + \big(f^m(x) - \varphi_m * f^m(x)\big) - \varphi_m * f^{m+1}(x).$$

The first term on the right-hand side is dominated by the maximal function of f, which has L^2-norm at most $C\|f\|_2 \leqslant C\big(\sum_{k=-\infty}^{\infty} \|f^k\|_2^2\big)^{1/2}$. The second term on the right is dominated by $\big(\sum_{m=-\infty}^{\infty} |f^m(x)|^2\big)^{1/2} + \big(\sum_{m=-\infty}^{\infty} |\varphi_m * f^m(x)|^2\big)^{1/2}$, which has L^2-norm $\leqslant \big(\sum_{k=-\infty}^{\infty} \|f^k\|_2^2\big)^{1/2} + \big(\sum_{k=-\infty}^{\infty} \|\varphi_k * f^k\|_2^2\big)^{1/2} \leqslant C\big(\sum_{k=-\infty}^{\infty} \|f_k\|_2^2\big)^{1/2}$ since $\|\varphi_k\|_1 \leqslant C$. The third term is handled similarly, completing the proof.

(c) For a function f on R^2, 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 R^1, $J^\sim$ denotes the dyadic interval containing J and twice as large.

III. The basic construction. As in the proof of the Calderón–Zygmund inequality, the main idea of our proof is to find an L^2 function $f^\sim$ corresponding to each $f \in L\log^+ L$, so that Af and $Af^\sim$ are approximately the same. This section gives the construction of $f^\sim$.

Start, then, with a function $f \in L\log^+ L(R^2)$ and a number $\alpha > 0$. For each fixed y, regard $f(x,y)$ as a function of x, and write $f = \sum_{k=-\infty}^{\infty} f_k$, as defined in II (a) above. Thus, for each fixed y, $f_k(\cdot, y) \in \mathscr{I}_k$. We abuse notation and write $f_k \in \mathscr{I}_k$. As we noted in II (a), $\|S(f)\|_1 \leqslant C(\|f\log^+ f\|_1 + C)$, where $S(f)(x,y)=\left(\sum_{k=-\infty}^{\infty}|f_k(x,y)|^2\right)^{1/2}$. Therefore, $M_y(S(f))$ belongs to weak L^1, and the set $\Omega_0=\{(x,y)\in R^2\,|\,M_y(Sf)(x,y)>\alpha\}$ has measure $|\Omega_0|\leqslant\frac{C}{\alpha}(\|f\log^+ f\|_1+C)$.

Using the set Ω_0, we are going to replace each f_k by an $f_k^{\sim}\in\mathscr{I}_k$. Our replacement for f will be $f^{\sim}=\sum_{k=-\infty}^{\infty}f_k^{\sim}\cdot f_k^{\sim}$ is defined as follows: Let x be a point of R^1, and suppose that I is the dyadic interval of length 2^{1-k} containing x. Let $\{J_l^{xk}\}_{l<\infty}$ be the set of all the maximal dyadic intervals $J\subseteq R^1$, for which the rectangle $I\times J\subseteq R^2$ has more than half its area contained in Ω_0. For fixed x, k, the intervals J_l^{xk} are pairwise disjoint. The function $f_k^{\sim}(x,\cdot)$ arises, simply by averaging $f_k(x,\cdot)$ over the intervals J_l^{xk}. That is,

$$f_k^{\sim}(x,y)\equiv\frac{1}{|J_l^{xk}|}\int_{J_l^{xk}}f_k(x,y')\,dy'\quad\text{ for } y\in J_l^{xk};$$

$$f_k^{\sim}(x,y)\equiv f_k(x,y)\quad\text{ for } y\notin\bigcup_l J_l^{xk}.$$

Clearly $f_k^{\sim}\in\mathscr{I}_k$. For future reference, define $F_k^{\sim}(x,\cdot)$ by averaging $|f_k(x,\cdot)|$ over the $\{J_l^{xk}\}$, i.e.,

$$F_k^{\sim}(x,y)\equiv\frac{1}{|J_l^{xk}|}\int_{J_l^{xk}}|f_k(x,y')|\,dy'\quad\text{ if }\quad y\in J_l^{xk},$$

$$F_k^{\sim}(x,y)\equiv|f_k(x,y)|\quad\text{ if }\quad y\notin\bigcup_l J_l^{xk}.$$

Having defined replacements $f_k^{\sim}$ for the f_k, we can now set

$$f^{\sim}\equiv\sum_{k=-\infty}^{\infty}f_k^{\sim}.$$

IV. Proof of Theorem 1. Now that we have constructed a replacement $f^{\sim}$ for f, our first order of business is to show that $f^{\sim}\in L^2(R^2)$ and has suitable norm. Since $f^{\sim}=\sum_{k=-\infty}^{\infty}f_k^{\sim}$ with $f_k^{\sim}\in\mathscr{I}_k$, this amounts to showing that $\sum_k\|f_k^{\sim}\|_2^2\leqslant C\alpha(\|f\log^+f\|_1+C)$. We shall prove more:

$$\sum_{k=-\infty}^{\infty}\|F_k^{\sim}\|_2^2\leqslant C\alpha(\|f\log^+f\|_1+C),\tag{3}$$

Thus, $f^{\sim}\in L^2$ will follow immediately from (3).

Our first step in proving (3) is to replace $F_k^{\sim}$ by a slightly smaller function F_k, defined as follows: $F_k(x,y)\equiv F_k^{\sim}(x,y)$ if

Case I. (a) $(x,y)\in R^2$, $y\in J_l^{xk}$, $\{x\}\times J_l^{\sim xk}\nsubseteq\Omega_0$ (J_l^{xk} is then called a "case I interval") or

(b) $(x,y)\in R^2$, $y\notin\bigcup_l J_l^{xk}$, $(x,y)\notin\Omega_0$ occurs.

In the contrary case,

Case II. (a) $(x,y)\in R^2$, $y\in J_l^{xk}$, $\{x\}\times J_l^{xk}\subseteq\Omega_0$ ($J_l^{\sim xk}$ is then called "a case II interval") or

(b) $(x,y)\in R^2$, $y\notin\bigcup_l J_l^{xk}$, $(x,y)\in\Omega_0$,

we define $F_k(x,y)\equiv 0$.

In other words, $F_k(x,y)$ is defined to be $F_k^{\sim}(x,y)$ for some (x,y), and zero for other (x,y). Obviously, then, $\|F_k\|_2\leqslant\|F_k^{\sim}\|_2$. We claim that $\|F_k^{\sim}\|_2\leqslant C\|F_k\|_2$ as well. To see this, note that $F_2^{\sim}$ is constant over each rectangle $I\times J_l^{xk}$ ($x\in I$, $|I|=2^{1-k}$) and over each interval $I\times\{y\}$ with $y\notin\bigcup_l J_l^{xk}$ ($x\in I$, $|I|=2^{1-k}$). 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 $I\times J_l^{\sim xk}$ is contained in Ω_0 by maximality of J_l^{xk}; 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 Ω_0 for some small J, contradicting $y\notin\bigcup_l J_l^{xk}$.

In any event, $\|F_k^{\sim}\|_2\leqslant C\|F_k\|_2$. Therefore, to prove (3) it will be enough to prove

$$\sum_{k=-\infty}^{\infty}\|F_k\|_2^2\leqslant C\alpha(\|f\log^+f\|_1+C).\tag{4}$$

For each $x\in R^1$, let $\{L_j^x\}$ be the collection of all the maximal dyadic intervals L with $\{x\}\times L^{\sim}\subseteq\Omega_0$.

Then set

$$G_k(x,y)\equiv\frac{1}{|L_j^x|}\int_{L_j^x}|f_k(x,y')|\,dy'\quad\text{ for } y\in L_j^x,$$

$$G_k(x,y)\equiv|f_k(x,y)|\quad\text{ for } y\notin\bigcup_j L_j^x.$$

We shall prove (4) by comparing F_k with G_k. 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 $\{J_l^{xk}\}$, and by setting $F_k(x,\cdot)\equiv 0$ on the case II intervals and points. On the other hand, $G_k(x,\cdot)$ arises from $|f_k(x,\cdot)|$ by averaging over certain other intervals $\{L_j^x\}$. We have to understand how the $\{L_j^x\}$ relate to the $\{J_l^{xk}\}$. It is not hard to see that for any x, k, any L_j^x, and any case I interval J_l^{xk}, we have $L_j^x\subseteq J_l^{xk}$ 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, $L^{\sim}\subseteq\Omega_0$ and $J^{\sim}\nsubseteq\Omega_0$.)

Therefore, for a fixed x, k, and y,

$$|F_k(x,y)|\leqslant|F_k^{\sim}(x,y)|=\frac{1}{|J_l^{xk}|}\int_{J_l^{xk}}G_k(x,y')\,dy'\quad\text{ if } y\in J_l^{xk}.$$

If $y\notin\bigcup_l J_l^{xk}$, then either $y\in$ some L_j^x, which implies $(x,y)\in\Omega_0$ and case II occurs, so that $F_k(x,y)=0$; or else $y\notin\bigcup_j L_j^x$, in which case $F_k(x,y)=G_k(x,y)=|f_k(x,y)|$. In all cases then, $|F_k(x,\cdot)|$ is smaller than an "averaged-out" version of G_k. Consequently, $\|F_k\|_2\leqslant\|G_k\|_2$, which is the desired comparison between F_k and G_k. (4) now reduces to $\sum_{k=-\infty}^{\infty}\|G_k\|_2^2\leqslant C\alpha(\|f\log^+f\|_1+C)$, or in other words

$$\Big\|\Big(\sum_{k=-\infty}^{\infty}|G_k(x,y)|^2\Big)^{1/2}\Big\|_2^2\leqslant C\alpha(\|f\log^+f\|_1+C). \tag{5}$$

So $f^{\sim}\in L^2$ reduces to (5). Mercifully, (5) can be proved directly, without further reductions, as follows: Regard $\big(G_k(x,y)\big)_{-\infty<k<\infty}$ as a vector in l^2. Then $\big(G_k(\cdot,\cdot)\big)$ is just an averaged-out version of $\big(|f_k(\cdot,\cdot)|\big)$, using the intervals $\{L_j^x\}$ to average over. Thus

$$\Big\|\Big(\sum_{k=-\infty}^{\infty}|G_k(\cdot,\cdot)|^2\Big)^{1/2}\Big\|_1\leqslant\Big\|\Big(\sum_{k=-\infty}^{\infty}|f_k(\cdot,\cdot)|^2\Big)^{1/2}\Big\|_1\equiv\|S(f)\|_1\leqslant C(\|f\log^+f\|_1+C).$$

On the other hand, for any x, j, the interval L_j^x is defined to be big enough (i.e., $\{x\}\times\tilde{L}_j^{x}\nsubseteq\Omega_0$) that the average of $Sf(x,y)=\big(\sum_{k=-\infty}^{\infty}|f_k(x,y)|^2\big)^{1/2}$ over L_j^x is at most $C\alpha$. Thus, for $y_0\in L_j^x$, $\big(\sum_{x=-\infty}^{\infty}|G_k(x,y_0)|^2\big)^{1/2}$, the l^2-norm of the average of the $|f_k(x,\cdot)|$ over L_j^x, is at most $C\alpha$, the average of the l^2-norm. Similarly, $\big(\sum_{k=-\infty}^{\infty}|G_k(x,y_0)|^2\big)^{1/2}\equiv\big(\sum_{k=-\infty}^{\infty}|f_k(x,y_0)|^2\big)^{1/2}\equiv Sf(x,y_0)\leqslant C\alpha$ for $y_0\notin\bigcup_j L_j^x$, since then $(x,y_0)\notin\Omega_0$. In both cases $\big(\sum_{k=-\infty}^{\infty}|G_k|^2\big)^{1/2}\leqslant C\alpha$, so that $\big\|\big(\sum_{k=-\infty}^{\infty}|G_k(\cdot,\cdot)|^2\big)^{1/2}\big\|_\infty\leqslant C\alpha$.

Now we have strong control of both L^1 and L^∞ norms of $\big(\sum_{k=-\infty}^{\infty}|G_k(\cdot,\cdot)|^2\big)^{1/2}$, and (5) follows at once. This proves that $f^{\sim}\in L^2$ and $\|f\|_2^2\leqslant C\alpha(\|f\log^+f\|_1+C)$, which was the goal of Section IV.

V. More preliminaries. Starting with $f\in L\log^+L(R^2)$, we now know how to find and $f^{\sim}$, and we know that $f^{\sim}\in L^2(R^2)$. It remains to show that Af and $Af^{\sim}$ are approximately the same. Again, the proof is technical, and 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_0 be a C^∞ function on R^1, equal to $\frac{1}{x}$ for $|x|\geqslant 1$, and equal to zero for $|x|\leqslant\frac{1}{2}$. For each integer k, set $H_k(x)\equiv 2^kH_0(2^kx)$. Similarly, let φ_1 be a non-negative C^∞ function with compact support and total integral 1 on R^1, and set $\varphi_\varepsilon(x)=\varepsilon^{-1}\varphi_1\Big(\frac{x}{\varepsilon}\Big)$.

For all practical purposes, we may replace

$$Af(x,y)=\sup_{\varepsilon_1,\varepsilon_2}\Big|\iint_{\substack{|x'|>\varepsilon_1\\|y'|>\varepsilon_2}}\frac{f(x-x',y-y')}{x'y'}\,dx'dy'\Big|$$

by the modified operator

$$A_0f(x,y)=\sup_{\substack{\varepsilon>0\\k}}\Big|\iint_{R^2}\Big(\varphi_\varepsilon*\frac{1}{x}\Big)(x')\cdot H_k(y')\cdot f(x-x',y-y')\,dx'dy'\Big|. \tag{6}$$

(For convenience, we write $\Big(\varphi_\varepsilon*\frac{1}{x}\Big)^x*H_k^y*f$ for the above integral.) To show that the two transforms are about the same, write

$$\iint_{\substack{|x'|>\varepsilon_1\\|y'|>\varepsilon_2}}\frac{f(x-x',y-y')}{x'y'}\,dx'dy'-\Big(\varphi_\varepsilon*\frac{1}{x}\Big)^x*H_k^y*f(x,y) \tag{7}$$

$$=\Big[\iint_{\substack{|x'|>\varepsilon_1\\|y'|>\varepsilon_2}}\frac{f(x-x',y-y')}{x'y'}\,dx'dy'$$

$$-\iint_{|y'|>\varepsilon_2}\Big(\varphi_\varepsilon*\frac{1}{x}\Big)(x')\cdot\frac{1}{y'}f(x-x',y-y')\,dx'dy'\Big]+$$

$$+\Big[\iint_{|y'|>\varepsilon_2}\Big(\varphi_\varepsilon*\frac{1}{x}\Big)(x')\cdot\frac{1}{y'}\cdot f(x-x',y-y')\,dx'dy'-$$

$$-\iint_{R^2}\Big(\varphi_\varepsilon*\frac{1}{x}\Big)(x')H_k(y')f(x-x',y-y')\,dx'dy'\Big].$$

The first term in brackets is dominated by $M_x(T_y f)(x, y)$ (see section II (c)), which is certainly in weak L^1 if $f \in L\log^+ L$. Similarly, if we take that k for which $2^{-1-k} < \varepsilon_1 \leqslant 2^{-k}$, the second term in brackets is dominated by $M_y(T_x f)(x, y)$. So, in fact, Af and $A_0 f$ are essentially the same.

(b) Let us analyze H_k. Writing $K_k = H_k - H_{k-1}$, we obtain $H_k = \sum_{l=-\infty}^{k} K_l$. Pick functions φ_j on the real line, with the following properties: (i) $\hat{\varphi_j}$ is supported in $2^{j-1} \leqslant |x| \leqslant 2^{j+1}$, (ii) $\sum_{j=-\infty}^{\infty} \hat{\varphi_j} = 1$, and (iii) $\left\| \frac{\partial^a}{\partial \xi^a} \hat{\varphi_j}(\xi) \right\|_\infty = O(2^{-ja})$ uniformly in j, for each a. By property (ii), each K_k is equal to $\sum_{l=-\infty}^{\infty} K_k * \varphi_l = \sum_{j=-\infty}^{\infty} K_{kj}$ where by definition $K_{kj} = K_k * \varphi_{k-j}$. We shall need a little precise information on the size and behavior of the K_{kj}. It is not very hard to convince oneself of the following:

If $j > 0$, then K_{kj} is a C^∞ function with "thickness" 2^{j-k} and L^∞-norm roughly 2^{k-2j}. In fact $\left\| \frac{\partial^a}{\partial x^a} K_{kj} \right\|_\infty = O(2^{k-2j+a(k-j)})$ for each a, uniformly in j and k.

If $j < 0$, then K_{kj} is a C^∞ function with "thickness" 2^{-k} and L^∞-norm smaller than 2^{k+j}. In fact $\left\| \frac{\partial^a}{\partial x^a} K_{kj} \right\|_\infty = O(2^{k+j+ka})$ for each a, uniformly in j and k.

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

LEMMA. *Suppose that $f = \sum_l f_l$, where each f_l lives on an interval J with center x_l, and has average zero. Then outside the union of the doubles of the J_l's, the maximal Hilbert transform of f is dominated by the Marcinkiewicz integral*

$$\sum \frac{\|f_l\|_1 |J_l|}{(|J_l| + |x - x_l|)^2}. \tag{8}$$

LEMMA. *Suppose that $f = \sum_l f_l$, where each f_l lives on an interval J_l and has average zero. Say that J_l has length 2^{-k_l} and center x_l. Then for each integer j*

$$\left| \sum_l \sum_{k=-\infty}^{k_l} K_{kj} * f_l(x) \right| \leqslant C \cdot 2^{-|j|} \cdot \sum_l \frac{\|f_l\|_1 \cdot 2^{-k_l}}{(2^{-k_l} + |x - x_l|)^2} \quad \textit{for all } x \in R^1. \tag{9}$$

VI. Proof of Theorem 1, continued. In this section we show that Af and $Af^{\sim}$ are roughly equal. By the results of V (a) this is equivalent to estimating $\left(\varphi_\varepsilon * \frac{1}{x} \right)^x * H_m^y * (f - f^{\sim})(x, y)$ for all ε and m. We have $H_m^y * (f - f^{\sim}) = \sum_{n=-\infty}^{\infty} K_n^y * (f - f^{\sim})$. For technical reasons, it will be expedient to split the sum into $\sum_{n \text{ even}}$ and $\sum_{n \text{ odd}}$, and henceforth any sum on the index n will be implicitly assumed to run only over n of one fixed parity. At the end of the section, we shall then have proved estimates for $A_0^{\text{even}}(f - f^{\sim}) = \sup_{\substack{\varepsilon > 0 \\ m}} \left| \left(\varphi_\varepsilon * \frac{1}{x} \right)^x * \sum_{\substack{n \leqslant m \\ n \text{ even}}} K_n^y * (f - f^{\sim}) \right|$ and for $A_0^{\text{odd}}(f - f^{\sim})$, defined similarly. Then we can write $A_0(f - f^{\sim}) \leqslant A_0^{\text{even}}(f - f^{\sim}) + A_0^{\text{odd}}(f - f^{\sim})$, to deduce the estimates we really want. With this minor embarrassment out of the way, we can proceed.

We have to estimate $\left(\varphi_\varepsilon * \frac{1}{x} \right)^x * H_m^y * (f - f^{\sim})$. Recall that $f = \sum_{k=-\infty}^{\infty} f_k$ and $f^{\sim} = \sum_{k=-\infty}^{\infty} f_k^{\sim}$ where $f_k, f_k^{\sim} \in \mathscr{I}_k$. We shall break up $H_m^y * (f - f^{\sim})$ into two parts. To do so, say that $f_k - f_k^{\sim} = \sum_l f_k^l$, where $f_k^l = (f_k - f_k^{\sim}) \cdot \chi_{I \times J_l^{x_0 k}}$ with $x_0 \in I$, $|I| = 2^{1-k}$. Thus, for fixed x, $f_k^l(x, \cdot)$ lives on the interval $J_l^{x_0 k}$, has average zero, and satisfies $\|f_k^l(x, \cdot)\|_1 \leqslant \|F_k^{\sim}(x, \cdot)\|_{L^1(J_l^{x_0 k})}$. Say also that J_l^{xk} has length $2^{-r(x,k,l)}$ and center $y(x, k, l)$.

Now

$$\begin{aligned} H_m^y * (f_k - f_k^{\sim})(x, \cdot) &= \sum_l H_m^y * f_k^l(x, \cdot) = \sum_l \sum_{n=-\infty}^{m} K_n^y * f_n^l(x, \cdot) \\ &= \sum_l \sum_{n=-\infty}^{\min(r(x,k,l),m)} K_n^y * f_k^l(x, \cdot) + \\ &\quad + \sum_l \sum_{n=\min(r(x,k,l),m)+1}^{m} K_n^y * f_k^l(x, \cdot) \\ &\equiv H'_m f_k(x, \cdot) + H''_m f_k(x, \cdot). \end{aligned}$$

Outside an exceptional set, we will be able to estimate both $\left(\varphi_\varepsilon * \frac{1}{x}\right)^x * H'_m f_k$ and $\left(\varphi_\varepsilon * \frac{1}{x}\right)^x * H''_m f_k$. Our estimates will be so sharp that we will be able to combine the results for different k to estimate the maximal double Hilbert transform of $(f - f^{\sim})$.

(a) The term $H'_m f_k$. We have

$$H'_m f_k(x, \cdot) = \sum_l \sum_{n=-\infty}^{\min(r(x,k,l),m)} \sum_{j=-\infty}^{\infty} K^y_{nj} * f^l_k(x, \cdot)$$
$$= \sum_{j=-\infty}^{\infty} \sum_l \sum_{n=-\infty}^{\min(r(x,k,l),m)} K^y_{nj} * f^l_k(x, \cdot) \equiv \sum_{j=-\infty}^{\infty} H'_{mj} f(x, \cdot).$$

Set $H'_{\infty j} f_k(x, \cdot) \equiv \sum_l \sum_{n=-\infty}^{r(x,k,l)} K^y_{nj} * f^l_k(x, \cdot)$. By (9),

$$|H'_{\infty j} f_k(x, y)| \leqslant C \cdot 2^{-|j|} \cdot \sum_l \frac{\|f^l_k(x, \cdot)\|_1 \cdot 2^{-r(x,k,l)}}{\left(2^{-r(x,k,l)} + |y - y(x, k, l)|\right)^2}$$
$$\leqslant C 2^{-|j|} \sum_l \frac{\|F^{\sim}_k(x, \cdot)\|_{L^1(J^{xk}_l)} \cdot 2^{-r(x,k,l)}}{\left(2^{-r(x,k,l)} + |y - y(x, k, l)|\right)^2}.$$

By a well-known inequality for Marcinkiewicz integrals (see [2]) $\|H_{\infty j} f_k\|_2 \leqslant C \cdot 2^{-|j|} \|F^{\sim}_k\|_2$. On the other hand, $H'_{\infty j} f_k \in \mathscr{I}_k$, so that for different k, the $H'_{\infty j} f_k$ are orthogonal. Hence

$$\Big\| \sum_{k=-\infty}^{\infty} H'_{\infty j} f_k \Big\|_2 \leqslant C \cdot 2^{-|j|} \cdot \Big(\sum_{k=-\infty}^{\infty} \|F^{\sim}_k\|_2^2 \Big)^{1/2} \leqslant C \cdot 2^{-|j|} C \alpha (\|f \log^+ f\|_1 + C)^{1/2}$$

by inequality (3). Since T_x, the Hilbert transform in the x variable, is bounded on L^2, it follows that

$$\Big\| \sum_{k=-\infty}^{\infty} T_x (H'_{\infty j} f_k) \Big\|_2^2 \leqslant C \cdot 2^{-|j|} \alpha (\|f \log^+ f\|_1 + C).$$

Recalling definitions, we have

$$\sum_{k=-\infty}^{\infty} T_x (H'_{\infty j} f_k) = \sum_{k=-\infty}^{\infty} T_x \Big[\sum_l \sum_{n=-\infty}^{r(x,k,l)} K^y_{nj} * f^l_k \Big]$$
$$= \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} T_x \Big[\sum_{l \text{ for which } r(x,k,l) \geqslant n} K^y_{nj} * f^l_k(x, \cdot) \Big] \equiv \sum_{n=-\infty}^{\infty} R^n_j.$$

So we have proved that $\big\| \sum_{n=-\infty}^{\infty} R^n_j \big\|_2^2 \leqslant C \cdot 2^{-2|j|} \alpha (\|f \log^+ f\|_1 + C)$. On the other hand, the term K^y_{nj} in R^n_j ensures that the Fourier transform, taken in the y variable, of R^n_j lives in the region $2^{j+n-1} \leqslant |y| \leqslant 2^{j+n+1}$. Because of our convention on $\sum_{n \text{ even}}$ and $\sum_{n \text{ odd}}$, these regions are pairwise disjoint for fixed j, and so the lemma of II (b) shows that

$$\Big\| \max_m \Big| \sum_{n=-\infty}^{m} R^n_j(\cdot, \cdot) \Big| \Big\|_2^2 \leqslant C \Big\| \sum_{n=-\infty}^{\infty} R^n_j \Big\|_2^2 \leqslant C 2^{-2|j|} \alpha (\|f \log^+ f\|_1 + C).$$

What does that mean? We know that

$$\sum_{k=-\infty}^{\infty} T_x * H'_{mj} f_k = \sum_{k=-\infty}^{\infty} T_x \Big[\sum_l \sum_{n=-\infty}^{\min(m, r(x,k,l))} K^y_{nj} * f^l_k \Big]$$
$$= \sum_{n=-\infty}^{m} \sum_{k=-\infty}^{\infty} T_x \Big[\sum_{\substack{l \text{ for which} \\ r(x,k,l) \geqslant n}} K^y_{nj} * f^l_k(x, \cdot) \Big] \equiv \sum_{n=-\infty}^{m} R^n_j.$$

So our L^2-inequality means simply that

$$\Big\| \sup_m \Big| T_x \Big[\sum_k H'_{mj} * f_k \Big] (\cdot, \cdot) \Big| \Big\|_2^2 \leqslant C \alpha \cdot 2^{-2|j|} \cdot (\|f \log^+ f\|_1 + C).$$

Now applying the maximal theorem for L^2, we get

$$\Big\| \sup_{\substack{\varepsilon > 0 \\ m}} \Big| \Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \Big(\sum_{k=-\infty}^{\infty} H'_{mj} f_k \Big) (\cdot, \cdot) \Big| \Big\|_2^2 \leqslant C \cdot 2^{-2|j|} \alpha (\|f \log^+ f\|_1 + C).$$

Summing over j yields

$$\Big\| \sup_{\substack{\varepsilon > 0 \\ m}} \Big| \Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \Big(\sum_{k=-\infty}^{\infty} H'_m f_k \Big) (\cdot, \cdot) \Big| \Big\|_2^2 \leqslant C \alpha (\|f \log^+ f\|_1 + C),$$

which implies that

$$(10) \qquad \Big| \Big\{ (x, y) \in R^2 \,\Big|\, \sup_{\substack{\varepsilon > 0 \\ m}} \Big| \Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \Big(\sum_{k=-\infty}^{\infty} H'_m f_k \Big) (x, y) \Big| > \alpha \Big\} \Big| \leqslant \frac{C}{\alpha} (\|f \log^+ f\|_1 + C).$$

This is exactly the estimate we need for the contribution of the $H'_m f_k$ to $A_0(f - f^{\sim})$.

By the techniques we just used, we could also have proved that

$$\Big\|\sup_m\Big|\sum_{k=-\infty}^{\infty} H'_m f_k(\cdot,\cdot)\Big|\Big\|_2^2 \leqslant C\alpha(\|f\log^+ f\|_1 + C).$$

(b) The term $H''_m f_k$. A look at the definition of $H''_m f_k$ will convince the reader that $H''_m f_k$ lives only on rectangles $I\times J^*$, where I is dyadic of length 2^{1-k}, J^* is an interval concentric with J and five times as large, and $J = J_l^{xk}$ for $x\in I$. By definition of J_l^{xk}, such rectangles have at least one tenth of their area inside Ω_0. Consequently, $H''_m f_k$ lives entirely on the set Ω_1 of all points of R^2 at which the strong maximal function of χ_{Ω_0} exceeds $\frac{1}{10}$. Furthermore, $H''_m f_k$ always has average zero over the component intervals of $\Omega_1\cap(R^1\times\{y_0\})$, for all $y_0\in R^1$. So by inequality (8) above, $\sup_{\varepsilon>0}\Big|\Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \sum_{k=-\infty}^{\infty} H''_m f_k\Big|$ is dominated outside Ω by a Marcinkiewicz integral (in the x variable) of the function $\Big|\sum_{k=-\infty}^{\infty} H''_m f_k\Big|$. Here, Ω denotes the union of all the intervals $I^*\times\{y_0\}$, where I is a component interval of $\{x\in R^1 | (x, y_0)\in\Omega_1\}$. Therefore, $\sup_{\substack{\varepsilon>0\\ m}}\Big|\Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \sum_{k=-\infty}^{\infty} H''_m f_k\Big|$ is dominated outside Ω by a Marcinkiewicz integral (again in the x variable) of $\sup_m\Big|\sum_{k=-\infty}^{\infty} H''_m f_k\Big|$. It follows then from the L^1 boundedness of Marcinkiewicz integrals (see [4]) that

$$\Big|\Big\{(x,y)\in R^2-\Omega \,\Big|\, \sup_{\substack{\varepsilon>0\\ m}}\Big|\Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \sum_{k=-\infty}^{\infty} H''_m f_k(x,y)\Big| > \alpha\Big\}\Big|$$

$$\leqslant \frac{C}{\alpha}\Big\|\sup_m\Big|\sum_{k=-\infty}^{\infty} H''_m f_k\Big|\Big\|_1.$$

So to finish off our weak-type estimate for $\sup_{\substack{\varepsilon>0\\ m}}\Big|\Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x \sum_{k=-\infty}^{\infty} H''_m f_k\Big|$, we need only show that

(11) $$\Big\|\sup_m\Big|\sum_{k=-\infty}^{\infty} H''_m f_k\Big|\Big\|_1 \leqslant C(\|f\log^+ f\|_1 + C).$$

To prove this inequality, write $H''_m f_k = H_m^y * (f_k - f_k^\sim) - H'_m f_k$, so that

(12) $$\sum_{k=-\infty}^{\infty} H''_m f_k = H_m^y(f - f^\sim) - \sum_{k=-\infty}^{\infty} H'_m f_k.$$

Now by standard results on the maximal Hilbert transform,

(13) $$\|\sup_m |H_m^y * f|\|_1 \leqslant C(\|f\log^+ f\|_1 + C) \quad\text{and}$$
$$\|\sup_m |H_m^y * f^\sim|\|_2^2 \leqslant C\alpha(\|f\log^+ f\|_1 + C).$$

by virtue of $f\in L\log^+ L$ and $f^\sim\in L^2$.

We saw in VI (a) above that

(14) $$\Big\|\sup_m\Big|\sum_{k=-\infty}^{\infty} H'_m f_k\Big|\Big\|_2^2 \leqslant C\alpha(\|f\log^+ f\|_1 + C).$$

Since $|\Omega| \leqslant 5|\Omega_1| \leqslant C|\Omega_0|$ (by the strong maximal theorem) $\leqslant \frac{C}{\alpha}(\|f\log^+ f\|_1 + C)$, Hölder's inequality and (13), (14) show that

$$\Big\|\sup_m\Big|\sum_{k=-\infty}^{\infty} H'_m f_k\Big|\Big\|_1 \leqslant C(\|f\log^+ f\|_1 + C),$$

$$\|\sup_m |H_m^y * f|\|_{L^1(\Omega)} \leqslant C(\|f\log^+ f\|_1 + C),$$

and

$$\|\sup_m |H_m^y * f^\sim|\|_{L^1(\Omega)} \leqslant C(\|f\log^+ f\|_1 + C).$$

Putting these estimates into (12) shows that

$$\Big\|\sup_m\Big|\sum_{k=-\infty}^{\infty} H''_m f_k\Big|\Big\|_{L^1(\Omega)} \leqslant C(\|f\log^+ f\|_1 + C),$$

which is exactly (11) since $H''_m f_k$ lives on Ω. This completes the proof of our main weak-type inequality for $H''_m f_k$:

(15) $$\Big|\Big\{(x,y)\in R^2-\Omega \,\Big|\, \sup_{\substack{\varepsilon>0\\ m}}\Big|\Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * \sum_{k=-\infty}^{\infty} H''_m f_k(x,y)\Big| > \alpha\Big\}\Big|$$
$$\leqslant \frac{C}{\alpha}(\|f\log^+ f\| + C)$$

(c) Conclusion. From (15) and (10), and our basic decomposition $H_m^y * (f - f^\sim) = \sum_{k=-\infty}^{\infty} H'_m f_k + \sum_{k=-\infty}^{\infty} H''_m f_k$, we obtain the weak-type inequality

(16) $$\Big|\Big\{(x,y)\in R^2-\Omega \,\Big|\, \sup_{\substack{\varepsilon>0\\ m}}\Big|\Big(\varphi_\varepsilon * \frac{1}{x}\Big)^x * H_m^y * (f - f^\sim)(x,y)\Big| > \alpha\Big\}\Big|$$
$$\leqslant \frac{C}{\alpha}(\|f\log^+ f\|_1 + C).$$

Since also $|\Omega| \leqslant \frac{C}{\alpha}(\|f\log^+ f\|_1 + C)$, we have

$$(17) \qquad \left|\{(x,y)\in R^2 \mid A_0(f-f^{\sim})(x,y) > \alpha\}\right| \leqslant \frac{C}{\alpha}(\|f\log^+ f\|_1 + C).$$

Thus $A_0(f)$ and $A_0(f^{\sim})$ are approximately equal.

VII. Proof of Theorem 1: Mop-Up. We have completed our program of constructing an $f^{\sim}$, showing that $f^{\sim}\in L^2$, and proving that $A(f) \approx A(f^{\sim})$. The proof of Theorem 1 is now a triviality. Clearly

$$\left|\{(x,y)\in R^2 \mid A(f)(x,y) > 2\alpha\}\right| \leqslant \left|\{(x,y)\in R^2 \mid A(f^{\sim})(x,y) > \alpha\}\right| + \left|\{(x,y)\in R^2 \mid A(f)(x,y) - A(f^{\sim})(x,y) > \alpha\}\right|.$$

The first term is at most $\frac{\|A(f^{\sim})\|_2^2}{\alpha^2} \leqslant \frac{C}{\alpha^2}\|f^{\sim}\|_2^2 \leqslant \frac{C}{\alpha}(\|f\log^+ f\|_1 + C)$, by the Chebyshev inequality, the L^2-boundedness of A, and our estimate for $\|f^{\sim}\|_2^2$. The second term is at most $\frac{C}{\alpha}(\|f\log^+ f\|_1 + C)$ by inequality (17). Thus, $|\{(x,y)\in R^2 | Af(x,y) > 2\alpha\}| \leqslant \frac{C}{\alpha}(\|f\log^+ f\|_1 + C)$. The proof of Theorem 1 is complete.

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

(b) Note that in order to find a non-trivial problem, we have to ask for Theorem 1 in its full strength. For instance, suppose we merely want to know that $A_1(f)$ is in weak L^1 for $f\in L\log^+ L$, where

$$A_1 f(x,y) = \sup_{\delta>0}\left|\iint_{\substack{|x'|>\delta\\|y'|>\delta}} \frac{f(x-x', y-y')}{x'y'}\,dx'\,dy'\right|.$$

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

1. θ is homogeneous of degree zero on R^2 and C^∞ on $R^2-\{0\}$.
2. $\theta(x,y) = 1$ for $|x| \leqslant \frac{1}{2}|y|$, except at the origin.
3. $\theta(x,y) = 0$ for $|y| \leqslant \frac{1}{2}|x|$, except at the origin.

Then if T is the double Hilbert transform, we have

$$\begin{aligned}[Tf(x,y)]^{\wedge} &= \operatorname{sgn}(x)\operatorname{sgn}(y) f^{\wedge}(x,y)\\ &= \operatorname{sgn}(x)\cdot[\theta(x,y)\operatorname{sgn}(y)]\cdot f^{\wedge}(x,y) + \operatorname{sgn}(y)\cdot\left[\left(1-\theta(x,y)\right)\sin(x)\right]\cdot f(x,y)\\ &\equiv \operatorname{sgn}(x)\cdot K_1^{\wedge}(x,y)\cdot f^{\wedge}(x,y) + \operatorname{sgn}(y)\cdot K_2^{\wedge}(x,y)\cdot f^{\wedge}(x,y),\end{aligned}$$

where K_1 and K_2 are Calderón–Zygmund kernels on R^2. In other words, $Tf = T_x(K_1 * f) + T_y(K_2 * f)$, where T_x and T_y are as in II (c).

Had we taken δ into account in the above calculation, we would have found $A_1 f \leqslant A_x(K_1 * f) + A_y(K_2 * f)$ + trivial error terms, where A_x and A_y are the maximal Hilbert transforms in the x and y directions, respectively. Therefore

$$\|A_1 f\|_{wkL^1} \leqslant C\|K_1 * f\|_1 + C\|K_2 * f\|_1 \leqslant C\|f\|_{L\log^+ L}. \quad \blacksquare$$

References

[1] A. P. Calderón and A. Zygmund, *On the existence of certain singular integrals*, Acta Math. 88 (1952), pp. 85-139.

[2] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. 116 (1966), pp. 135-157.

[3] A. Zygmund, *Trigonometric Series*, Cambridge 1959.

[4] — *On the boundary values of functions of several complex variables*, Fund. Math. 36 (1949), pp. 207-235.

Received December 28, 1970 (282)