A SHORT TOUR OF HARMONIC ANALYSIS

BLAINE TALBUT

ABSTRACT. We present a selection of introductory results from harmonic analysis, beginning with the fundamentals of harmonic analysis on the circle $\mathbb{T}$ and the real line.

CONTENTS

1. INTRODUCTION

Our goal in this paper is to introduce the reader to some of the fundamental techniques and results of harmonic analysis. We begin by introducing the Fourier series of functions on the circle $\mathbb{T}$. The most important property of the Fourier series is that it converges in L^p norm back to the original function for $1 < p < \infty$; we will show an equivalent result, that the Hilbert transform is bounded as an operator from $L^p(\mathbb{T})$ to $L^p(\mathbb{T})$ for all such p. We then turn to the real-variable analogue of the Fourier series, the Fourier transform. Using the Fourier transform and a technique called Calderón-Zygmund decomposition, we will show that the real-variable Hilbert transform and the class of its natural generalizations—namely, the Calderón-Zygmund operators—are also bounded as operators from $L^p(\mathbb{R})$ to $L^p(\mathbb{R})$ for $1 < p < \infty$. We will conclude with a proof of a somewhat weaker version of the previous result which does not require the Fourier transform and which introduces Cotlar's lemma, an important result, which is of enduring usefulness in harmonic analysis.

We assume familiarity with the basic results of functional analysis—in particular, the theory of Hilbert spaces. We also assume familiarity with Lebesgue integral and L^p spaces, including convolution and interpolation results. The first chapter of Muscalu and Schlag [3] contains the background in convolution and interpolation needed to read this paper, though it merely states the Marcinkiewicz and Riesz-Thorin interpolation theorems; the proofs of these are contained in the classic

Date: August 30, 2013.

textbook on Fourier analysis, Stein and Weiss [4]. All integrals that appear in this paper are in the sense of Lebesgue.

2. Fourier Analysis on the Circle—Basic Results

Our goal in this section is to introduce the Fourier series, a technique for representing general periodic functions as (usually infinite) sums of trigonometric functions—the most fundamental periodic functions. Euler's formula tells us that we may equivalently consider sums over complex exponential functions, which allows us to write many of our results in a much more elegant form. Consider the space $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Intuitively, $\mathbb{T}$ can be thought of as the unit interval $[0, 1]$ with the endpoints identified. Alternatively, it can be realized as the unit circle in $\mathbb{C}$ via the homeomorphism $t \mapsto e^{2\pi it}$. The former picture makes clear that we can consider any periodic function on $\mathbb{R}$ as a function on $\mathbb{T}$, possibly with some horizontal scaling.

Definition 2.1. Let $f \in L^1(\mathbb{T})$. We define the *Fourier coefficients* of f for all $n \in \mathbb{Z}$ by

$$\hat{f}(n) = \int_{\mathbb{T}} f(x)e^{-2\pi inx}dx.$$

Note that this integral is well-defined for all $f \in L^1(\mathbb{T})$. Since $\mathbb{T}$ is compact, we have that $L^p(\mathbb{T}) \subseteq L^1(\mathbb{T})$ for all $1 \leq p \leq \infty$. Thus, $\hat{f}$ is well-defined for $f \in L^p(\mathbb{T})$, $1 \leq p \leq \infty$.

We now state some basic properties of the Fourier coefficients.

Proposition 2.2. *For all $f, g \in L^1(\mathbb{T})$ and all $n \in \mathbb{Z}$, we have*

(1) $(\widehat{f+g})(n) = \hat{f}(n) + \hat{g}(n)$.
(2) $\widehat{\lambda f}(n) = \lambda \hat{f}(n)$.
(3) $(\widehat{f * g})(n) = \hat{f}(n)\hat{g}(n)$.
(4) $(\widehat{f(\cdot + y)})(n) = \hat{f}(n)e^{-2\pi iny}$.
(5) If f is continuously differentiable, then $\widehat{f'}(n) = (2\pi in)\hat{f}(n)$.
(6) $\sup_{n\in\mathbb{Z}} |\hat{f}(n)| \leq \|f\|_{L^1(\mathbb{T})}$.
(7) (Riemann-Lebesgue lemma) $\lim_{|n|\to\infty} |\hat{f}(n)| = 0$.

The proof of these properties is left to the reader; we note, however, that (3) is proved with Fubini's theorem, (5) with integration by parts, and (7) by the density of compactly supported C^∞ functions in $\mathbb{T}$.

Definition 2.3. Given $f \in L^1(\mathbb{T})$, we define the *N^{th} partial sum of the Fourier series of f* by

$$S_N f(x) = \sum_{n=-N}^{N} \hat{f}(n)e^{2\pi nx}$$

for each $N \in \mathbb{N}$. We associate formally with f the *Fourier series* defined by

$$\mathscr{F} f(x) = \lim_{N\to\infty} S_N f(x)$$

wherever the above limit exists.

It is not clear whether this limit exists in any sense or if its limit is f. However, we will shortly see a reason to suspect this is the case for nice enough f.

First, note that we can write

$$\begin{aligned} S_N f(x) &= \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x} \\ &= \sum_{n=-N}^{N} \int_{\mathbb{T}} f(t) e^{2\pi i n (x-t)} dt \\ &= \int_{\mathbb{T}} f(t) \left(\sum_{n=-N}^{N} e^{2\pi i n (x-t)} \right) dt. \end{aligned}$$

If $D_N(x) = \sum_{n=-N}^{N} e^{2\pi n x}$, then the above gives us $S_N f(x) = (D_N * f)(x)$.

Definition 2.4. The N^{th} *Dirichlet kernel* D_N is defined by

$$D_N(x) = \sum_{n=-N}^{N} e^{2\pi i n x}.$$

Proposition 2.5. *For all $N \in \mathbb{N}$ and $x \in \mathbb{T}$, we have that*

$$D_N(x) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)}.$$

The precise derivation of the above is excluded for reasons of tedium; we merely remark that it is based on the equality

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

The above proposition suggests the reason a function's Fourier series might converge to the original function: the Dirichlet kernel resembles an approximation to the identity, growing without bound at the origin and decaying to 0 away from it. If the Dirichlet kernel *were* an approximation to the identity, then the job of showing that the Fourier series converges would be simple indeed; nearly everything we would want to say about the Fourier series would follow from the basic properties of an approximation to the identity. Unfortunately, this is not the case; it is a consequence of the following theorem that the Dirichlet kernel cannot be an approximation to the identity.

Theorem 2.6. *For all $N \in \mathbb{N}$ and some constant $C > 0$ we have*

$$\|D_N\|_{L^1(\mathbb{T})} \geq C \log N.$$

Proof. We have the following inequality for all $x \in [0, \frac{\pi}{2}]$:

$$\frac{|x|}{2} \leq |\sin(x)| \leq |x|.$$

Thus

$$\begin{aligned}\|D_N\|_{L^1(\mathbb{T})} &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{|\sin((2N+1)\pi x)|}{|\sin(\pi x)|}\,\mathrm{d}x \\ &= 2\int_0^{\frac{1}{2}} \frac{|\sin((2N+1)\pi x)|}{|\sin(\pi x)|}\,\mathrm{d}x \\ &\geq \int_0^{\frac{1}{2}} \frac{|\sin((2N+1)\pi x)|}{|x|}\,\mathrm{d}x.\end{aligned}$$

Now, consider the intervals $I_n = \left(\frac{n+1/4}{2N+1}, \frac{n+3/4}{2N+1}\right)$, $n = 0, \ldots, N-1$. Each I_n has length $\frac{1}{4N+2}$, and we have $|\sin((2N+1)\pi x)| > \frac{1}{2}$ on each I_n. We therefore have

$$\begin{aligned}\int_0^{\frac{1}{2}} \frac{|\sin((2N+1)\pi x)|}{|x|}\,\mathrm{d}x &\geq \sum_{n=0}^{N-1} \int_{I_n} \frac{|\sin((2N+1)\pi x)|}{|x|}\,\mathrm{d}x \\ &= \frac{1}{2}\sum_{n=0}^{N-1} \int_{I_n} \frac{1}{x}\,\mathrm{d}x \\ &= \frac{1}{2}\sum_{n=0}^{N-1} \frac{2N+1}{(4N+2)(n+1/4)} \\ &\geq C\log N\end{aligned}$$

for some $C > 0$. □

Despite this failure, if the Fourier series of a function is to converge to the original function, we expect the Dirichlet kernel to exhibit some approximate identity–esque behavior. We might therefore ask if the Dirichlet kernel can be "smoothed out" in such a way as to produce an approximation to the identity. This "smoothing out" is done by taking arithmetic means of the family $\{D_N\}_{N\in\mathbb{N}}$.

Definition 2.7. We define the *Fejér kernel* F_N by

$$F_N = \frac{1}{N+1}\sum_{k=0}^{N} D_k.$$

We have the following identity for $F_N(x)$, analogous to that of Proposition 3.5:

Proposition 2.8. *For all $N \in \mathbb{N}$ and all $x \in \mathbb{T}$, we have that*

$$F_N(x) = \frac{1}{N+1}\left(\frac{\sin((2N+1)\pi x)}{\sin(\pi x)}\right)^2.$$

Theorem 2.9. *The sequence $\{F_N\}_{N\in\mathbb{N}}$ is an approximation to the identity.*

Proof. It is clear from Definition 2.4 that $\int_{\mathbb{T}} D_N(x)\,\mathrm{d}x = 1$ for all $N \in \mathbb{N}$; from this it follows that $\int_{\mathbb{T}} F_N(x)\,\mathrm{d}x = 1$ for all $N \in \mathbb{N}$. Since $F_N \geq 0$, it follows that $\sup_{N\in\mathbb{N}} \|F_N\|_{L^1(\mathbb{T})} = 1 < \infty$. Finally, it follows from Proposition 2.8 that for any neighborhood U of the origin we have $\int_{\mathbb{T}\setminus U} |F_N(t)|\,\mathrm{d}t \to 0$ as $N \to \infty$. □

Definition 2.10. For $N \in \mathbb{N}$ the expressions

$$(f * F_N)(x) = \sum_{|k|\leq N}\left(1 - \frac{|k|}{N}\right)\hat{f}(k)e^{2\pi ikx}$$

are called the *Fejér means* of f.

For any set of complex coefficients $\{a_k\}_{|k|\leq N}$, we call the function f defined by $f(x) = \sum_{k=-N}^{N} a_k e^{2\pi ikx}$ a *trigonometric polynomial* of degree N. The Fejér kernel allows us to quite easily show the following fundamental and useful results:

Theorem 2.11. *The set of trigonometric polynomials is dense in $L^p(\mathbb{T})$ for $1 \leq p < \infty$.*

Proof. Let $f \in L^p(\mathbb{T})$ for p as above. Note that $f * F_N$ is a trigonometric polynomial. Theorem 2.5 gives that $f * F_N$ converges to f in L^p norm as $N \to \infty$. □

Theorem 2.12. *Let $f, g \in L^1(\mathbb{T})$. If $\hat{f}(m) = \hat{g}(m)$ for all $m \in \mathbb{Z}$, then $f = g$ a.e.*

Proof. If $\widehat{(f-g)}(m) = \hat{f}(m) - \hat{g}(m) = 0$ for all $m \in \mathbb{Z}$, then $(f-g) * F_N = 0$ for all $N \in \mathbb{N}$. Theorem 2.5 therefore gives that

$$\|(f-g) - f * F_N\|_{L^1(\mathbb{T})} \to 0$$

as $N \to \infty$. It follows that $\|f-g\|_{L^1(\mathbb{T})} = 0$, so $f = g$ a.e. □

The Fejér means are a key to unlocking a number of useful results about pointwise convergence. Since the Fejér kernel is an approximation to the identity, we know that the Fejér means of any continuous function f converge pointwise (in fact, uniformly) to f. If the Fourier series of any continuous function converges pointwise, then we know it converges to f, since any convergent sequence must converge to the limit of its arithmetic means. Unfortunately, it is not the case that the Fourier series of every continuous function converges pointwise, although we will not show this here. We do, however, have the following encouraging result:

Theorem 2.13. *Let $f \in L^1(\mathbb{T})$ satisfy*

$$\int_{\mathbb{T}} \frac{|f(x) - f(a)|}{|x-a|} \, \mathrm{d}x < \infty$$

for some $a \in \mathbb{T}$. Then $\mathscr{F}(f)(a) = f(a)$.

Proof. We may assume without loss of generality that $a = 0$ and $f(a) = 0$ (since translation and addition of a constant do not change the convergence of the Fourier series). Our problem then reduces to that of showing that $(D_N * f)(0)$ vanishes as $N \to \infty$. Note that we can write

$$\begin{aligned}(D_N * f)(0) &= \int_0^1 f(x) \frac{\sin(-(2N+1)\pi x)}{\sin(-\pi x)} \, \mathrm{d}x \\ &= \mathrm{Im}(\hat{g}(2N+1))\end{aligned}$$

where

$$g(x) = -\frac{f(x)}{\sin(\pi x)}.$$

Note that we used that sine is even to simplify the expression in the first line. g is clearly integrable over $|x| > \frac{1}{2}$, but, since $|\sin(\pi x)| \geq \frac{\pi|x|}{2}$ for $|x| \leq \frac{1}{2}$, it follows from the hypothesis that g is integrable over $|x| \leq \frac{1}{2}$. The desired result therefore follows from the Riemann-Lebesgue lemma (Proposition 2.2(7)). □

Since many functions encountered in physical applications are Lipschitz continuous, this is a powerful result. Nevertheless, for many applications—and for purely mathematical satisfaction—it is necessary to consider a more global form of convergence.

3. L^p Convergence of the Fourier Series on the Circle

Our goal in this section is to answer the question: "For what p does the Fourier series of a general $L^p(\mathbb{T})$ function converge in L^p norm back to the original function?" It is trivial to observe that we cannot have $p = \infty$, since a uniformly convergent series of continuous functions must converge to a continuous function, and most $L^\infty(\mathbb{T})$ functions are not continuous. We therefore pose the question: "Does the Fourier series of a general $C(\mathbb{T})$ function converge in L^∞ norm back to the original function?" We ask the reader to interpret all references to the Fourier series of $L^p(\mathbb{T})$ functions in the following section as referring to $C(\mathbb{T})$ in the case of $p = \infty$.

Proposition 3.1. *The set $\{\phi_k\}_{k\in\mathbb{Z}}$ with $\phi_k(x) = e^{2\pi ikx}$ is an orthonormal basis for $L^2(\mathbb{T})$ with respect to the inner product $\langle f|g\rangle = \int_{\mathbb{T}} f(x)\overline{g(x)}dx$.*

Proof. For $k \neq \ell$, we have

$$\langle \phi_k|\phi_\ell\rangle = \int_{\mathbb{T}} e^{2\pi i(k-\ell)x}dx = 0.$$

This shows that $\{\phi_k\}_{k\in\mathbb{Z}}$ is an orthonormal set. Theorem 3.13 gives that, if $f \in L^1(\mathbb{T})$ satisfies $\hat{f}(m) = \langle f|\phi_m\rangle = 0$ for all $m \in \mathbb{Z}$, then $f = 0$ a.e. This completes the proof. □

The theory of Hilbert spaces immediately yields the following corollary:

Corollary 3.2. *Let $f \in L^2(\mathbb{T})$. We have that*

(1) (Plancherel's theorem)

$$\|f\|^2_{L^1(\mathbb{T})} = \sum_{m\in\mathbb{Z}} |\hat{f}(m)|^2.$$

(2) $\|f - f * D_N\|_{L^2(\mathbb{T})} \to 0$ *as* $N \to \infty$.

Thus, we have that the Fourier series converges in L^p for at least one p. However, our proof of this result relied on the unique structure of the L^2 space and is clearly not generalizable to the other L^p spaces. Rather than attempt to deal with general L^p convergence directly, we will immediately consider a reformulation of the problem.

Lemma 3.3. *Let $1 \leq p \leq \infty$. The following are equivalent:*

(1) $\|S_N f - f\|_{L^p(\mathbb{T})} \to 0$ *as* $n \to \infty$ *for all* $f \in L^p(\mathbb{T})$.

(2) $\sup_{N\in\mathbb{N}} \|S_N\|_{L^p\to L^p} < \infty$.

If $p = \infty$, we replace $L^p(\mathbb{T})$ with $C(\mathbb{T})$ above.

Proof. That (1) implies (2) is an immediate consequence of the uniform boundedness principle. Conversely, note that (1) is clear if f is a trigonometric polynomial. Therefore, if (2), then by density of the trigonometric polynomials (1) holds for all $f \in L^p(\mathbb{T})$. □

The above lemma is often useful in its own right; indeed, it is a fundamental result in any discussion of the L^p convergence of the Fourier transform. However, it interests us here only as a stepping-stone to the next theorem, which is the desired reformulation of the problem of L^p convergence.

Definition 3.4. Given $f \in L^1(\mathbb{T})$, we associate formally with f the *Hilbert transform* of f given by

$$Hf(x) = \sum_{n\in\mathbb{Z}} \operatorname{sgn}(n)\hat{f}(n)e^{2\pi inx}.$$

We furthermore associate with f the *Riesz projection* of f given by

$$P_+f(x) = \sum_{m=0}^{\infty} \hat{f}(m)e^{2\pi imx}.$$

Note that, in practice, we shall only have to consider the Hilbert transforms and Riesz projections of trigonometric polynomials, for which both of the above expressions are well-defined.

The following proposition will prove useful:

Proposition 3.5. *The Hilbert transform is self-adjoint; that is, for any trigonometric polynomials f, g we have*

$$\langle Hf|g\rangle = \int_0^1 Hf(x)\overline{g(x)}\,\mathrm{d}x = \int_0^1 f(x)\overline{Hg(x)}\,\mathrm{d}x = \langle f|Hg\rangle.$$

Proof. We have

$$\begin{aligned}\int_0^1 Hf(x)\overline{g(x)}\,\mathrm{d}x &= \int_0^1 \left(\sum_{m\in\mathbb{Z}} \operatorname{sgn}(m)\hat{f}(m)e^{2\pi imx}\right)\overline{g(x)}\,\mathrm{d}x \\ &= \sum_{m\in\mathbb{Z}} \hat{f}(m)\operatorname{sgn}(m)\int_0^1 \overline{g(x)}e^{2\pi imx}\,\mathrm{d}x \\ &= \sum_{m\in\mathbb{Z}} \operatorname{sgn}(m)\overline{\hat{g}(m)}\int_0^1 f(y)e^{2\pi imy}\,\mathrm{d}y \\ &= \int_0^1 f(y)\left(\sum_{m\in\mathbb{Z}} \operatorname{sgn}(m)\overline{\hat{g}(m)}e^{2\pi imy}\right)\mathrm{d}y \\ &= \int_0^1 f(y)\overline{Hg(y)}\,\mathrm{d}y.\end{aligned}$$

Note that we can interchange integrals and sums freely because here all sums are finite. □

It is worth mentioning that the Hilbert transform, as usually defined, instead takes the form

$$Hf(x) = \sum_{n\in\mathbb{Z}} -i\operatorname{sgn}(n)\hat{f}(n)e^{2\pi inx}.$$

However, as will shortly become clear, the main property of the Hilbert transform that interests us is its boundedness, and clearly the boundedness of the operator under one definition is equivalent to its boundedness under the other.

Theorem 3.6. *Let $1 \le p \le \infty$. The following are equivalent:*

(1) $\|S_N f - f\|_{L^p(\mathbb{T})} \to 0$ *as* $N \to \infty$ *for all* $f \in L^p(\mathbb{T})$.
(2) There exists a constant $C_p > 0$ *such that* $\|Hf\|_{L^p(\mathbb{T})} < C_p\|f\|_{L^p(\mathbb{T})}$ *for all trigonometric polynomials* f.

Again, for $p = \infty$ *we replace* $L^p(\mathbb{T})$ *with* $C(\mathbb{T})$.

Proof. We begin by noting that $P_+ f = \frac{1}{2}(f + Hf) + \frac{1}{2}\hat{f}(0)$. Thus, boundedness of the Hilbert transform is equivalent to boundedness of the Riesz projection. We also note that the Riesz projection is a bounded operator if and only if $\sup_{N\in\mathbb{N}} \|S_{N,+}\|_{L^p\to L^p} < \infty$, where

$$S_{N,+}f(x) := \sum_{m=0}^{N} \hat{f}(m)e^{2\pi imx}.$$

The proof of this result is similar to the proof of Lemma 5.5 and is omitted here; we note, however, that the Riesz transform was defined only for trigonometric polynomials, while $S_{N,+}$ is defined for all L^p functions. Density of the trigonometric functions in $L^p(\mathbb{T})$ allows us to pass from one domain to the other in this way. It now remains to show that

$$\sup_{N\in\mathbb{N}} \|S_N\|_{L^p\to L^p} < \infty \iff \sup_{N\in\mathbb{N}} \|S_{N,+}\|_{L^p\to L^p}\infty.$$

Note that

$$\begin{aligned}
\|S_n f\|_{L^p(\mathbb{T})} &= \left\| \sum_{m=-N}^{N} \hat{f}(m)e^{2\pi imx} \right\|_{L^p(\mathbb{T})} \\
&= \left\| e^{-2\pi iNx} \sum_{m=0}^{2N} \hat{f}(m-N)e^{2\pi imx} \right\|_{L^p(\mathbb{T})} \\
&= \left\| S_{n,+}\left(e^{2\pi iN(\cdot)}f\right) \right\|_{L^p(\mathbb{T})}.
\end{aligned}$$

Since $\|f\|_{L^p(\mathbb{T})} = \left\|e^{2\pi iN(\cdot)}f\right\|_{L^p(\mathbb{T})}$, this completes the proof. □

We have now reduced the problem of determining the L^p convergence of the Fourier series to the problem of determining the L^p boundedness of the Hilbert transform. It is then a consequence of the next theorem that the Fourier series converges in L^p norm for all $1 < p < \infty$.

Theorem 3.7. *For* $1 < p < \infty$*, there exists a constant* $C_p > 0$ *such that, for all trigonometric functions* f*, we have*

$$\|Hf\|_{L^p(\mathbb{T}} \le C_p\|f\|_{L^p(\mathbb{T})}.$$

Proof. Our strategy is to first show that the theorem holds for even p. Once we have established this case, interpolation allows us to extend our result to all $2 \le p < \infty$. Then, duality allows us to extend our result to $1 < p < 2$.

Let f be a real-valued trigonometric polynomial of degree N with $\hat{f}(0) = 0$. Since f is real-valued, we have that $\hat{f}(-m) = \overline{\hat{f}(m)}$ for all $m \in \mathbb{Z}$. Thus,

$$\begin{aligned} Hf(t) &= \sum_{m>0} \hat{f}(m) e^{2\pi imt} - \sum_{m>0} \hat{f}(-m) e^{-2\pi imt} \\ &= 2i \operatorname{Im}\left(\sum_{m>0} \hat{f}(m) e^{2\pi imt} \right) \end{aligned}$$

so that Hf is imaginary-valued.

Now, we can write

$$(f + Hf)(t) = 2 \sum_{m=1}^{N} \hat{f}(m) e^{2\pi imt}$$

so that, for $k \in \mathbb{N}$, we have

$$\begin{aligned} \int_0^1 (f(t) + Hf(t))^{2k} \,\mathrm{d}t &= \int_0^1 \left(2 \sum_{m=1}^{N} \hat{f}(m) e^{2\pi imt} \right)^{2k} \mathrm{d}t \\ &= 4^k \int_0^1 \sum_{m=2}^{2N} \left(\sum_{\ell=\max(1,m-N)}^{\min(N,m-1)} \hat{f}(\ell)\hat{f}(m-\ell) \right) e^{2\pi imt} \,\mathrm{d}t \\ &= 4^k \sum_{m=2}^{2N} \left(\sum_{\ell=\max(1,m-N)}^{\min(N,m-1)} \hat{f}(\ell)\hat{f}(m-\ell) \right) \int_0^1 e^{2\pi imt} \,\mathrm{d}t \\ &= 0. \end{aligned}$$

Expanding the $2k^{\text{th}}$ power in the first expression above and taking real parts yields

$$\sum_{j=0}^{k} \binom{2k}{2j} \int_0^1 (Hf)(t)^{2k-2j} f(t)^{2j} \,\mathrm{d}t = 0,$$

where we use that f is real-valued and Hf imaginary-valued. Thus, a simple application of the triangle inequality yields

$$\begin{aligned} \|Hf\|_{L^{2k}(\mathbb{T})}^{2k} &\le \sum_{j=0}^{k} \binom{2k}{2j} \int_0^1 |(Hf)(t)|^{2k-2j} f(t)^{2j} \,\mathrm{d}t \\ &\le \sum_{j=1}^{k} \binom{2k}{2j} \|Hf\|_{L^2k(\mathbb{T})}^{2k} \|f\|_{L^{2k}(\mathbb{T})}^{2k}, \end{aligned}$$

where we pass to the second line by applying Hölder's inequality with exponents $\frac{2k}{2k-2j}$ and $\frac{2k}{2j}$, respectively, to the j^{th} term of the sum. If we let

$$R_f = \|Hf\|_{L^{2k}(\mathbb{T})} / \|f\|_{L^{2k}(\mathbb{T})},$$

then we have

$$1 \le \sum_{j=1}^{k} \binom{2k}{2j} R_f^{-2j}.$$

If we could find f such that R_f were arbitrarily large, then we could make the right-hand side of the above arbitrarily small, obtaining a contradiction. Hence there exists a constant A_{2k} such that $R_f \le A_{2k}$; that is,

$$\|Hf\|_{L^p(\mathbb{T})} \le A_p \|f\|_{L^p(\mathbb{T})}, \qquad p = 2k.$$

We now remove some of the assumptions which we made on f in the beginning. Suppose $\hat{f}(0) \ne 0$. Note that the Hilbert transform of a constant is 0 and that $|\hat{f}(0)| \le \|f\|_{L^p(\mathbb{T})}$. Thus,

$$\begin{aligned}\|Hf\|_{L^p(\mathbb{T})} &= \|H(f - \hat{f}(0))\|_{L^p(\mathbb{T})} \\ &\le A_p \|f - \hat{f}(0)\|_{L^p(\mathbb{T})} \\ &\le 2A_p \|f\|_{L^p(\mathbb{T})}.\end{aligned}$$

Now, suppose f is a general trigonometric polynomial (not necessarily real-valued). We can write $f = P + iQ$, where P and Q are real-valued trigonometric polynomials. We then obtain

$$\begin{aligned}\|Hf\|_{L^p(\mathbb{T})} &\le \|HP\|_{L^p(\mathbb{T})} + \|HQ\|_{L^p(\mathbb{T})} \\ &\le 2A_p(\|P\|_{L^p(\mathbb{T})} + \|Q\|_{L^p(\mathbb{T})}) \\ &\le 4A_p \|f\|_{L^p(\mathbb{T})}.\end{aligned}$$

This proves the theorem for $p = 2k$.

Let $2 \le p < \infty$. We can find some $k \in \mathbb{N}$ such that $2k \le p < 2(k+1)$. The Riesz-Thorin interpolation theorem and the above result therefore yields a constant $C_p > 0$ such that

$$\|Hf\|_{L^p(\mathbb{T})} \le C_p \|f\|_{L^p(\mathbb{T})}$$

for all trigonometric functions f. Now, let $1 < p < 2$, and p' satisfy $\frac{1}{p} + \frac{1}{p'} = 1$. Since the Hilbert transform is self-adjoint, duality gives us that

$$\|Hf\|_{L^p(\mathbb{T})} \le C_{p'} \|f\|_{L^p(\mathbb{T})}.$$

This completes the proof. □

It is worth remarking that the above proof, due to Salomon Bochner, is fairly short relative to the classical proof of this result. The argument is nevertheless not by any means simple, and relies upon both interpolation and duality arguments. We will now show that the result of Theorem 3.7 cannot be improved; that is, the Hilbert transform is unbounded in operator norm on $L^1(\mathbb{T})$ and $C(\mathbb{T})$. The job of showing this will be expedited significantly by the following proposition, which we state without proof.

Proposition 3.8. *For any $f \in L^1(\mathbb{T})$, we have*

$$Hf(\theta) = \lim_{\epsilon \to 0} i \int_{|\theta - \varphi| > \epsilon} \cot(\pi(\theta - \varphi)) f(\varphi)\, d\varphi$$

almost everywhere.

Theorem 3.9. *The Hilbert transform is unbounded in operator norm on $L^1(\mathbb{T})$ and $C(\mathbb{T})$.*

Proof. To prove the theorem, we take advantage of the fact that $\mathbb{T}$ can also be written as the interval $[-\frac{1}{2}, \frac{1}{2}]$, with the endpoints $-\frac{1}{2}$ and $\frac{1}{2}$ identified. When $\mathbb{T}$ is defined this way, we have

$$Hf(\theta) = \lim_{\epsilon\to 0} i \int_{|\varphi|>\epsilon} \cot(\pi\varphi) f(\theta - \varphi)\, \mathrm{d}\varphi.$$

Clearly, we can ignore the coefficient i in the above expression. Now, since H is given by convolution with the cotangent function, which is not integrable, one would expect that the L^1 norm of the Hilbert transform of an approximation to the identity would be unbounded. Indeed, this is the case; consider $f_N = N\chi_{(-\frac{1}{N},\frac{1}{N})}$. Then we have

$$Hf_N(\theta) = \lim_{\epsilon\to 0} \int_{\substack{\theta-\frac{1}{N} \\ |\varphi|>\epsilon}}^{\theta+\frac{1}{N}} N\cot(\pi\varphi)\, \mathrm{d}\theta.$$

For any $\epsilon > 0$ we can find N_ϵ large enough that $\frac{1}{N_\epsilon} < \epsilon$, which implies that, for all $|\theta| > 2\epsilon$,

$$Hf_{N_\epsilon}(\theta) = \int_{\theta-\frac{1}{N_\epsilon}}^{\theta+\frac{1}{N_\epsilon}} N_\epsilon \cot(\pi\varphi)\, \mathrm{d}\varphi$$

for all $|\theta| > 2\epsilon$. Hence

$$\begin{aligned}
\|Hf_{N_\epsilon}\|_{L^1(\mathbb{T})} &= \int_{-\frac{1}{2}}^{\frac{1}{2}} |Hf_{N_\epsilon}(\theta)|\, \mathrm{d}\theta \\
&\geq \int_{|\theta|>2\epsilon} \left| \int_{\theta-\frac{1}{N_\epsilon}}^{\theta+\frac{1}{N_\epsilon}} N_\epsilon \cot(\pi\varphi)\, \mathrm{d}\varphi \right| \mathrm{d}\theta \\
&\geq 2\int_{\theta>2\epsilon} \int_{\theta-\frac{1}{N_\epsilon}}^{\theta+\frac{1}{N_\epsilon}} N_\epsilon \cot\left(\pi\left(\theta + \frac{1}{N_\epsilon}\right)\right) \mathrm{d}\varphi\, \mathrm{d}\theta \\
&= 4\int_{\theta>2\epsilon} \cot\left(\pi\left(\theta + \frac{1}{N_\epsilon}\right)\right) \mathrm{d}\theta.
\end{aligned}$$

The above expression tends to ∞ as $\epsilon \to 0$ (note that $N_\epsilon \to \infty$). Since $\|f_N\|_{L^1(\mathbb{T})} = 1$ for all $N \in \mathbb{N}$, it follows that

$$\|H\|_{L^1\to L^1} \geq \|Hf_{N_\epsilon}\|_{L^1(\mathbb{T})} \to \infty$$

as $\epsilon \to 0$. This suffices to show L^1 unboundedness.

For the $C(\mathbb{T})$ case, take functions $g_N \in C(\mathbb{T})$ such that $g_N(x) = 1$ for all $x > \frac{1}{N}$, $g_N(x) = 0$ for all $x \leq 0$, and $\|g_N\|_{L^\infty(\mathbb{T})} = 1$ for all $N \in \mathbb{N}$. We then have that:

$$\begin{aligned}
\|Hg_N\|_{L^\infty(\mathbb{T})} &\geq |Hg_N(0)| \\
&= \int_{\frac{1}{N}}^{\frac{1}{2}} \cot(\pi\varphi)\, \mathrm{d}\varphi
\end{aligned}$$

which clearly grows without bound as $N \to \infty$. This completes the proof. □

The failure of the Fourier series to converge in $L^1(\mathbb{T})$ and $C(\mathbb{T})$ can also be proved using only Lemma 3.3 and basic properties of the Dirichlet kernel—in particular, Proposition 2.6. The reader who is uncomfortable with our use of the unjustified Proposition 3.8 is encouraged to work out the details of the former approach. We choose to use the Hilbert transform here so as to introduce the identity given in

Proposition 3.8, which forms an important connection between our discussion in this section and the real-variable results in Section 6.

4. The Fourier Transform on the Line

We now shift our attention from the circle $\mathbb{T}$ to the whole real line $\mathbb{R}$, and hence from periodic functions to general integrable functions. Instead of representing functions in terms of a series of discrete terms, we shall in this setting endeavor to represent a function as an integral over the continuum. Our goal in this section is to quickly develop the fundamental results needed to render this representation useful to our discussion in the next section of some important results in real-variable harmonic analysis.

Definition 4.1. Let $f \in L^1(\mathbb{R})$. We define the *Fourier transform* of f by

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x.$$

The real-variable Fourier transform satisfies the following properties, analogous to those of Proposition 2.2:

Proposition 4.2. *For all $f, g \in L^1(\mathbb{R})$ and all $\xi \in \mathbb{R}$, we have*

(1) $(\widehat{f+g})(\xi) = \hat{f}(\xi) + \hat{g}(\xi)$.
(2) $\widehat{\lambda f}(\xi) = \lambda \hat{f}(\xi)$.
(3) $(\widehat{f * g})(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.
(4) $(\widehat{f(\cdot + y)})(\xi) = \hat{f}(\xi) e^{-2\pi i \xi y}$.
(5) $(\widehat{e^{2\pi i (\cdot) y} f})(\xi) = \hat{f}(\xi - y)$.
(6) $\widehat{f(a(\cdot))} = \frac{1}{a} \widehat{f(a^{-1}(\cdot))}$
(7) *If f is continuously differentiable, then* $(\widehat{f'})(\xi) = (2\pi i \xi)\hat{f}$.
(8) $\|\hat{f}\|_{L^\infty(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})}$.
(9) (Riemann-Lebesgue lemma) $\lim_{|\xi| \to \infty} |\hat{f}(\xi)| = 0$.

Again, the proofs of these results are left to the reader. On the circle, it was useful in many cases to restrict our attention to trigonometric polynomials, where many results could be easily obtained, and then extend by density. On the line, we have a space of similarly nice functions, which will prove to be a natural setting for the theorems we wish to consider.

Definition 4.3. We say that $f \in C^\infty(\mathbb{R})$ is a *Schwarz function* if

$$\sup_{x \in \mathbb{R}} |x|^m \left| f^{(n)}(x) \right| < \infty$$

for all $m, n \in \mathbb{N}$. We define the *Schwarz space* $\mathcal{S}(\mathbb{R})$ to be the space of all Schwarz functions.

Intuitively speaking, the Schwarz space is the space of smooth functions that decay with their derivatives faster than any polynomial at infinity. It is easy to see that any smooth function with compact support is a Schwarz function. Since the C_c^∞ functions are dense in $L^p(\mathbb{R})$ for all $1 \leq p < \infty$, it follows that the Schwarz functions are as well.

We now prove Fourier inversion for Schwarz functions.

Theorem 4.4 (Fourier inversion)**.** *Let* $f \in \mathcal{S}(\mathbb{R})$*. Then*

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi$$

for all $x \in \mathbb{R}$*.*

Proof. Our proof is based on the identity

$$(\widehat{e^{-\pi(\cdot)^2}})(\xi) = e^{-\pi\xi^2},$$

which can be shown using contour integration. It follows, by Proposition 4.2(5) and (6), that

$$(\widehat{e^{-\pi(\epsilon(\cdot))^2}})(\xi) = \frac{1}{\epsilon} e^{-\pi(\xi/\epsilon)^2}.$$

The dominated convergence theorem allows us to write

$$\begin{aligned}
\int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi &= \lim_{\epsilon \to 0} \int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i x \xi} e^{-\pi(\epsilon\xi)^2} \, \mathrm{d}\xi \\
&= \lim_{\epsilon \to 0} \int_{\mathbb{R}} e^{2\pi i x \xi} e^{-\pi|\epsilon\xi|^2} \int_{\mathbb{R}} f(y) e^{-2\pi i \xi y} \, \mathrm{d}y \, \mathrm{d}\xi \\
&= \lim_{\epsilon \to 0} \int_{\mathbb{R}} f(y) \int_{\mathbb{R}} e^{-\pi|\epsilon\xi|^2} e^{-2\pi i \xi(y-x)} \, \mathrm{d}\xi \, \mathrm{d}y \\
&= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{\mathbb{R}} f(y) e^{2\pi i x y} e^{-\pi((y-x)/\epsilon)^2} \, \mathrm{d}y.
\end{aligned}$$

Since $\frac{1}{\epsilon} e^{-\pi((\cdot)/\epsilon)}$ is an approximate identity as $\epsilon \to 0$, we have that

$$\int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi = f(x)$$

as desired. □

The second important result that we consider is the real-variable version of Corollary 3.2(1), the Plancherel Theorem.

Theorem 4.5 (Plancherel Theorem)**.** *For any* $f \in L^2(\mathbb{R})$ *we have that*

$$\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})}$$

Proof. By density, it suffices to show that the theorem holds for $f \in \mathcal{S}(\mathbb{R})$. For such f we have

$$\begin{aligned}
\|\hat{f}\|_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} \hat{f}(\xi) \overline{\hat{f}(\xi)} \, \mathrm{d}\xi \\
&= \int_{\mathbb{R}} f(y) \int_{\mathbb{R}} \overline{\hat{f}(\xi)} e^{-2\pi i \xi y} \, \mathrm{d}\xi \, \mathrm{d}y \\
&= \int_{\mathbb{R}} f(y) \int_{\mathbb{R}} \hat{\bar{f}}(-\xi) e^{-2\pi i \xi y} \, \mathrm{d}\xi \, \mathrm{d}y \\
&= \int_{\mathbb{R}} f(y) \overline{f(y)} \, \mathrm{d}y
\end{aligned}$$

where we use Theorem 4.4 to pass to the last line. □

5. Calderón-Zygmund Theory

Many of the most important techniques used in harmonic analysis involve "breaking up" functions into smaller pieces that are easier to study than the function as a whole. The Fourier series on $\mathbb{T}$ is one such technique, as is the Fourier transform on $\mathbb{R}$ if one allows for talk of infinitesimal "pieces." We now consider another fundamental such technique: the Calderón-Zygmund decomposition. Intuitively speaking, the following theorem allows us to break a function up into two pieces, one which is "good" (small) and one which is "bad" (large) but which latter piece we can control in a useful way. In the following theorem, the "good" part is denoted g and the "bad" part b.

Theorem 5.1 (Calderón-Zygmund decomposition)**.** *Let $f \in L^1(\mathbb{R})$ and $\lambda > 0$. Then one can write $f = g + b$ with $|g| \leq \lambda$ and*

$$b = \sum_{Q \in \mathcal{B}} \chi_Q f \tag{5.2}$$

where $\mathcal{B} = \{Q\}$ is a collection of disjoint intervals with the property that for each Q one has

$$\lambda < \frac{1}{|Q|} \int_Q |f| \leq 2\lambda. \tag{5.3}$$

Moreover,

$$\left| \bigcup_{Q \in \mathcal{B}} Q \right| < \frac{1}{\lambda} \|f\|_1. \tag{5.4}$$

Proof. For each $m \in \mathbb{Z}$, let $\mathcal{D}_m$ be defined by

$$\mathcal{D}_m = \{2^m lk, 2^m(k+1) : k \in \mathbb{Z}\}.$$

Note that, for $Q \in \mathcal{D}_m$ and $Q' \in \mathcal{D}_{m'}$ with $m \leq m'$, we have either $Q \cap Q' = \emptyset$ or $Q \subseteq Q'$. Since $f \in L^1(\mathbb{R})$, we can pick an m_0 large enough that

$$\frac{1}{|Q|} \int_Q |f| \leq \lambda \tag{5.5}$$

for all $Q \in \mathcal{D}_{m_0}$. Now, for each $Q \in \mathcal{D}_{m_0}$ consider the two elements of $\mathcal{D}_{m_0-1}$ given by splitting Q down the middle. Let Q' be one of these halves; if

$$\frac{1}{|Q'|} \int_{Q'} |f| > \lambda,$$

then we have

$$\lambda < \frac{1}{|Q'|} \int_{Q'} |f| \leq \frac{2}{|Q|} \int_Q |f| \leq 2\lambda$$

and so we include Q' in $\mathcal{B}$. Otherwise, Q' satisfies 5.5 with Q' replacing Q, and we subdivide Q' in just the same way that we subdivided Q. Continuing inductively, we produce a disjoint collection of dyadic intervals $\mathcal{B}$ that satisfies 5.3 and 5.4. Let $G = \mathbb{R} / \bigcup_{Q \in \mathcal{B}} \overline{Q}$. Then for $x \in G$ we have that x is contained in the intersection of a decreasing sequence of intervals Q satisfying 5.5. By the Lebesgue differentiation theorem, $|f(x)| \leq \lambda$ for almost every such x. Disregarding the measure-zero set $\left(\bigcup_{\mathcal{B}} Q\right)^c / G$, the functions $g := \chi_G f$ and $b := \sum_{Q \in \mathcal{B}} \chi_Q f$ satisfy the conclusion of the theorem. □

With the Calderón Zygmund decomposition in hand, we consider the real-variable equivalent of the Hilbert transform we saw in Section 4.

Definition 5.6. Let $f \in \mathcal{S}(\mathbb{R})$. The *Hilbert transform* of f is given by

$$Hf(x) = \lim_{\epsilon \to 0} \int_{|y|>\epsilon} \frac{f(x-y)}{y} \, \mathrm{d}x = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} \frac{f(y)}{x-y} \, \mathrm{d}x.$$

We call the kernel $x \mapsto \frac{1}{x}$, $x \neq 0$ the *Hilbert kernel* when speaking in the context of the Hilbert transform.

Note the similarity between this definition and the formulation of the Hilbert transform on the circle given in Proposition 3.8. Indeed, since $\frac{|x|}{2} \leq |\sin x| \leq |x|$ and $\frac{\sqrt{2}}{2} \leq |\cos x| \leq 1$ for all $|x| \leq \frac{\pi}{4}$, we have that

$$\frac{\sqrt{2}}{2|x|} \leq |\cot x| \leq \frac{2}{|x|}$$

for all such x. With this, the following proposition follows easily from our work in Theorem 3.9.

Proposition 5.7. *The real-variable Hilbert transform is unbounded in operator norm in $L^1(\mathbb{R})$ and $C(\mathbb{R})$.*

It is natural to ask if we have the same boundedness conditions for the Hilbert transform on the line as we obtained for the Hilbert transform on the circle; namely, L^p boundedness for all $1 < p < \infty$. We will tackle this question by immediately considering a generalization of the Hilbert transform.

Definition 5.8. Let $K : \mathbb{R}/\{0\} \to \mathbb{C}$ satisfy, for some constant B,

(1) $|K(x)| \leq \frac{B}{|x|}$ for all $x \neq 0$,
(2) $\int_{|x|>2|y|} |K(x) - K(x-y)| \, \mathrm{d}x \leq B$ for all $y \neq 0$,
(3) $\int_{r<|x|<s} K(x) \, \mathrm{d}x = 0$ for all $0 < r < s < \infty$.

We call K satisfying the above a *Calderón-Zygmund kernel*. For a Calderón-Zygmund kernel K we define a *Calderón-Zygmund operator* with kernel K by

$$Tf(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} K(x-y) f(y) \, \mathrm{d}y = \lim_{\epsilon \to 0} \int_{|y|>0} K(y) f(x-y) \, \mathrm{d}y.$$

It is easy to verify that the Hilbert kernel is a Calderón-Zygmund kernel. Hence, all results that we obtain for Calderón-Zygmund operators will also apply to the Hilbert transform.

Our strategy for showing L^p boundedness of the Calderón-Zygmund operators will be to show L^2 boundedness and weak-L^1 boundedness; interpolation and duality will then suffice to prove L^p boundedness for all $1 < p < \infty$. We begin with L^2 boundedness:

Theorem 5.9. *Let K be as in Definition 5.8 and let T be the Calderón-Zygmund operator with kernel K. Then there exists a constant $C > 0$ such that, for all $f \in \mathcal{S}(\mathbb{R})$, we have $\|T\|_{L^2(\mathbb{R})} \leq CB\|f\|_{L^2(\mathbb{R})}$.*

Proof. In the proof below, let $C > 0$ be a constant that may vary from line to line. Let $0 < r < s < \infty$ and define

$$T_{r,s} f(x) = \int_{r<|y|<s} K(y) f(x-y) \, \mathrm{d}y,$$

$$m_{r,s}(\xi) = \int_{r<|x|<s} K(x)e^{-2\pi i x\xi}\,\mathrm{d}x.$$

Since $m_{r,s}$ is the Fourier transform of the kernel $K(y)\chi_{r<|y|<s}$, it is easy to see (by Plancherel's theorem) that

$$\|T_{r,s}\|_{L^2\to L^2} = \|m_{r,s}\|_{L^\infty(\mathbb{R})}$$

Suppose that $m_{r,s}$ satisfies

$$\sup_{r,s} \|m_{r,s}\|_{L^\infty(\mathbb{R})} \leq CB. \tag{5.10}$$

Since

$$Tf(x) = \lim_{r\to 0, s\to\infty} T_{r,s}f(x)$$

for all $x \in \mathbb{R}/\{0\}$, Fatou's lemma implies that

$$\|Tf\|_{L^2(\mathbb{R})} \leq CB\|f\|_{L^2(\mathbb{T})}$$

for all $f \in \mathcal{S}(\mathbb{R})$, as desired.

We now prove 5.10. To do so, we write

$$m_{r,s}(\xi) = \int_{r<|x|\leq|\xi|^{-1}} K(x)e^{-2\pi i x\xi}\,\mathrm{d}x + \int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i x\xi}\,\mathrm{d}x. \tag{5.11}$$

Consider first the first integral in 5.11. We have

$$\begin{aligned}
\left|\int_{r<|x|\leq|\xi|^{-1}} K(x)e^{-2\pi i x\xi}\,\mathrm{d}x\right| &= \left|\int_{r<|x|\leq|\xi|^{-1}} K(x)\left(e^{-2\pi i x\xi} - 1\right)\,\mathrm{d}x\right| \\
&\leq \int_{|x|\leq|\xi|^{-1}} 2\pi|K(x)||x||\xi|\,\mathrm{d}x \\
&\leq 2\pi \int_{|x|\leq|\xi|^{-1}} B\,\mathrm{d}x \\
&= CB
\end{aligned}$$

for some $C > 0$. Note that we use condition (3) of Definition 5.8 in the first line and condition (1) in the third, while we use the inequality $|1 - e^{-z}| < |z|$ to pass to the second line. Now we consider the second integral in 5.11. Observe that

$$\begin{aligned}
\int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i x\xi}\,\mathrm{d}x &= -\int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i(x+\frac{1}{2\xi})\xi}\,\mathrm{d}x \\
&= -\int_{|\xi|^{-1}<|x-\frac{1}{2\xi}|<s} K\left(x - \frac{1}{2\xi}\right)e^{-2\pi i x\xi}\,\mathrm{d}x.
\end{aligned}$$

Hence

$$2\int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i x\xi}\,\mathrm{d}x = \int_{|\xi|^{-1}<|x|<s}\left(K(x) - K\left(x - \frac{1}{2\xi}\right)\right)e^{-2\pi i x\xi}\,\mathrm{d}x + R$$

where

$$R = \int_{|\xi|^{-1}<|x|<s} K\left(x - \frac{1}{2\xi}\right)e^{-2\pi i x\xi}\,\mathrm{d}x - \int_{|\xi|^{-1}<|x-\frac{1}{2\xi}|<s} K\left(x - \frac{1}{2\xi}\right)e^{-2\pi i x\xi}\,\mathrm{d}x.$$

Now, we have

$$\left|\int_{|\xi|^{-1}<|x|<s}\left(K(x)-K\left(x-\frac{1}{2\xi}\right)\right)e^{-2\pi ix\xi}\,\mathrm{d}x\right| \leq \int_{|\xi|^{-1}<|x|}\left|K(x)-K\left(x-\frac{1}{2\xi}\right)\right|\,\mathrm{d}x$$
$$\leq B$$

by condition (2) of Definition 5.8. It remains only to bound $|R|$. It is clear from the definition that R is given by the integral of $K(x-1/(2\xi))e^{-2\pi ix\xi}$ over some region A of measure no greater than $1/|\xi|$. Moreover, for all x in this region we have $|x| \geq \frac{1}{2|\xi|}$. Hence

$$\begin{aligned}|R| &\leq \int_A \left|K\left(x-\frac{1}{2\xi}\right)\right|\,\mathrm{d}x \\ &\leq CB\int_A |\xi|\,\mathrm{d}x \\ &\leq CB.\end{aligned}$$

This completes the proof. □

We now turn to proving the weak-L^1 boundedness of the Calderón-Zygmund operators. Here we rely crucially upon the Calderón-Zygmund decomposition.

Theorem 5.12. *Let K be as in Definition 5.8 and let T be the Calderón-Zygmund operator with kernel K. Then there exists a constant $C > 0$ such that, for all $f \in \mathcal{S}(\mathbb{R})$, we have $\|T\|_{L^{1\infty}(\mathbb{R})} \leq CB\|f\|_{L^1(\mathbb{R})}$.*

Proof. As in the proof of Theorem 5.9, we allow $C > 0$ to be a constant that varies from line to line. We may assume without loss of generality that $B = 1$ (otherwise consider $K' = K/B$). Now let $f \in \mathcal{S}(\mathbb{R})$ and $\lambda > 0$ be arbitrary. Let $f = b + g$ be the Calderón-Zygmund decomposition of f with respect to λ and $\mathcal{B} = \{Q\}$ be the relevant collection of cubes. We wish to modify g and b in such a way that the average value of b on any cube $Q \in \mathcal{B}$ is 0. In particular, we consider the functions

$$g' = g + \sum_{Q\in\mathcal{B}} \frac{\chi_Q}{|Q|}\int_Q f,$$

$$b' = b - \sum_{Q\in\mathcal{B}} \frac{\chi_Q}{|Q|}\int_Q f = \sum_{Q\in\mathcal{B}} f_Q$$

with

$$f_Q = \chi_Q\left(f - \frac{1}{|Q|}\int_Q f\right).$$

Our functions g' and b' satisfy the following:

$$f = g' + b', \qquad [\|g'\|_{L^\infty(\mathbb{R})} \leq 2\lambda,$$
$$\|b'\|_{L^1(\mathbb{R})} \leq 2\|f\|_{L^1(\mathbb{R})}, \qquad [\|g'\|_{L^1(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})},$$
$$\int_Q b' = 0.$$

Now, we have:

$$\begin{aligned}\|Tf\|_{L^{1,\infty}(\mathbb{R})} &= \left|\{x\in\mathbb{R} : |Tf(x)| > \lambda\}\right| \\ &\leq \left|\left\{x\in\mathbb{R} : |Tg'(x)| > \frac{\lambda}{2}\right\}\right| + \left|\left\{x\in\mathbb{R} : |Tb'(x)| > \frac{\lambda}{2}\right\}\right|. \qquad (5.13)\end{aligned}$$

As a bound for the first term in 5.13, we have:

$$\begin{aligned}\left|\{x \in \mathbb{R} : |Tg'(x)| > \frac{\lambda}{2}\}\right| &= \int_{|Tg'(x)|>\frac{\lambda}{2}} \mathrm{d}x \\ &\leq \frac{C}{\lambda^2}\int_{\mathbb{R}} |Tg'(x)|^2 \,\mathrm{d}x \\ &= \frac{C}{\lambda^2}\|Tg'\|^2_{L^2(\mathbb{R})} \\ &\leq \frac{C}{\lambda^2}\|g'\|^2_{L^2(\mathbb{R})} \\ &\leq \frac{C}{\lambda^2}\|g'\|_{L^\infty(\mathbb{R})}\|g'\|_{L^1(\mathbb{R})} \\ &\leq \frac{C}{\lambda}\|f\|_{L^1(\mathbb{R})}.\end{aligned}$$

Note that we used Theorem 5.9 to pass to the fourth line above.

To bound the second term in 5.13, we first define, for each $Q \in \mathcal{B}$, a cube Q^* with the same center y_Q as Q but with side length dilated by some factor D, to be determined later. Then

$$\begin{aligned}\left|\left\{x \in \mathbb{R} : Tb'(x) > \frac{\lambda}{2}\right\}\right| &\leq \left|\bigcup_{Q\in\mathcal{B}} Q^*\right| + \left|\left\{x \in \mathbb{R}/\cup Q^* : |Tb'(x)| > \frac{\lambda}{2}\right\}\right| \\ &\leq D\sum_{Q\in\mathcal{B}}|Q| + \frac{2}{\lambda}\int_{\mathbb{R}/\cup Q^*}|Tg'(x)|\,\mathrm{d}x \\ &\leq \frac{C}{\lambda}\|f\|_{L^1(\mathbb{R})} + \frac{C}{\lambda}\sum_{Q\in\mathcal{B}}\int_{\mathbb{R}/Q^*}|Tf_Q(x)|\,\mathrm{d}x.\end{aligned}$$

Because b' has mean value 0, we can write

$$\begin{aligned}Tf_Q(x) &= \int_{\mathbb{Q}} K(x-y)f_Q(x)\,\mathrm{d}y \\ &= \int_{\mathbb{Q}} \left(K(x-y) - K(x-y_Q)\right) f_Q(x)\,\mathrm{d}y\end{aligned}$$

We can choose D large enough that $|x - y_Q| > 2|y - y_Q|$ for all $x \in \mathbb{R} \setminus Q^*$ and all $y \in Q$. For such D and any $Q \in \mathcal{B}$, we obtain

$$\begin{aligned}\int_{\mathbb{R}\setminus Q^*}|Tf_Q(x)|\,\mathrm{d}x &\leq \int_{\mathbb{R}\setminus Q^*}\int_Q |K(x-y) - K(x-y_Q)||f_Q(y)|\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_Q |f_Q(y)| \int_{\mathbb{R}/Q^*}|K(x-y) - K(x-y_Q)|\,\mathrm{d}x\,\mathrm{d}y \\ &\leq \int_Q |f_Q(y)|\,\mathrm{d}y \\ &\leq 2\int_Q |f(y)|\,\mathrm{d}y.\end{aligned}$$

To pass to the third line, we use condition (2) of Definition 5.8. Hence

$$\frac{2}{\lambda}\sum_{Q\in\mathcal{B}}\int_{\mathbb{R}\setminus Q^*}|Tf_Q(x)|\,\mathrm{d}x \leq \frac{C}{\lambda}\|f\|_{L^1(\mathbb{R})}.$$

This completes the proof. □

A few remarks are in order, here. Note that, aside from the invocation of L^2-boundedness, we only used condition (2) of Definition 5.8 in the above proof. It follows that both the above theorem and Theorem 5.15, below, apply to all L^2-bounded linear operators given by a kernel which satisfies condition (2) of Definition 5.8. Because of its importance, this condition is called the *Hörmander condition.* Although it might appear awkward or unintuitive on its own, the following proposition gives a simple necessary condition for the Hörmander condition to hold.

Proposition 5.14. *Let $K : \mathbb{R} \setminus \{0\} \to \mathbb{C}$ satisfy $|\frac{\mathrm{d}}{\mathrm{d}x} K(x)| \leq C|x|^{-2}$ for all $x \neq 0$ and some $C > 0$. Then K satisfies the Hörmander condition.*

Proof. This follows immediately from the mean value theorem. □

We are now within easy reach of the result we set out to prove.

Theorem 5.15. *Let T be the Calderón-Zygmund operator given by a kernel K. Then for every $1 < p < \infty$ one can extend T to a bounded operator on $L^p(\mathbb{R})$ with the bound $\|T\|_{L^p \to L^p} \leq CB$ with B as in Definition 5.8 and for some $C > 0$.*

Proof. By Theorems 5.9 and 5.12 and the density of the Schwarz functions in $L^p(\mathbb{R})$, we have $\|T\|_{L^2 \to L^2} \leq CB$ and $\|T\|_{L^1 \to L^{1,\infty}} \leq CB$. By the Marcinkiewicz interpolation theorem, we obtain the result for $1 < p \leq 2$. Now, note that the adjoint of T satisfies

$$T^* f(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} \overline{K(-x)} g(y)\, \mathrm{d}y.$$

Since $\overline{K(-x)}$ is a Calderón-Zygmund kernel, by duality, we obtain the result for $2 < p < \infty$. □

Note that the result above cannot be improved; indeed, we already have a familiar example of a Calderón-Zygmund operator that fails the above theorem for $p = 1, \infty$; namely, the Hilbert transform.

To prove Theorem 5.9, we "cut up" the function on which the Calderón-Zygmund operator acts by taking its Fourier transform. An alternative method (used elsewhere in harmonic analysis) is to "cut up" the operator itself. To develop this method we turn to the notion of almost orthogonality.

Definition 5.16. Let H be a Hilbert space and let $\{T_j\}_{j \in \mathbb{N}}$ be a family of linear operators on H. We say that the family $\{T_j\}_{j \in \mathbb{N}}$ is *almost orthogonal* if there exists a nonnegative function $\lambda : \mathbb{Z} \to \mathbb{R}^+$ such that

$$\sum_{k \in \mathbb{Z}} \lambda(k) < \infty$$

which satisfies

$$\|T_j T_k^*\| \leq \lambda^2(j-k), \qquad \|T_j^* T_k\| \leq \lambda^2(j-k)$$

for all $j, k \in \mathbb{N}$.

Note that this is a generalization of the notion of orthogonality; a family of strictly orthogonal operators $\{T_j\}_{j \in \mathbb{N}}$ would satisfy the above with $\lambda = 0$. We can now prove the following theorem, which is the crucial tool that we wish to introduce here.

Theorem 5.17 (Cotlar's lemma). *Let H be a Hilbert space and $\{T_j\}_I$ be a finite family of almost orthogonal linear operators on H, with λ as in Definition 5.16. If we let*

$$M = \sum_{k\in\mathbb{Z}} \lambda(k)$$

then

$$\left\| \sum_{j\in I} T_j \right\| \leq M.$$

Proof. Let $n \in \mathbb{N}$. Since T^*T is a self-adjoint linear operator, if it is bounded, then the spectral theorem gives us that

$$\|(T^*T)^n\| = \|T^*T\|^n = \|T\|^{2n}.$$

We therefore endeavor to bound $\|(T^*T)^n\|$. We write

$$(T^*T)^n = \sum_{\substack{j_1,\dots,j_n\in I\\ k_1,\dots,k_n\in I}} \prod_{i=1}^{n} T_{j_i}^* T_{k_i}.$$

Observe that

$$\left\| \prod_{i=1}^{n} T_{j_i}^* T_{k_i} \right\| \leq \prod_{i=1}^{n} \|T_{j_i}^*\| \|T_{k_i}\|,$$

$$\left\| \prod_{i=1}^{n} T_{j_i}^* T_{k_i} \right\| \leq \|T_{j_1}^*\| \|T_{k_n}\| \prod_{i=1}^{n-1} \|T_{j_{i+1}}^* T_{k_i}\|.$$

Taking the geometric mean of the two bounds above yields

$$\left\| \prod_{i=1}^{n} T_{j_i}^* T_{k_i} \right\| \leq \sqrt{\|T_{j_1}^* T_{k_n}\|} \prod_{i=1}^{n} \sqrt{\|T_{j_i}^* T_{k_i}\|} \prod_{i=1}^{n-1} \sqrt{\|T_{j_{i+1}}^* T_{k_i}\|}$$

We can now write

$$\begin{aligned}
\|(T^*T)^n\| &\leq \sum_{\substack{j_1,\dots,j_n\in I\\ k_1,\dots,k_n\in I}} \sqrt{\|T_{j_1}^* T_{k_n}\|} \prod_{i=1}^{n} \sqrt{\|T_{j_i}^* T_{k_i}\|} \prod_{i=1}^{n-1} \sqrt{\|T_{j_{i+1}}^* T_{k_i}\|} \\
&\leq \sum_{\substack{j_1,\dots,j_n\in I\\ k_1,\dots,k_n\in I}} B \prod_{i,\ell=1,\dots,n} \lambda(j_i - k_\ell) \\
&\leq NBA^{2n-1}
\end{aligned}$$

for $B = \sup_{j\in I} \|T_j\| \leq A < \infty$. We now have

$$\|T\| \leq (NBA^{2n-1})^{1/(2n)}$$

for all $n \in \mathbb{N}$. Letting n tend to infinity yields the desired result. □

The applications of Cotlar's lemma are many and varied. We will explore only one such application here: namely, the promised proof of the L^2-boundedness of the Calderón-Zygmund operators. Actually, we will prove this for the smaller class of operators whose kernels additionally satisfy the hypothesis of Proposition 5.14. While this is a weakening of Theorem 5.9, many of the most important Calderón-Zygmund kernels do indeed satisfy the hypothesis of Proposition 5.14, including the Hilbert kernel.

Theorem 5.18. *Let K be a Calderón-Zygmund kernel with the additional assumption that $\left|\frac{\mathrm{d}}{\mathrm{d}x}K(x)\right| \leq B|x|^{-1}$ with B as in Definition 5.8, and let T be the associated Calderón-Zygmund operator. Then*

$$\|T\|_{L^2\to L^2} \leq CB$$

for some constant $C > 0$.

To prove Theorem 5.18, we will need the following *partition of unity*:

Proposition 5.19. *There exists a nonnegative function $\psi \in C^\infty(\mathbb{R})$ with compact support and that vanishes at 0 such that*

$$\sum_{k\in\mathbb{Z}} \psi(2^{-k}x) = 1$$

for all $x \neq 0$. Moreover, at most two terms of the above sum are nonzero.

Proof. Let $\phi \in C^\infty(\mathbb{R}\setminus\{0\})$ satisfy $\phi(x) = 1$ for all $|x| \leq 1$ and $\phi(x) = 0$ for all $|x| \geq 2$. The function $\psi : \psi(x) = \phi(x) - \phi(2x)$ satisfies the above proposition. □

We now prove Theorem 5.18.

Proof of Theorem 5.18. As usual, let $C > 0$ be a constant that may vary from line to line. As in the proof of Theorem 5.12, we may assume $B = 1$. Let ψ be as in Proposition 5.19 and define, for $j \in \mathbb{Z}$

$$K_j(x) = K(x)\psi(2^{-j}x),$$

$$T_jf(x) = \int_{\mathbb{R}} K_j(x-y)f(y)\,\mathrm{d}y.$$

We note that, due to the smoothness of ψ, for all $j \in \mathbb{Z}$ we have

$$\int_{\mathbb{R}} K_j(x)\,\mathrm{d}x = 0, \tag{5.20}$$

$$\int_{\mathbb{R}} |K_j(x)|\,\mathrm{d}x \leq C.$$

so that T_j is absolutely convergent for any compactly supported $f \in L^1(\mathbb{R})$. Since such functions are dense in $L^2(\mathbb{R})$, it is no problem to restrict our attention to these. Since ψ is a partition of unity, we have that $\sum_{|j|\leq N} T_j \to T$ pointwise. We note the following easily verified bounds:

$$\left\|\frac{\mathrm{d}}{\mathrm{d}x}K_j\right\|_{L^\infty(\mathbb{R})} \leq C2^{-2j},$$

$$\int_{\mathbb{R}} |x||K_j(x)|\,\mathrm{d}x \leq C2^{-j}.$$

If $\widetilde{K}_j(x) = \overline{K}_j(-x)$, then, for $j,k \in \mathbb{Z}$

$$(T_j^*T_k)(x) = \int_{\mathbb{R}} \left(\widetilde{K}_j * K_k\right)(y)f(x-y)\,\mathrm{d}y,$$

Young's inequality therefore yields

$$\|T_j^*T_k\|_{L^2\to L^2} \leq \left\|\widetilde{K}_j * K_k\right\|_{L^1(\mathbb{R})}.$$

Likewise, we have that

$$\|T_j T_k^*\|_{L^2 \to L^2} \leq \left\| K_j * \widetilde{K}_k \right\|_{L^1(\mathbb{R})}.$$

Assume without loss of generality that $j \geq k$. Then, by 5.20, we have

$$\begin{aligned} \left| \left(\widetilde{K}_j * K_k \right)(x) \right| &= \left| \int_{\mathbb{R}} \overline{K}_j(y-x) K_k(y) \, \mathrm{d}y \right| \\ &= \left| \int_{\mathbb{R}} \left(\overline{K}_j(y-x) - \overline{K}_j(-x) \right) K_k(y) \, \mathrm{d}y \right| \\ &\leq \int_{\mathbb{R}} \left\| \frac{\mathrm{d}}{\mathrm{d}x} K_j \right\|_{L^\infty(\mathbb{R})} |y| |K_k(y)| \, \mathrm{d}y \\ &\leq C 2^{k-2j}. \end{aligned}$$

Since $\widetilde{K}_j * K_k$ is supported in $|x| \leq C2^j$, we have that

$$\left\| \widetilde{K}_j * K_k \right\|_{L^1(\mathbb{R})} \leq C 2^{-|j-k|}.$$

Applying Cotlar's lemma with $\lambda(k) = C 2^{-\frac{1}{2}|k|}$ yields

$$\left\| \sum_{|k| \leq N} T_k \right\|_{L^2 \to L^2} \leq C$$

for a constant $C > 0$ not dependent on N. Fatou's lemma therefore yields the desired result. ☐

Acknowledgments. I wish to thank my mentor, Casey Rodriguez, for his guidance in the writing of this paper and his frequent insights into the doing of harmonic analysis. I also wish to thank Peter May and the Department of Mathematics at the University of Chicago for making this REU possible.

References

[1] Loukas Grafakos. Classical Fourier Analysis. Second Edition. Springer. 2008.
[2] Steven Krantz. A Panorama of Harmonic Analysis. The Mathematical Association of America. 1999.
[3] Camil Muscalu and Wilhelm Schlag. Classical and Multilinear Harmonic Analysis, Volume 1. Cambridge University Press. 2013.
[4] Elias Stein and Guido Weiss. Introduction to Fourier Analysis on Euclidean Spaces. Princeton University Press. 1971.