INTERPOLATION, MAXIMAL OPERATORS, AND THE HILBERT TRANSFORM

MICHAEL WONG

Abstract. Real-variable methods are used to prove the Marcinkiewicz Interpolation Theorem, boundedness of the dyadic and Hardy-Littlewood maximal operators, and the Calderón-Zygmund Covering Lemma. The Hilbert transform is defined, and its boundedness is investigated. All results lead to a final theorem on the pointwise convergence of the truncated Hilbert transform

Contents

1. Introduction

The Hilbert transform H on $\mathbb{R}$ is formally defined by

$$Hf(x) = \lim_{\epsilon \to 0^+} H_\epsilon(x) \tag{1.1}$$

where H_ϵ is the truncated Hilbert transform at $\epsilon > 0$,

$$H_\epsilon f(x) = \int_{|y|>\epsilon} \frac{f(x-y)}{y}\, dy$$

and dx is the Lebesgue measure. While H_ϵ is well-defined for a large class of functions, the principal value integral implicit in Equation (1.1) is finite for only well-behaving functions. However, H has boundedness properties by which it can be extended to larger function spaces.

In the following exposition, we build the real-variable tools needed to extend the Hilbert transform to $L^p(\mathbb{R}, \mu)$ for all $p \in [1, \infty)$, where μ is the Lebesgue measure.

Date: August 31, 2010.

These instruments include the Marcinkiewicz Interpolation Theorem, Calderón-Zygmund decomposition, Schwartz functions, and tempered distributions. In the end, we prove that Equation (1.1) is valid for $f \in L^p$, $p \in [1, \infty)$, up to a μ-null set.

2. Preliminaries

We begin by defining different types of boundedness. The basic domain considered in this paper is $L^p(X, \mu)$, where $p \in (0, \infty]$, X is an arbitrary set, and μ is a nonnegative, extended real-valued measure. Intuitively, the image of L^p under a given operator T determines the strength of its boundedness: does T map L^p to q-integrable functions, $0 < q \leq \infty$, or to functions satisfying only a weaker condition? One such weaker condition is that which defines the *weak-L^p space*.

Definition 2.1. Let (X, μ) be a measure space, and let $f : X \to \mathbb{C}$ be a measurable function. For measurable $A \subseteq X$, we denote $\mu(A)$ by $|A|$. The *distributional function* of f is the function $d_f : \mathbb{R}_+ \to \overline{\mathbb{R}}$ given by

$$d_f(\lambda) = |\{x \in X : |f(x)| > \lambda\}|$$

If $X = \mathbb{R}^n$, we set μ to be the Lebesgue measure.

Note that d_f is a measurable function.

Definition 2.2. For $p \in (0, \infty)$, *weak-$L^p(X, \mu)$*, denoted by $L^{p,\infty}(X, \mu)$, is the space of all measurable functions $f : X \to \mathbb{C}$ such that

$$||f||_{p,\infty} := \inf_{\lambda > 0}\{C > 0 : d_f(\lambda) \leq \frac{C^p}{\lambda^p}\} = \sup_{\lambda > 0}\{\lambda d_f(\lambda)^{\frac{1}{p}}\} < \infty$$

We set *weak-$L^\infty(X, \mu)$* to be $L^\infty(X, \mu)$ and $|| \cdot ||_{\infty,\infty}$ to be $|| \cdot ||_\infty$.

The map $|| \cdot ||_{p,\infty}$ is a quasinorm on the linear space $L^{p,\infty}(X, \mu)/\mathbb{C}$. Chebyshev's inequality shows that $L^p \subseteq L^{p,\infty}$. If $X = \mathbb{R}^n$, a counterexample demonstrating that this containment is strict is $f(x) = |x|^{-n/p}$. So L^p and $L^{p,\infty}$ determine two types of boundedness.

Definition 2.3. An operator T from $L^p(X, \mu)$ to the space of complex-valued, measurable functions on a measure space (Y, ν) is *sublinear* if

(1) $\forall f, g \in L^p(X, \mu), \quad |T(f + g)(y)| \leq |Tf(y)| + |Tg(y)|$
(2) $\forall \alpha \in \mathbb{C}, \quad |T(\alpha f)(y)| = |\alpha||Tf(y)|$

Definition 2.4. A sublinear operator T is *weakly bounded from p to q*, $0 < p, q \leq \infty$, if there exists $C > 0$ such that

$$||Tf||_{q,\infty} \leq C||f||_p \quad \forall f \in L^p(X, \mu)$$

We say that such an operator T is weak (p, q) for short. T is *strongly bounded from p to q* if there exists a $C > 0$ such that

$$||Tf||_q \leq C||f||_p \quad \forall f \in L^p(X, \mu)$$

We say that T is strong (p, q).

Remark 2.5. Note that by our definitions, weak (p, ∞) boundedness is identical to strong (p, ∞) boundedness. We will say that an operator satisfying Definition 2.4 for $q = \infty$ is *bounded* (p, ∞).

In other words, T is strong (p, q) if T maps L^p into L^q; T is weak (p, q) if T maps L^p only into $L^{q,\infty}$. By the fact that $L^q \subseteq L^{q,\infty}$, T is strong (p, q) implies T is weak (p, q). Note that T is weak (p, q) if and only if for all $f \in L^p$ and $\lambda > 0$,

$$d_{Tf}(\lambda) \leq (\frac{C}{\lambda}||f||_p)^q$$

The following theorem about the pointwise convergence of linear operators, based only on the above definitions, will be of use later.

Theorem 2.6. *Let $\{T_t\}$ be a family of linear operators mapping $L^p(X, \mu)$ into the space of complex-valued, measurable functions over (X, μ). Define the maximal operator T^* by*

$$T^* f(x) = \sup_t \{|T_t f(x)|\}$$

If T^* is weak (p, q), then the set

$$A = \{f \in L^p : \lim_{t \to t_0} T_t f(x) \text{ exists a.e.}\}$$

is closed in L^p.

Proof. Assume that $T_t f$ is real-valued. If $T_t f$ is complex-valued, apply the following argument to the real and imaginary parts of $T_t f$ separately. First, observe that for all $f \in L^p$,

$$\limsup_{t \to t_0} T_t f(x) - \liminf_{t \to t_0} T_t f(x) \leq 2T^* f(x) \tag{2.7}$$

Now suppose $\{f_n\} \subset A$ converges in L^p norm to f. Each f_n satisfies

$$|\{x \in X : \limsup_{t \to t_0} T_t f_n(x) - \liminf_{t \to t_0} T_t f_n(x) > 0\}| = 0$$

and it suffices to show f satisfies the same equation. For all $\lambda > 0$,

$$\begin{aligned} &|\{x \in X : \limsup_{t \to t_0} T_t f(x) - \liminf_{t \to t_0} T_t f(x) > \lambda\}| \\ &\quad \leq |\{x \in X : \limsup_{t \to t_0} T_t(f - f_n)(x) - \liminf_{t \to t_0} T_t(f - f_n)(x) > \lambda\}| \\ &\quad \leq |\{x \in X : 2T^*(f - f_n)(x) > \lambda\}| \qquad \text{(Equation (2.7))} \\ &\quad = d_{T^*(f-f_n)}(\frac{\lambda}{2}) \leq (\frac{2C}{\lambda}||f - f_n||_p)^q \qquad (T \text{ is weak } (p, q)) \end{aligned}$$

The limit of the last term as $n \to \infty$ is 0. Hence,

$$\begin{aligned} &|\{x \in X : \limsup_{t \to t_0} T_t f(x) - \liminf_{t \to t_0} T_t f(x) > 0\}| \\ &\quad \leq \sum_{k=1}^{\infty} |\{x \in X : \limsup_{t \to t_0} T_t f(x) - \liminf_{t \to t_0} T_t f(x) > \frac{1}{k}\}| = 0 \end{aligned}$$

□

Remark 2.8. By a similar proof, one could show that the set

$$A' = \{f \in L^p : \lim_{t \to t_0} T_t f(x) = f(x) \text{ a.e.}\}$$

is closed in L^p. One would disregard Equation (2.7) and thereafter replace $\liminf T_t f_n(x)$ with $f_n(x)$ and $\liminf T_t f(x)$ with $f(x)$.

3. Marcinkiewicz Interpolation Theorem

The next step is to determine the $p \in (0, \infty]$ for which a given sublinear operator is bounded (p,p). The Marcinkiewicz interpolation theorem asserts strong boundedness for all values of p between two values for which weak boundedness is established. In our proof of the theorem, the following two lemmas will be used. We set $d\mu(x) = dx$.

Lemma 3.1. *Let $L^{p_0}(X,\mu) + L^{p_1}(X,\mu)$, $1 \le p_0 < p_1 \le \infty$, be the direct sum of $L^{p_0}(X,\mu)$ and $L^{p_1}(X,\mu)$. If $p_0 < p < p_1$ and $f \in L^p(X,\mu)$, then $f = f_0 + f_1$ for some $f_0 \in L^{p_0}$ and $f_1 \in L^{p_1}$.*

Proof. Fix $\lambda > 0$. Given $f \in L^p$, define f_0 and f_1 by

$$f_0 = f\chi_{\{x:|f(x)|\ge c\lambda\}} \qquad f_1 = f\chi_{\{x:|f(x)|<c\lambda\}}$$

where the value of the constant c will be chosen in Theorem 3.3. To see that $f_0 \in L^{p_o}$, observe that

$$\begin{aligned}\int_X |f_0(x)|^{p_0}\,dx &= \int_{\{|f(x)|\ge c\lambda\}} |f(x)|^{p_0}(\frac{c\lambda}{c\lambda})^{p-p_0}\,dx \\ &\le \int_{\{|f(x)|\ge c\lambda\}} |f(x)|^{p_0}(\frac{|f(x)|}{c\lambda})^{p-p_0}\,dx \\ &\le (\frac{1}{c\lambda})^{p-p_0}||f||_p^p\end{aligned}$$

The fact that $f_1 \in L^{p_1}$ is shown similarly. □

Lemma 3.2. *If $f \in L^p(X,\mu)$, $1 \le p < \infty$, then $||f||_p^p = \int_0^\infty p\lambda^{p-1}d_f(\lambda)\,d\lambda$.*

Proof.

$$\begin{aligned}||f||_p^p &= \int_X |f(x)|^p\,dx \\ &= \int_X \int_0^{|f(x)|} p\lambda^{p-1}\,d\lambda dx \\ &= \int_0^\infty \int_{\{x:|f(x)|>\lambda\}} p\lambda^{p-1}\,dx d\lambda \qquad \text{(Fubini's theorem)} \\ &= \int_0^\infty p\lambda^{p-1}d_f(\lambda)\,d\lambda\end{aligned}$$

□

Theorem 3.3 (Marcinkiewicz Interpolation Theorem)**.** *Suppose T is a sublinear operator from $L^{p_0}(X,\mu) + L^{p_1}(X,\mu)$, $1 \le p_0 < p_1 \le \infty$, to the space of complex-valued, measurable functions on (Y,ν). If T is weak (p_0,p_0) and weak (p_1,p_1), then T is strong (p,p) for all $p \in (p_0,p_1)$.*

Proof. Fix $\lambda > 0$. Given $f \in L^p$, define f_0 and f_1 as in Lemma 3.1. Because T is sublinear,

$$d_{Tf}(\lambda) \le d_{Tf_0}(\frac{\lambda}{2}) + d_{Tf_1}(\frac{\lambda}{2}) \tag{3.4}$$

We choose c by cases for the value of p_1.

First, suppose $p_1 < \infty$, and let $c = 1$. T is weak (p_0, p_0) and weak (p_1, p_1), so Inequality (3.4) becomes

$$d_{Tf}(\lambda) \leq (\frac{2C_0}{\lambda}||f_0||_{p_0})^{p_0} + (\frac{2C_1}{\lambda}||f_1||_{p_1})^{p_1}$$

from which we derive the following estimate:

$$\begin{aligned}
||Tf||_p^p &= \int_0^\infty p\lambda^{p-1} d_{Tf}(\lambda)\, d\lambda && \text{(Lemma 3.2)} \\
&\leq \int_0^\infty (2C_0)^{p_0} p\lambda^{p-p_0-1} ||f_0||_{p_0}^{p_0}\, d\lambda + \int_0^\infty (2C_1)^{p_1} p\lambda^{p-p_1-1} ||f_1||_{p_1}^{p_1}\, d\lambda \\
&= \int_0^\infty (2C_0)^{p_0} p\lambda^{p-p_0-1} \int_{\{y:|f(y)|\geq\lambda\}} |f(y)|^{p_0}\, dy\, d\lambda \\
&\quad + \int_0^\infty (2C_1)^{p_1} p\lambda^{p-p_1-1} \int_{\{y:|f(y)|<\lambda\}} |f(y)|^{p_1}\, dy\, d\lambda \\
&= (2C_0)^{p_0} p \int_X |f(y)|^{p_0} \int_0^{|f(y)|} \lambda^{p-p_0-1}\, d\lambda\, dy \\
&\quad + (2C_1)^{p_1} p \int_X |f(y)|^{p_1} \int_{|f(y)|}^\infty \lambda^{p-p_1-1}\, d\lambda\, dy && \text{(Fubini's theorem)} \\
&= (\frac{(2C_0)^{p_0} p}{p-p_0} + \frac{(2C_1)^{p_1} p}{p_1 - p})||f||_p^p
\end{aligned}$$

Now, suppose $p_1 = \infty$. Then $||Tf||_\infty \leq C_1||f||_\infty$ for all $f \in L^\infty$. If $c = 1/(2C_1)$, then $f_1 = f\chi_{\{x:|f(x)|<\lambda/(2C_1)\}}$, implying $||Tf_1||_\infty \leq \frac{\lambda}{2}$. So $d_{Tf_1}(\frac{\lambda}{2}) = 0$ in Inequality (3.4), and with just one term, the result follows as above. □

4. The Dyadic Maximal Operator and Calderón-Zygmund Decomposition

The Calderón-Zygmund decomposition of a function $f \in L^1(\mathbb{R}^n)$ will be essential to our study of the Hilbert transform. This decomposition is a corollary to the Calderón-Zygmund Covering Lemma, which we prove using the *dyadic maximal operator*. The construction of this operator involves the partition of $\mathbb{R}^n$ into dyadic cubes.

Consider the cube $[0,1)^n$ generated by half-open intervals. The collection $\mathscr{D}_0 = [0,1)^n + \mathbb{Z}^n$ of translates of the cube is a pairwise disjoint cover of $\mathbb{R}^n$, and every point in $\mathbb{Z}^n$ is a vertex of 2^n cubes in $\mathscr{D}_0$. For $k \in \mathbb{Z}$, let $\mathscr{D}_k = [0, 2^{-k}) + 2^{-k}\mathbb{Z}^n$. In words, $\mathscr{D}_{k-1}$ is formed from $\mathscr{D}_k$ by partitioning each cube in $\mathscr{D}_k$ into 2^n disjoint, equal cubes. The following observations are clear:

1. for all k, $\mathscr{D}_k$ is pairwise disjoint.
2. each cube in $\mathscr{D}_k$ contains exactly 2^n cubes in $\mathscr{D}_{k-1}$
3. for every $x \in \mathbb{R}^n$, there exists a unique sequence of dyadic cubes $\{Q_k\}_{k\in\mathbb{Z}}$ such that $x \in Q_k$ and $Q_k \in \mathscr{D}_k$ for all k.

Given $f \in L^1_{loc}(\mathbb{R}^n)$, we define an operator E_k whose value $E_kf(x)$ is the average of f over $Q \in \mathscr{D}_k$ where $x \in Q$. Then the dyadic maximal operator is given by the supremum over k of these averages:

Definition 4.1. The operator E_k on $L^1_{loc}(\mathbb{R}^n)$ is defined by

$$E_k f(x) = \sum_{Q \in \mathscr{D}_k} \Big(\frac{1}{|Q|} \int_Q f \Big) \chi_Q(x)$$

The *dyadic maximal operator* E is then defined by

$$Ef(x) = \sup_{k \in \mathbb{Z}} \{ |E_k f(x)| \}$$

Clearly, E is sublinear. The Calderón-Zygmund Covering Lemma will be a small step from the following theorem.

Theorem 4.2. .

(1) E is weak $(1,1)$
(2) If $f \in L^1_{loc}(\mathbb{R}^n)$, then $\lim_{k\to\infty} E_k f(x) = f(x)$ a.e.

Proof. .

(1) Suppose $f \in L^1$, and fix $\lambda > 0$. Let

$$\Omega_k = \{ x \in \mathbb{R}^n : |E_k f(x)| > \lambda \text{ and } |E_j f(x)| \leq \lambda \; \forall j < k \} \tag{4.3}$$

To see that this definition makes sense, observe that for all x, $\lim_{k\to-\infty} E_k f(x) = 0$. So if there exists k' such that $|E_{k'} f(x)| > \lambda$, then there exists a k as in the definition of Ω_k: namely, the smallest k'. It follows from the definition of E_k that there exists a subcollection $S_k \subseteq \mathscr{D}_k$ such that

$$\Omega_k = \bigcup_{Q \in S_k} Q \tag{4.4}$$

Clearly, $\{\Omega_k\}_{k\in\mathbb{Z}}$ is pairwise disjoint. Then writing

$$\{x : Ef(x) > \lambda\} = \bigcup_k \Omega_k \tag{4.5}$$

we see that

$$\begin{aligned} d_{Ef}(\lambda) = \sum_k |\Omega_k| &\leq \frac{1}{\lambda} \sum_k \int_{\Omega_k} |E_k f| && \text{(Chebyshev's inequality)} \\ &= \frac{1}{\lambda} \sum_k \sum_{Q \in S_k} \Big| \frac{1}{|Q|} \int_Q f \Big| && \text{(Equation (4.4))} \\ &\leq \frac{1}{\lambda} ||f||_1 && (4.6) \end{aligned}$$

(2) Assume $f \in L^1$ is continuous. For a fixed $x \in \mathbb{R}^n$,

$$E_k f(x) = \frac{1}{|Q_k|} \int_{Q_k} f$$

where $\{Q_k\}$ is the unique sequence in observation (3) above. By the continuity of f, for all $\epsilon > 0$, there exists $K \in \mathbb{Z}$ such that for all $k \geq K$,

$$|E_k f(x) - f(x)| = \Big| \frac{1}{|Q_k|} \int_{Q_k} f(y) - f(x)\, dx \Big| \leq \frac{1}{|Q_k|} \int_{Q_k} |f(y) - f(x)|\, dx < \epsilon$$

Therefore, $\lim_{k\to\infty} E_k f(x) = f(x)$ everywhere. The set of continuous, integrable functions is dense in L^1. Hence, by Theorem 2.6, the equation

$$\lim_{k\to\infty} E_k f(x) = f(x) \text{ a.e.} \tag{4.7}$$

holds for all $f \in L^1$.

Now, if $f \in L^1_{loc}$, then for all compact cubes $R_m = [-m, m]^n$, $m \in \mathbb{N}$, $f\chi_{R_m}$ is integrable. Thus, Equation (4.7) is satisfied for a.e. $x \in R_m$. So $E_k f(x)$ converges to $f(x)$ for a.e. $x \in \bigcup_m R_m = \mathbb{R}^n$.

□

Lemma 4.8 (Calderón-Zygmund Covering Lemma)**.** *Suppose $f \in L^1(\mathbb{R}^n)$ is real-valued and nonnegative, and fix $\lambda > 0$. There exists a pairwise disjoint sequence of dyadic cubes $\{Q_j\}$ such that*

(1) $f(x) \leq \lambda$ *a.e.* $x \notin \bigcup_j Q_j$
(2) $|\bigcup_j Q_j| \leq \frac{1}{\lambda}||f||_1$
(3) $\lambda < \frac{1}{|Q_j|}\int_{Q_j} f \leq 2^n \lambda$

Proof. Define Ω_k and S_k as in Theorem 4.2, and let $\bigcup_k S_k = \{Q_j\}$. Then

(1) by Equation (4.3), $E_k f(x) \leq \lambda$ for all $x \notin \bigcup_j Q_j$. Hence, by Theorem 4.2, Part 2, the result follows.
(2) Observe that $|\bigcup_j Q_j| = \sum_k |\Omega_k|$. Therefore, by Inequality (4.6), the result follows.
(3) The first inequality is a consequence of the definition of Ω_k. Let $\tilde{Q}_j$ be the unique cube in $\mathscr{D}_{j-1}$ containing Q_j. Note that the average of f over $\tilde{Q}_j$ is at most λ. Then

$$\frac{1}{|Q_j|}\int_{Q_j} f \leq \frac{|\tilde{Q}_j|}{|Q_j|}\frac{1}{|\tilde{Q}_j|}\int_{\tilde{Q}_j} f \leq 2^n\lambda$$

□

Given $\lambda > 0$ and any function f satisfying the hypotheses of Lemma 4.8, we can write $\mathbb{R}^n$ as the union of two disjoint sets, $\Omega := \bigcup_j Q_j$ and $\mathbb{R}^n \setminus \Omega$. In turn, this cover gives a way to decompose f into a sum of two functions g and b, defined by

$$g(x) = \begin{cases} f(x) & \text{if } x \notin \Omega, \\ \frac{1}{|Q_j|}\int_{Q_j} f & \text{if } x \in Q_j \end{cases}, \quad b(x) = \sum_j b_j(x) \tag{4.9}$$

where

$$b_j(x) = \Big(f(x) - \frac{1}{|Q_j|}\int_{Q_j} f\Big)\chi_{Q_j}(x)$$

The function g is called the bounded or "good" part of f, for

$$g(x) \leq 2^n\lambda \quad \forall x \in \mathbb{R}^n \tag{4.10}$$

whereas b is the oscillatory or "bad" part of f, for each b_j has zero average. Note also that b_j vanishes outside Q_j. This way of writing f is called the *Calderón-Zygmund (C-Z) decomposition of f at height λ.*

5. The Hardy-Littlewood Maximal Operator

Later on, we will need to bound convolution-type operators of functions with radial symmetry. It will be convenient to use the *Hardy-Littlewood maximal operator* instead of the dyadic maximal operator for this task. Like the dyadic operator E, the Hardy-Littlewood operator M is defined by the supremum over averages of f. However, the averages in the definition of M are over balls of arbitrary radius, not dyadic cubes.

Definition 5.1. The *Hardy-Littlewood maximal operator* M is defined by

$$Mf(x) = \sup_{r>0} \Big\{ \frac{1}{|B_r|} \int_{B_r} |f(x-y)|\, dy \Big\}$$

where B_r is the ball of radius r centered at 0 and f is any function on $\mathbb{R}^n$ for which this quantity is finite.

Evidently, M is sublinear. Replacing balls with cubes of side length r centered at 0 defines a new operator M' given by

$$M'f(x) = \sup_{r>0} \Big\{ \frac{1}{|Q_r|} \int_{Q_r} |f(x-y)|\, dy \Big\}$$

By considering balls nested in cubes and vice versa, one may verify that there exist constants c_n and C_n, depending only on the dimension n, such that

$$c_n M'f(x) \leq Mf(x) \leq C_n M'f(x) \tag{5.2}$$

Concerning the boundedness of M, it follows immediately from the definition that $||Mf||_\infty \leq ||f||_\infty$, so M is bounded (∞, ∞). By the Marcinkiewicz Interpolation Theorem, if M is also weak $(1,1)$, then M is strong (p,p) for all $p \in (1, \infty)$.

Theorem 4.2 implies that $d_{Ef}(\lambda) \leq 1/\lambda ||f||_1$ where $f \in L^1(\mathbb{R}^n)$. So bounding d_{Mf} by $d_{E|f|}$ pointwise suffices to show M is weak $(1,1)$. (The absolute value is needed because we will use C-Z decomposition, which demands that f is real-valued and nonnegative). To avoid the issue of comparing arbitrary balls to dyadic cubes, however, we prove that $d_{M'f}$ is dominated by $d_{E|f|}$ and apply Equation (5.2).

Proposition 5.3. *If $f \in L^1(\mathbb{R}^n)$, then $d_{M'f}(4^n\lambda) \leq 2^n d_{E|f|}(\lambda)$*

Proof. Note that $M'f = M'|f|$, so without loss of generality, assume f is real-valued and nonnegative. Let $\{Q_j\}$ be the sequence of dyadic cubes in the C-Z decomposition of f at height λ, and let $\tilde{Q}_j$ be the cube with the same center as Q_j but twice the side length. By Equation (4.5),

$$|\bigcup_j \tilde{Q}_j| \leq 2^n |\bigcup_j Q_j| = 2^n d_{Ef}(\lambda)$$

So it suffices to show $\{x \in \mathbb{R}^n : M'f(x) > 4^n\lambda\} \subset \bigcup_j \tilde{Q}_j$.

Suppose $x \notin \bigcup_j \tilde{Q}_j$, and let Q be any cube centered at x. Choose $k \in \mathbb{Z}$ such that $2^{k-1} \leq l(Q) < 2^k$, where $l(Q)$ is the side length of Q. A simple geometric argument shows that Q is covered by m cubes in $\mathscr{D}_{-k}$, where $m \leq 2^n$. Let $\{R_i\}_{i=1}^m$ be this cover. Note that no R_i is contained in any Q_j, for if there were such an R_i, then we would have $x \in \bigcup_j \tilde{Q}_j$. Hence, by Equation (4.3), the average of f over R_i is at most λ. So

$$\begin{aligned}
\frac{1}{|Q|}\int_Q f &= \sum_{i=1}^m \frac{1}{|Q|}\int_{Q\cap R_i} f \\
&\leq \sum_{i=1}^m \frac{2^{kn}}{|Q|}\frac{1}{|R_i|}\int_{R_i} f \qquad (R_i \in \mathscr{D}_{-k}) \\
&\leq 2^n m\lambda \qquad (2^{k-1} < l(Q)) \\
&\leq 4^n \lambda
\end{aligned}$$

This inequality holds for all Q, implying $x \notin \{x \in \mathbb{R}^n : M'f(x) > 4^n\lambda\}$. □

The following proposition shows how M is a pointwise bound to convolution-type operators on $L^1_{loc}(\mathbb{R}^n)$, defined by a particular class of functions. For a function ϕ, we define ϕ_ϵ by $\phi_\epsilon(x) = \epsilon^{-n}\phi(\epsilon^{-1}x)$.

Proposition 5.4. *Suppose $f \in L^1_{loc}(\mathbb{R}^n)$. If $\phi : \mathbb{R}^n \to \mathbb{R}$ is nonnegative, radial, decreasing [as a function on $(0,\infty)$], and integrable, then*

$$\sup_{\epsilon>0}\{|\phi_\epsilon * f(x)|\} \leq ||\phi||_1 Mf(x)$$

Proof. First, assume ϕ is the simple function

$$\phi(x) = \sum_{i=1}^{m} a_i \chi_{B_i}(x)$$

where each $a_i > 0$ and B_i is the ball of radius r_i centered at 0. This simple function satisfies the hypothesis. Then

$$|\phi * f(x)| = \Big|\sum_{i=1}^{m} a_i |B_i| \frac{1}{|B_i|}\int_{B_i} f(x-y)\,dy\Big| \leq ||\phi||_1 Mf(x)$$

For general ϕ satisfying the hypothesis, let $\{\phi_k\}$ be a monotonic sequence of simple functions converging to ϕ pointwise. Then by the monotone convergence theorem,

$$\begin{aligned} |\phi * f(x)| &\leq \phi * |f|(x) \\ &= \lim_{k\to\infty} \phi_k * |f|(x) \\ &\leq \lim_{k\to\infty} ||\phi_k||_1 Mf(x) \\ &= ||\phi||_1 Mf(x) \end{aligned}$$

Any dilation ϕ_ϵ is also nonnegative, radial, decreasing [as a function on $(0,\infty)$], and integrable, with $||\phi_\epsilon||_1 = ||\phi||_1$ by change of variables. So for all $\epsilon > 0$,

$$|\phi_\epsilon * f(x)| \leq ||\phi||_1 Mf(x)$$

and the desired result follows. □

6. Schwartz Functions and Tempered Distributions

As stated in the introduction, the Hilbert transform defined by Equation (1.1) must be restricted to well-behaving functions. But we would like to extend H to $L^p(\mathbb{R})$ via its boundedness. So the domain of H must be a function space satisfying a few conditions:

(1) the functions overcome the kernel at the singularity $x = 0$.
(2) the functions decay sufficiently quickly.
(3) the function space is dense in $L^p(\mathbb{R})$, $1 \leq p < \infty$.

A candidate for this function space is the Schwartz space, the collection of smooth functions which decrease rapidly in the following sense:

Definition 6.1. The *Schwartz space* over $\mathbb{R}^n$, denoted by $\mathscr{S}(\mathbb{R}^n)$, is the set of all infinitely differentiable functions $f : \mathbb{R}^n \to \mathbb{C}$ such that

$$||f||_{\alpha,\beta} := \sup_{\mathbb{R}^n} |x^\alpha D^\beta f(x)| < \infty$$

for all multi-indices α, β.

One may show that $\{||\cdot||_{\alpha,\beta}\}$ is a countable collection of seminorms which separates points, so $(\mathscr{S}(\mathbb{R}^n), \{||\cdot||_{\alpha,\beta}\})/\mathbb{C}$ is a locally convex space. Under the natural topology induced by these seminorms, the Schwartz space is complete and metrizable. In addition, $\mathscr{S}(\mathbb{R}^n)$ is dense in L^p, $1 \leq p < \infty$.

Recall that the Fourier transform of a Schwartz function f, denoted by $\mathscr{F}(f) = \hat{f}$, is defined by

$$\hat{f}(x) = \int_{\mathbb{R}^n} f(y)e^{-2\pi i x\cdot y}\,dy$$

and the inverse Fourier transform of f, denoted by $\mathscr{F}^{-1}(f) = \check{f}$, is defined by

$$\check{f}(x) = \int_{\mathbb{R}^n} f(y)e^{2\pi i x\cdot y}\,dy$$

One can prove that $\mathscr{F}$ and $\mathscr{F}^{-1}$ are continuous linear transformations on $\mathscr{S}$ and, moreover, that $\mathscr{F}^{-1}\mathscr{F}(f) = \mathscr{F}\mathscr{F}^{-1}(f) = f$.

The dual of $\mathscr{S}(\mathbb{R}^n)$ will also be important in defining the Hilbert transform and demonstrating its boundedness.

Definition 6.2. The *space of tempered distributions*, denoted by $\mathscr{S}'(\mathbb{R}^n)$, is the space of continuous linear functionals $W : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$.

Remark 6.3. From the theory of locally convex spaces, we know that a linear functional W on $\mathscr{S}(\mathbb{R}^n)$ is continuous if and only if W is bounded by a linear combination of seminorms:

$$|W(f)| \leq \sum_{|\alpha|,|\beta|<m} ||f||_{\alpha,\beta} \quad \forall f \in \mathscr{S}(\mathbb{R}^n)$$

The classic example of a tempered distribution is integration against a fixed function $g \in \mathscr{S}$, defined by

$$W_g(f) =< g, f >= \int_{\mathbb{R}^n} g(x)f(x)\,dx \qquad \forall f \in \mathscr{S} \tag{6.4}$$

This functional clearly is linear. To see that it is continuous, observe that

$$| < g, f > | \leq \int_{\mathbb{R}^n} |g(x)f(x)|\,dx \leq ||g||_1||f||_{0,0}$$

In fact, Equation (6.4) defines a tempered distribution if g is any function satisfying

$$|g(x)| \leq C(1 + |x|)^k \tag{6.5}$$

for constants $C > 0$ and $k \in \mathbb{R}$.

Not every tempered distribution has this concrete form. Nevertheless, Equation (6.4) motivates the use of inner-product notation:

$$W(f) =< W, f >$$

With this notation, adjoint identities suggest a way to extend certain operations on $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$. A simple application of Fubini's theorem shows that for $g, f \in \mathscr{S}$,

$$< \hat{g}, f >= \int \hat{g}(x)f(x)\,dx = \int g(x)\hat{f}(x)\,dx =< g, \hat{f} > \tag{6.6}$$

Taking this equation as a template, we define the Fourier transform $\widehat{W}$ of a tempered distribution W by

$$< \widehat{W}, f >=< W, \hat{f} > \tag{6.7}$$

One can show that $\widehat{W}$ is itself a tempered distribution. Additionally, the inverse Fourier transform of a tempered distribution can be defined in the same way.

In a similar manner, we define multiplication and convolution of a Schwartz function and a tempered distribution. The Schwartz space is closed under both operations. Observe that for $g, f, h \in \mathscr{S}$,

$$< f \cdot g, h >=< g, f \cdot h >$$

Defining $\tilde{f}$ by $\tilde{f}(x) = f(-x)$, we see that

$$\begin{aligned} < g * f, h > &= \int \int g(y) f(x-y) h(x)\, dydx \\ &= \int g(y) \int \tilde{f}(y-x) h(x)\, dxdy \qquad \text{(Fubini's theorem)} \\ &= < g, \tilde{f} * h > \end{aligned}$$

So we define the product of $f \in \mathscr{S}$ and $W \in \mathscr{S}'$ by

$$< f \cdot W, h >=< W \cdot f, h >=< W, f \cdot h > \tag{6.8}$$

and the convolution by

$$< f * W, h >=< W * f, h >=< W, \tilde{f} * h > \tag{6.9}$$

Note that convolution (as well as multiplication) is associative, in the sense that

$$< (f_1 * f_2) * W, h >=< f_1 * (f_2 * W), h >$$

The Convolution-Multiplication Theorem states that the Fourier transform of the convolution of two Schwartz functions is the product of the two functions transformed:

$$\mathscr{F}(f * g) = \hat{f} \cdot \hat{g}$$

An identical equality holds with $\mathscr{F}^{-1}$ in place of $\mathscr{F}$. Furthermore, there is an analogous equation for the convolution defined above.

Proposition 6.10. *If $f \in \mathscr{S}(\mathbb{R}^n)$ and $W \in \mathscr{S}'(\mathbb{R}^n)$, then*

$$\mathscr{F}(f * W) = \hat{f} \cdot \widehat{W}$$

Proof. Observe that for all $h \in \mathscr{S}$,

$$\begin{aligned} < \mathscr{F}(f * W), h > &= < f * W, \hat{h} > \\ &= < W, \tilde{f} * \hat{h} > \\ &= < \widehat{W}, \mathscr{F}^{-1}(\tilde{f} * \hat{h}) > \\ &= < \widehat{W}, \mathscr{F}^{-1}(\tilde{f}) \cdot h > \end{aligned}$$

But $\mathscr{F}^{-1}(\tilde{f}) = \hat{f}$. Therefore,

$$< \mathscr{F}(f * W), h >=< \widehat{W}, \hat{f} \cdot h >=< \hat{f} \cdot \widehat{W}, h >$$

□

An alternative way to define convolution of $f \in \mathscr{S}$ and $W \in \mathscr{S}'$ uses the translation operator τ_x, given by $\tau_x f(y) = f(x+y)$:

$$(f * W)(x) = W(\tau_{-x}\tilde{f}) \tag{6.11}$$

Here, the convolution is a function. But one may verify that the tempered distribution defined by this function is the same as that defined by Equation (6.9).

For later reference, it is worth noting that Plancherel's Theorem is an immediate corollary of Equation (6.6). The theorem states that the (inverse) Fourier transform is an isometry from $L^2(\mathbb{R}^n)$ onto itself:

$$||\hat{f}||_2 = ||f||_2$$

To prove this equation for $f \in \mathscr{S}$, simply set $f = \overline{\hat{g}}$ in Equation (6.6), where the bar indicates complex conjugation. Then use the density of $\mathscr{S}$ to extend $\mathscr{F}$ to L^2, and the result follows.

For a more thorough treatment of the material in this section, see Reed and Simon [4] and Strichartz [6].

7. The Hilbert Transform

The convolution which defines the Hilbert transform is a *principal value integral*. The principal value of integration against the kernel $1/x$ may be generally defined as

$$p.v.(\frac{1}{x})f = \lim_{\epsilon \to 0^+} \int_{|x|>\epsilon} \frac{f(x)}{x}\,dx$$

where f is any function on $\mathbb{R}$ for which this limit is finite. As suggested before, the limit is finite if $f \in \mathscr{S}(\mathbb{R})$. In fact, more is true:

Proposition 7.1. *The functional $W = p.v.(\frac{1}{x})$ is a tempered distribution.*

Proof. Clearly, the operator is linear. Given $f \in \mathscr{S}(\mathbb{R})$, observe that

$$W(f) = \lim_{\epsilon \to 0^+} \int_{\epsilon<|x|<1} \frac{f(x)}{x}\,dx + \int_{|x|\geq 1} \frac{f(x)}{x}\,dx \tag{7.2}$$

We want to show that this quantity is bounded by a linear combination of seminorms. Because $1/x$ is odd and the range of integration is symmetric about the origin, the first term equals

$$\lim_{\epsilon \to 0^+} \int_{\epsilon<|x|<1} \frac{f(x)-f(0)}{x}\,dx$$

But f is smooth, so by the mean value theorem, there exists $x^* \in (0,x)$ such that

$$\frac{f(x)-f(0)}{x} = f'(x^*)$$

Hence, the first term in Equation (7.2) is dominated by $2||f||_{0,1}$. For the second term,

$$\int_{|x|\geq 1} \frac{f(x)}{x}\,dx = \int_{|x|\geq 1} \frac{f(x)}{x}\frac{x^2}{x^2}\,dx \leq 2||f||_{1,0}$$

Thus, by Remark 6.3, W is continuous. □

Now the Hilbert transform can be defined as a convolution with W in the sense of Equation (6.11).

Definition 7.3. The *Hilbert transform* is the operator H on $\mathscr{S}(\mathbb{R})$ defined by

$$Hf(x) = \frac{1}{\pi} f * W(x) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \int_{|y|>\epsilon} \frac{f(x-y)}{y}\, dy$$

The fact that W is a tempered distribution ensures that H is well-defined on the Schwartz space. We use the tools developed in Section 6 to prove that H satisfies the strong $(2,2)$ inequality on $\mathscr{S}$ and thereby extend H to all of L^2.

Theorem 7.4. *H is strong $(2,2)$.*

Proof. Suppose $f \in \mathscr{S}(\mathbb{R})$. By Equation (6.7),

$$\begin{aligned} < \widehat{W}, f > &= < W, \hat{f} > \\ &= \lim_{\epsilon \to 0^+} \int_{|x|>\epsilon} \frac{1}{x} \int_{\mathbb{R}} f(y) e^{-2\pi i x y}\, dy dx \\ &= \lim_{\epsilon \to 0^+} \int_{\mathbb{R}} f(y) \int_{|x|>\epsilon} \frac{-isin(2\pi x y)}{x}\, dx dy \qquad \text{(Fubini's Theorem)} \\ &= \lim_{\epsilon \to 0^+} \int_{\mathbb{R}} -isgn(y) f(y) \int_{|x|>2\pi|y|\epsilon} \frac{sin(x)}{x}\, dx dy \end{aligned}$$

The inner integral is uniformly bounded above by

$$\int_{\mathbb{R}} \frac{sin(x)}{x} = \pi$$

Hence, by the dominated convergence theorem,

$$< \widehat{W}, f >= \int_{\mathbb{R}} -i\pi sgn(y) f(y)\, dy$$

So in this case, $\widehat{W}$ is a function satisfying Equation (6.5): namely, $\widehat{W}(y) = -i\pi sgn(y)$. Then by Proposition 6.10,

$$\widehat{Hf}(x) = \frac{1}{\pi} \mathscr{F}(f * W)(x) = -isgn(x) \hat{f}(x)$$

Therefore, $||\widehat{Hf}||_2 = ||\hat{f}||_2$, and by Plancherel's Theorem,

$$||Hf||_2 = ||f||_2 \tag{7.5}$$

Now, take $f \in L^2(\mathbb{R})$. Because $\mathscr{S}$ is dense in L^2, there exists a sequence $\{f_n\} \subset \mathscr{S}(\mathbb{R})$ converging to f in L^2 norm. Equation (7.5) implies that $\{Hf_n\}$ is Cauchy and thus converges to an L^2 function. Defining Hf, the Hilbert transform of f, as this limit, we see that Equation (7.5) is satisfied on all of L^2. □

Equation (7.5) is used along with C-Z decomposition to prove that H satisfies the weak $(1,1)$ inequality on $\mathscr{S}(\mathbb{R})$.

Theorem 7.6. *If $f \in \mathscr{S}(\mathbb{R})$, then $||Hf||_{1,\infty} \leq C||f||_1$.*

Proof. Fix $\lambda > 0$, and assume $f \in \mathscr{S}(\mathbb{R})$ is real-valued and nonnegative. Let $\{I_j\}$ be the sequence of dyadic intervals in the C-Z decomposition of f at height λ. Let $\Omega = \bigcup_j I_j$, and write $f = g + b$ in accordance with Equations (4.9). H is linear, so if Hg and Hb are well-defined by Equation (1.1), then

$$d_{Hf}(\lambda) \leq d_{Hg}(\frac{\lambda}{2}) + d_{Hb}(\frac{\lambda}{2}) \tag{7.7}$$

Consider the first term on the right side:

$$\begin{aligned} d_{Hg}(\frac{\lambda}{2}) &\leq \frac{4}{\lambda^2}\int_{\mathbb{R}} Hg(x)^2\,dx && \text{(Chebyshev's inequality)} \\ &= \frac{4}{\lambda^2}\int_{\mathbb{R}} g(x)^2\,dx && \text{(Equation (7.5))} \\ &\leq \frac{8}{\lambda}\int_{\mathbb{R}} g(x)\,dx && \text{(Equation (4.10))} \\ &= \frac{8}{\lambda}(\int_{\Omega} g + \int_{\mathbb{R}\setminus\Omega} g) \leq \frac{16}{\lambda}||f||_1 && \text{(Equations (4.9))} \end{aligned}$$

So Hg is well-defined.

Now, consider the second term on the right side of Inequality (7.7). Let $\Omega^* = \bigcup_j \tilde{I}_j$, where $\tilde{I}_j$ is the the interval with the same center c_j as I_j but is twice as long. Then

$$\begin{aligned} d_{Hb}(\frac{\lambda}{2}) &\leq |\Omega^*| + |\{x \notin \Omega^* : |Hb(x)| > \frac{\lambda}{2}\}| \\ &\leq \frac{2}{\lambda}||f||_1 + \frac{2}{\lambda}\int_{\mathbb{R}\setminus\Omega^*} |Hb(x)|\,dx \qquad \text{(Lemma 4.8 and Chebyshev)} \end{aligned} \tag{7.8}$$

We want to show that the last integral is bounded above by $||f||_1$. To this end, we make two remarks.

(1) For $x \notin \Omega^*$,

$$Hb_j(x) = \frac{1}{\pi}\lim_{\epsilon\to 0^+}\int_{|y|>\epsilon}\frac{b_j(y)}{x-y}\,dy = \frac{1}{\pi}\int_{I_j}\frac{b_j(y)}{x-y}\,dy < \infty$$

because b_j vanishes outside I_j. Hence, the Hilbert transform of b_j in the sense of Equation (1.1) is well-defined.

(2) We claim that $|Hb(x)| \leq \sum_j |Hb_j(x)|$ a.e. This inequality follows immediately if the sum has a finite number of terms; otherwise, one may prove it using the fact that Hb_j converges to Hb in L^2 norm.

So we reduce the problem to showing

$$\int_{\mathbb{R}\setminus\Omega^*}\sum_j |Hb_j(x)|\,dx \leq C||f||_1 \tag{7.9}$$

Note that

$$\sum_j\int_{\mathbb{R}\setminus\Omega^*}|Hb_j(x)|\,dx = \frac{1}{\pi}\sum_j\int_{\mathbb{R}\setminus\Omega^*}\Big|\int_{I_j}\frac{b_j(y)}{x-y}\,dy\Big|\,dx \tag{7.10}$$

First, consider the inner integral. Because b_j has zero average and $x \notin \Omega^*$,

$$\Big|\int_{I_j}\frac{b_j(y)}{x-y}\,dy\Big| = \Big|\int_{I_j}\frac{b_j(y)}{x-y} - \frac{b_j(y)}{x-c_j}\,dy\Big| \leq \int_{I_j}\Big|\frac{b_j(y)(y-c_j)}{(x-y)(x-c_j)}\Big|\,dy$$

But $|y-c_j| \leq |I_j|/2$ and $|x-y| \geq |x-c_j|/2$, so the last term is at most

$$\int_{I_j}|b_j(y)|\frac{|I_j|}{(x-c_j)^2}\,dy \tag{7.11}$$

Now, observe that

$$\int_{\mathbb{R}\setminus\Omega^*} \frac{|I_j|}{(x-c_j)^2}\,dx \le \int_{\mathbb{R}\setminus I_j} \frac{|I_j|}{(x-c_j)^2}\,dx \le 4 \tag{7.12}$$

So substituting Expression (7.11) into Equation (7.10) and applying Fubini's Theorem gives

$$\begin{aligned}\sum_j \int_{\mathbb{R}\setminus\Omega^*} |Hb_j(x)|\,dx &\le \frac{1}{\pi}\sum_j \int_{\mathbb{R}\setminus\Omega^*}\int_{I_j} |b_j(y)|\frac{|I_j|}{(x-c_j)^2}\,dydx \\ &\le \frac{4}{\pi}\sum_j \int_{I_j} |b_j(y)|\,dy \\ &\le \frac{8}{\pi}||f||_1\end{aligned}$$

Therefore, Inequality (7.7) can be rewritten as

$$d_{Hf}(\lambda) \le (\frac{16}{\lambda} + \frac{2}{\lambda} + \frac{16}{\pi\lambda})||f||_1 \tag{7.13}$$

If f is complex, apply this argument to the positive and negative parts of the real and imaginary parts of f. □

For $f \in L^1(\mathbb{R})$, let $\{f_n\} \subset \mathscr{S}(\mathbb{R})$ be a sequence converging to f in L^1 norm. Then by Equation (7.13), $\{Hf_n\}$ is Cauchy in measure and thus converges to a measurable function. Defining Hf, the Hilbert transform of f, to be this measurable function, we see that H satisfies Equation (7.13) on all L^1. So H is weak $(1,1)$.

Since H is weak $(1,1)$ and strong $(2,2)$, the Marcinkiewicz Interpolation Theorem indicates that H is strong (p,p) for all $p \in (1,2)$. Consequently, the following duality argument, which uses the adjoint of H, shows that H satisfies the strong (p,p) inequality on $\mathscr{S}$, where $p \in (2,\infty)$.

The adjoint of H is the operator H' defined by

$$\int_{\mathbb{R}} Hf \cdot \overline{g} = \int_{\mathbb{R}} f \cdot \overline{H'g} \qquad \forall f, g \in \mathscr{S} \tag{7.14}$$

From this equation, one can determine that $\widehat{H'f}(x) = isgn(x)\hat{f}(x)$, implying that $H' = -H$. Then a density argument shows that Equation (7.14) holds with $g \in L^q$, $q \in (1,2)$.

Theorem 7.15. *If $f \in \mathscr{S}(\mathbb{R})$, then $||Hf||_p \le ||f||_p$ for all $p \in (2,\infty)$.*

Proof. Suppose $f \in \mathscr{S}$, and let q be the Hölder conjugate of p. The map

$$g \to \int_{\mathbb{R}} Hf\,\overline{g}$$

is a linear functional on L^q with norm equal to $||Hf||_p$. So

$$\begin{aligned}||Hf||_p &= \sup_{||g||_q=1} |\int Hf\,\overline{g}| \\ &= \sup_{||g||_q=1} |\int f\,\overline{Hg}| \\ &\le \sup_{||g||_q=1} ||f||_p||Hg||_q \qquad \text{(Hölder's inequality)} \\ &\le C||f||_p \qquad \text{(strong } (q,q) \text{ boundedness)}\end{aligned}$$

□

As before, we extend H to L^p by considering sequences of Schwartz functions converging in L^p norm.

8. The Truncated Hilbert Transform

Dropping the limit in Definition 7.3 gives a new operator which is defined on L^p:

Definition 8.1. The *truncated Hilbert transform* at height $\epsilon > 0$ is the operator H_ϵ on $L^p(\mathbb{R})$, $1 \leq p < \infty$, given by

$$H_\epsilon f(x) = \frac{1}{\pi}(\frac{1}{y}\chi_{\{|y|>\epsilon\}}) * f(x) = \frac{1}{\pi}\int_{|y|>\epsilon} \frac{f(x-y)}{y}\,dy$$

To see that H_ϵ is well-defined, observe that

$$|H_\epsilon f(x)| \leq C||f||_p||\frac{1}{y}\chi_{\{|y|>\epsilon\}}||_q < \infty \qquad \text{(Hölder's inequality)}$$

if and only if $q > 1$. Hence, any $p \in [1, \infty)$ is permitted.

Similar calculation as that in Proposition 7.4 shows that H_ϵ is strong $(2, 2)$ with a uniform bound for all ϵ; that is, the constant C in the definition of boundedness is independent of ϵ. Then the proof that H_ϵ is weak $(1, 1)$ and strong (p, p), $p \in (1, \infty)$, with uniform bounds follows as in Theorems 7.6 and 7.15. By these inequalities, if $\{f_n\}$ converges to f in L^p norm or in measure for $p = 1$, then $\{H_\epsilon f_n\}$ converges to $H_\epsilon f$ in norm or in measure, respectively. But we want to know if and how H_ϵ converges to H.

Proposition 8.2. *Suppose $f \in L^p$. Then $H_\epsilon f$ converges to Hf in norm for $p \in (1, \infty)$ and in measure for $p = 1$.*

Proof. Suppose $\{f_n\}$ converges to f in L^p norm, $1 < p < \infty$. Then the following chain of equalities holds.

$$||Hf||_p = \lim_{n\to\infty} ||Hf_n||_p = \lim_{n\to\infty}\lim_{\epsilon\to 0^+} ||H_\epsilon f_n||_p = \lim_{\epsilon\to 0^+}\lim_{n\to\infty} ||H_\epsilon f_n||_p = \lim_{\epsilon\to 0^+} ||H_\epsilon f||_p$$

The second and third equalities result from the strong (uniform) (p, p) boundedness of H and H_ϵ. If $p = 1$, replace convergence in norm with convergence in measure. □

Proposition 8.2 implies that, for a given $f \in L^p$, there exists a subsequence of $\{H_\epsilon f\}$ converging to Hf pointwise a.e. We want to show that the sequence itself converges to Hf pointwise:

Theorem 8.3. *If $f \in L^p(\mathbb{R})$, $1 \leq p < \infty$, then*

$$\lim_{\epsilon\to 0^+} H_\epsilon f(x) = Hf(x) \text{ a.e.} \tag{8.4}$$

The burden is only to show that the limit exists. By Definition 7.3, Equation (8.4) holds for $f \in \mathscr{S}$. The Schwartz space is dense in L^p, so by Theorem 2.6, it suffices to prove that the maximal operator H^* given by

$$H^* f(x) = \sup_{\epsilon>0}\{|H_\epsilon f(x)|\}$$

is weak (p, p) for all $p \in [1, \infty)$. In fact, H^* is strongly bounded if $p > 1$, as implied by the next theorem.

Theorem 8.5 (Cotlar's Inequality)**.** *If $f \in \mathscr{S}(\mathbb{R})$, then*

$$H^* f(x) \le MHf(x) + CMf(x)$$

where C is independent of f.

Proof. It suffices to show

$$|H_\epsilon f(x)| \le MHf(x) + CMf(x)$$

for all $\epsilon > 0$. Fix $\phi \in \mathscr{S}(\mathbb{R})$ that is nonnegative, even, decreasing, and integrable with $||\phi||_1 = 1$, and that has support $\{x \in \mathbb{R} : |x| \le \frac{1}{2}\}$. Recall that $\phi_\epsilon(x) = \epsilon^{-1}\phi(\epsilon^{-1}x)$. Letting $W = p.v.(\frac{1}{x})$, we have

$$\frac{1}{y}\chi_{\{|y|>\epsilon\}} = \phi_\epsilon * W(y) + \Big(\frac{1}{y}\chi_{\{|y|>\epsilon\}} - \phi_\epsilon * W(y)\Big)$$

which implies that

$$|H_\epsilon f(x)| \le |\frac{1}{\pi}(\phi_\epsilon * W) * f(x)| + \frac{1}{\pi}\Big|\frac{1}{y}\chi_{\{|y|>\epsilon\}} - \phi_\epsilon * W(y)\Big| * |f|(x)$$

By Proposition 5.4, the first term on the right side satisfies

$$|\frac{1}{\pi}(\phi_\epsilon * W) * f(x)| = |\frac{1}{\pi}\phi_\epsilon * (W * f)(x)| \le MHf(x)$$

Now consider the second term. Assume $\epsilon = 1$. We find a pointwise estimate for the kernel by examining two cases for the value of y.

(1) $|y| > 1$. Then

$$\begin{aligned} |\frac{1}{y} - \phi * W(y)| &= |\frac{1}{y} - \int_{|x|<1/2} \frac{\phi(x)}{y-x}\,dx| \\ &= |\int_{|x|<1/2} \phi(x)(\frac{1}{y} - \frac{1}{y-x})\,dx| \qquad (||\phi||_1 = 1) \\ &\le \int_{|x|<1/2} \frac{\phi(x)|x|}{|y||y-x|}\,dx \end{aligned}$$

Observe that

$$\frac{|y|}{|y-x|} \le \frac{|y|}{|y \pm 1/2|} \le 2$$

where $(+)$ is chosen if y is negative and $(-)$ is chosen if y is positive. So the last integral above is at most

$$\int_{|x|<1/2} \frac{\phi(x)}{y^2}\,dx = \frac{1}{y^2}$$

(2) $|y| < 1$. Then by the argument in Proposition 7.1,

$$\begin{aligned} |-\phi * W(y)| &= |\lim_{\delta\to 0}\int_{|x|>\delta} \frac{\phi(y-x)}{x}\,dx| \\ &= |\lim_{\delta\to 0}\int_{\delta<|x|<2} \frac{\phi(y-x)-\phi(y)}{x}\,dx| \\ &\le 4||\phi||_{0,1} := C' \end{aligned}$$

To combine these two bounds, note that

(1) if $|y| > 1$, then $(1+y^2)/y^2 \le 2$, and
(2) if $|y| < 1$, then $C'(1+y^2) \le 2C'$.

Hence,

$$\frac{1}{\pi}|\frac{1}{y}\chi_{\{|y|>1\}} - \phi * W(y)| \leq \frac{C''}{1+y^2}$$

So by Proposition 5.4,

$$\frac{1}{\pi}\Big|\frac{1}{y}\chi_{\{|y|>1\}} - \phi * W(y)\Big| * |f|(x) \leq \frac{C''}{1+y^2} * |f|(x) \leq CMf(x) \tag{8.6}$$

The following dilation argument proves that this inequality holds when we replace 1 by ϵ. For any $f \in \mathscr{S}$, let f^ϵ be the dilation of f by ϵ, given by $f^\epsilon(x) = f(\epsilon x)$. Let $g(y) = \frac{1}{\pi}|\frac{1}{y}\chi_{\{|y|>1\}} - \phi * W(y)|$. One may show that $g_\epsilon(y) = \frac{1}{\pi}|\frac{1}{y}\chi_{\{|y|>\epsilon\}} - \phi_\epsilon * W(y)|$. Then we have the following equations:

$$\begin{aligned} g * f^\epsilon(\epsilon^{-1}x) &= \int_{\mathbb{R}} f(\epsilon y) g(\epsilon^{-1}x - y)\, dy \\ &= \int_{\mathbb{R}} f(y)\epsilon^{-1} g(\epsilon^{-1}x - \epsilon^{-1}y)\, dy \\ &= g_\epsilon * f(x) \end{aligned}$$

$$\begin{aligned} Mf^\epsilon(x) &= \sup \frac{1}{|I_r|}\int_{I_r} |f(\epsilon x - \epsilon y)|\, dy \\ &= \sup \frac{1}{|\epsilon I_r|}\int_{|\epsilon I_r|} |f(\epsilon x - y)|\, dy \\ &= Mf(\epsilon x) \end{aligned}$$

Therefore, by Inequality (8.6),

$$|g_\epsilon * f(x)| = |g * f^\epsilon(\epsilon^{-1}x)| \leq CMf^\epsilon(\epsilon^{-1}x) = CMf(x)$$

□

By the theorem, if $p \in (1, \infty)$ and $f \in \mathscr{S}$, then

$$||H^* f||_p \leq ||MHf||_p + ||CMf||_p \leq C'||f||_p$$

where the last inequality holds because both M and H are strong (p, p). The fact that H^* is strong (p, p) thus follows. It remains to show

Theorem 8.7. *H^* is weak $(1, 1)$.*

Proof. The argument proceeds initially as in Theorem 7.6. Fix $\lambda > 0$, and assume $f \in L^1$ is real-valued and nonnegative. Let $\{I_j\}$ be the sequence of dyadic intervals in the C-Z decomposition of f at height λ, and write $f = g + b$ in accordance with Equations (4.9). We have

$$d_{H^*f}(\lambda) \leq d_{H^*g}(\frac{\lambda}{2}) + d_{H^*b}(\frac{\lambda}{2})$$

Because H^* is strong $(2, 2)$, bounding d_{H^*g} follows in precisely the same way as bounding d_{Hg} in Theorem 7.6. Recalling Inequality (7.8) and noting that $||b||_1 \leq 2||f||_1$, we reduce the problem to showing

$$|\{x \notin \Omega^* : H^*b(x) > \lambda\}| \leq \frac{C}{\lambda}||b||_1$$

where we have made the change $\lambda/2 \to \lambda$ for convenience.

Fix $x \notin \Omega^*$ and $\epsilon > 0$, and consider b_j, which has zero average and vanishes outside I_j. Only one of the following holds:

(1) $(x-\epsilon, x+\epsilon)\bigcap I_j = I_j$. Then

$$H_\epsilon b_j(x) = \frac{1}{\pi}\int_{|y|>\epsilon} \frac{b_j(x-y)}{y}\,dy = 0$$

because $x-y \notin I_j$ if $|y|>\epsilon$.

(2) $(x-\epsilon, x+\epsilon)\bigcap I_j = \emptyset$. Then

$$H_\epsilon b_j(x) = \frac{1}{\pi}\int_{I_j} \frac{b_j(x-y)}{y}\,dy = Hb_j(x)$$

because $I_j \subset \{x-y : |y|>\epsilon\}$. So by Equation (7.11),

$$|H_\epsilon b_j(x)| \leq \frac{|I_j|}{(x-c_j)^2}||b_j||_1$$

(3) either $x-\epsilon$ or $x+\epsilon$ is in I_j. Then $(x-3\epsilon, x+3\epsilon)\supset I_j$, and for all $y\in I_j$, $|x-y|>\frac{\epsilon}{3}$. Thus,

$$|H_\epsilon b_j(x)| \leq \int_{I_j}\frac{|b_j(y)|}{|x-y|}\,dy \leq \frac{3}{\epsilon}\int_{x-3\epsilon}^{x+3\epsilon}|b_j(y)|$$

Hence, summing $|H_\epsilon b_j(x)|$ over all j, we see

$$\begin{aligned}
|H_\epsilon b(x)| &\leq \sum_j \Big[\frac{|I_j|}{(x-c_j)^2}||b_j||_1 + \frac{3}{\epsilon}\int_{x-3\epsilon}^{x+3\epsilon}|b_j(y)|\Big] \\
&= \sum_j \frac{|I_j|}{(x-c_j)^2}||b_j||_1 + \frac{3}{\epsilon}\int_{x-3\epsilon}^{x+3\epsilon}|b(y)| \\
&= \sum_j \frac{|I_j|}{(x-c_j)^2}||b_j||_1 + C'Mb(x) \qquad \text{(Definition 5.1)}
\end{aligned}$$

The last estimate is independent of ϵ and so holds if we replace $|H_\epsilon b(x)|$ with $H^*b(x)$. Therefore,

$$\begin{aligned}
&|\{x\notin\Omega^* : H^*b(x)>\lambda\}| \\
&\quad\leq |\{x\notin\Omega^* : \sum_j \frac{|I_j|}{(x-c_j)^2}||b_j||_1 > \frac{\lambda}{2}\}| + |\{x\in\mathbb{R} : Mb(x) > \frac{\lambda}{2C'}\}| \\
&\quad\leq \frac{2}{\lambda}\int_{\mathbb{R}\setminus\Omega^*}\sum_j \frac{|I_j|}{(x-c_j)^2}||b_j||_1\,dx + \frac{2C'}{\lambda}||b||_1 \\
&\quad\leq (\frac{C''+2C'}{\lambda})||b||_1
\end{aligned}$$

where the second inequality follows from Chebyshev and the weak $(1,1)$ boundedness of M and the last inequality from Equation (7.12). If f is complex, apply this argument to the positive and negative parts of the real and imaginary parts of f. □

9. Conclusion

Using real-variable methods, we proved that Equation (1.1) holds for $f\in L^p(\mathbb{R},\mu)$, $p\in[1,\infty)$, up to a μ-null set. But the techniques we used and the tools we constructed go beyond the special focus of this paper. The Hilbert transform is an important operator to investigate on its own, but it naturally segues into broader studies in harmonic analysis.

Acknowledgments. I would like to thank my mentors Olga Turanova, Timur Akhunov, and Jessica Lin for their patience, their guidance, and their corrections. I especially thank Ms. Turanova for repeatedly editing this paper and Ms. Lin for suggestions on several proofs. I would also like to thank Professor Peter May for his feedback and for organizing the REU, making all this possible.

References

[1] Kunal Chaudhury. L^p-boundedness of the Hilbert Transform. big-www.epfl.ch/chaudhury/Lp_boundedness_HT.pdf.
[2] Javier Duoandikoetxea. Fourier Analysis. Trans. by David Cruz-Uribe. American Mathematical Society. 2001
[3] Loukas Grafakos. Classical and Modern Fourier Analysis. Springer. 2008.
[4] Michael Reed and Barry Simon. Methods of Modern Mathematical Physics I: Functional Analysis Academic Press. 1972.
[5] Walter Rudin. Functional Analysis. McGraw-Hill. 2006.
[6] Robert Strichartz. A Guide to Distribution Theory and Fourier Transforms. World Scientific. 2003.