Harmonic Analysis: From Fourier to Wavelets

María Cristina Pereyra

Lesley A. Ward

Department of Mathematics and Statistics, MSC03 2150, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA

E-mail address: `crisp@math.unm.edu`

School of Mathematics and Statistics, University of South Australia, Mawson Lakes SA 5095, Australia

E-mail address: `Lesley.Ward@unisa.edu.au`

2000 *Mathematics Subject Classification.* Primary 42-01, Secondary 42-02, 42A16, 42A20, 42A24, 42A38, 42A65, 42B20, 42B25, 42C40

Contents

Chapter 11

The Hilbert transform

In this chapter we introduce a very important operator in harmonic analysis, the *Hilbert transform*[1]. It is a prototypical example for two general classes of operators: *Fourier multipliers* and *singular integral operators*. These operators arise in many areas of mathematics, including Fourier analysis, complex analysis, partial differential equations, and stochastic processes.

We define the Hilbert transform in two equivalent ways: first via its action in the frequency domain (Section 11.1), and second via its action in the time domain (Section 11.2). We are interested in the mapping properties of the Hilbert transform, and in particular what it does to functions in $L^p(\mathbb{R})$ (Section 11.3; see Section 8.6 for a parallel discussion of the Fourier transform's mapping properties). The upshot is that the Hilbert transform maps $L^p(\mathbb{R})$ into itself when $1 < p < \infty$, but not when $p = 1$ or $p = \infty$. However, weaker results hold for $L^1(\mathbb{R})$ and $L^\infty(\mathbb{R})$. These results are expressed in terms of strong and weak boundedness of operators and the maximal function (Section 11.4), as well as interpolation techniques (Section 11.5).

The interpolation results lead to short proofs of some of the most beautiful and useful inequalities in analysis (Section 11.6). We have used these inequalities throughout the book: the Hausdorff–Young inequality for the Fourier transform, the general Hölder inequality,

[1]Named after the German mathematician David Hilbert (1862–1943).

and Young's inequality for convolutions. We also take the opportunity to prove some other inequalities mentioned earlier: Hölder's inequality, Minkowski's inequality (the triangle inequality for L^p), and Minkowski's integral inequality.

We close the chapter by outlining the connections of the Hilbert transform with Fourier theory and complex analysis (Section 11.7), returning in particular to our book's theme of Fourier analysis.

Excellent accounts of the Hilbert transform and its generalizations can be found in, for example, [**SW2**, Chapter 6], [**Graf2**, Chapter 4], [**Sad**, Chapter 6], [**Duo**, Chapter 3], and [**Tor**, Chapter V].

11.1. The Hilbert transform in the frequency domain: Fourier multipliers

Definition 11.1. The *Hilbert transform* H is defined on the Fourier side by the formula

$$(Hf)^\wedge(\xi) := -i \,\operatorname{sgn}(\xi)\widehat{f}(\xi), \tag{11.1}$$

for f in $L^2(\mathbb{R})$. Here the *signum function* sgn is defined by $\operatorname{sgn}(\xi) := 1$ if $\xi > 0$; $\operatorname{sgn}(\xi) := -1$ if $\xi < 0$; and $\operatorname{sgn}(0) := 0$. $\Diamond$

Since the Fourier transform is a bijection on $L^2(\mathbb{R})$, formula (11.1) does define an operator on $L^2(\mathbb{R})$: the Hilbert transform of f is the function whose Fourier transform is $-i\operatorname{sgn}(\xi)$ times the Fourier transform of f. Moreover, the Hilbert transform is an *isometry* on $L^2(\mathbb{R})$, since it preserves L^2-norms, as we can see by using Plancherel's Identity twice:

$$\|Hf\|_2 = \|(Hf)^\wedge\|_2 = \| -i \,\operatorname{sgn}(\cdot)\widehat{f}(\cdot)\|_2 = \|\widehat{f}\|_2 = \|f\|_2.$$

Definition 11.1 also implies that $H^2 = -I$, where $-I$ denotes the negative of the identity operator, since for all $\xi \neq 0$ we have

$$(H^2 f)^\wedge(\xi) = -i \,\operatorname{sgn}(\xi)(Hf)^\wedge(\xi) = (-i \,\operatorname{sgn}(\xi))^2\widehat{f}(\xi) = -\widehat{f}(\xi).$$

Exercise 11.2. ($\mathcal{F}^4$ *is the Identity Operator*) Let $\mathcal{F}$ denote the Fourier transform: $\mathcal{F}f = \widehat{f}$. Show that $\mathcal{F}^4$ is the identity operator on those spaces where the Fourier transform is a bijection, namely $\mathcal{S}(\mathbb{R})$, $L^2(\mathbb{R})$, and $\mathcal{S}'(\mathbb{R})$. $\Diamond$

Exercise 11.3. (*Commutativity, Take 1*) Show that the Hilbert transform commutes with translations and dilations, and anticommutes with reflections, as follows.

(i) Let τ_h denote the translation operator: $\tau_h f(x) := f(x-h)$ for $h \in \mathbb{R}$. Then $\tau_h H = H\tau_h$.

(ii) Let δ_a denote the dilation operator: $\delta_a f(x) := f(ax)$ for $a > 0$. Then $\delta_a H = H\delta_a$.

(iii) Let $\tilde{f}(x) := f(-x)$ denote the reflection across the origin. Then $\widetilde{Hf}(x) = -H(\tilde{f})$. ◊

The Hilbert transform is an example of a *Fourier multiplier*.

Definition 11.4. A *Fourier multiplier* is an operator T that is defined on the Fourier side, at least for nice, smooth functions f, by multiplication of the Fourier transform $\widehat{f}$ of f by a function $m(\xi)$:

$$(Tf)^\wedge(\xi) := m(\xi)\widehat{f}(\xi). \tag{11.2}$$

The function $m(\xi)$ is called the *symbol* of the operator. ◊

The symbol m_H of the Hilbert transform is the bounded function

$$m_H(\xi) = -i \operatorname{sgn}(\xi).$$

Exercise 11.5. Verify that the partial Fourier integral operator $S_R f$, defined in Exercise 7.27, and the Cesàro means $\sigma_R f$ on $\mathbb{R}$, defined in equation (8.3), are Fourier multipliers. ◊

Exercise 11.6. Is the Fourier transform a Fourier multiplier? ◊

The results of Exercise 11.5 are no accident. Both operators, S_R and σ_R, are convolution operators: $Tf = K * f$. The time–frequency dictionary tells us that on the Fourier side $(Tf)^\wedge(\xi) = \widehat{K}(\xi)\,\widehat{f}(\xi)$. So as long as the function K is nice enough the convolution operator will be a Fourier multiplier with multiplier $m(\xi) := \widehat{K}(\xi)$. Exercise 11.3(i) can be generalized as follows.

Exercise 11.7. (*Commutativity, Take 2*) Let T_m be a Fourier multiplier with bounded symbol m. Show that T_m commutes with translations. For the converse, see Remark 11.12. ◊

We now develop the direct definition in the time domain of the Hilbert transform, given in formula (11.3) below.

11.2. The Hilbert transform in the time domain: singular integrals

Using the set-up and notation we have just developed, and the properties of the Fourier transform, we see that the Hilbert transform of a function f in $L^2(\mathbb{R})$ is given by the inverse Fourier transform of its Fourier transform, namely,

$$Hf = (m_H\widehat{f}\,)^{\vee},$$

Our experience with the time–frequency dictionary suggests that in the time domain, at least formally, the Hilbert transform should be given by convolution with a kernel k_H:

$$Hf = (m_H\widehat{f}\,)^{\vee} = k_H * f,$$

where $k_H(x) = (m_H)^{\vee}(x)$. Here we have used the fact that the Fourier transform converts convolution into multiplication. The question is: what is the inverse Fourier transform of the symbol? For a bounded symbol such as m_H, the inverse Fourier transform can be calculated in the sense of distributions, and in fact k_H turns out to be a distribution and not a function. This heuristic works in general: all bounded Fourier multipliers correspond to convolution operators with distributional kernels [**Graf2**, Chapter 2.5].

One can compute the inverse Fourier transform of $(Hf)^{\wedge}$ explicitly, at least when f is a Schwartz function. It is given formally by the following principal value integral:

$$\begin{aligned} Hf(x) \quad &:= \quad \text{p.v.}\frac{1}{\pi}\int \frac{f(y)}{x-y}\,dy \\ &= \quad \lim_{\varepsilon\to 0}\frac{1}{\pi}\int_{|x-y|>\varepsilon}\frac{f(y)}{x-y}\,dy \;=\; k_H * f(x), \end{aligned}$$

where

$$k_H(x) := \text{p.v.}\frac{1}{\pi x}$$

is the principal value distribution introduced in Section 8.7. See Example 8.52, and Exercises 8.54, 8.55, and 8.56.

We use formula (11.3) as the *definition* of the Hilbert transform of functions f in $L^2(\mathbb{R})$, not just of Schwartz functions.

We justify this broader definition as follows. Consider the well defined and nonsingular kernel

$$k_{\varepsilon,R}(y) := \frac{1}{\pi y}\chi_{\{y\in\mathbb{R}:\varepsilon<|y|<R\}}(y),$$

which lies in $L^1(\mathbb{R})\cap L^2(\mathbb{R})$ for $0<\varepsilon<R<\infty$. This function $k_{\varepsilon,R}(y)$ is a *truncated* or *cut-off* version of the kernel k_H; the values of k_H have been set to zero for $|y|\leq\varepsilon$ and for $|y|\geq R$. Note that $(K_{\varepsilon,R})^\wedge(0)=0$, since $K_{\varepsilon,R}$ is an odd function. We calculate the Fourier transform of $k_{\varepsilon,R}$, for $\xi\neq 0$:

$$\begin{aligned}
(k_{\varepsilon,R})^\wedge(\xi) &= \int_{\varepsilon<|y|<R}\frac{1}{\pi y}\,e^{-2\pi iy\xi}\,dy\\
&= \int_{\varepsilon}^{R}\frac{1}{\pi y}\,e^{-2\pi iy\xi}\,dy+\int_{-R}^{-\varepsilon}\frac{1}{\pi y}\,e^{-2\pi iy\xi}\,dy\\
&= \int_{\varepsilon}^{R}\frac{1}{\pi y}\,e^{-2\pi iy\xi}\,dy-\int_{\varepsilon}^{R}\frac{1}{\pi y}\,e^{2\pi iy\xi}\,dy\\
&= -2i\int_{\varepsilon}^{R}\frac{\sin 2\pi y\xi}{\pi y}\,dy\\
&= -\frac{2i}{\pi}\int_{2\pi\varepsilon\xi}^{2\pi R\xi}\frac{\sin t}{t}\,dt\\
&= -i\,\operatorname{sgn}(\xi)\left[\frac{2}{\pi}\int_{2\pi\varepsilon|\xi|}^{2\pi R|\xi|}\frac{\sin t}{t}\,dt\right].
\end{aligned}$$

Notice that for each real $\xi\neq 0$,

$$\frac{2}{\pi}\int_{2\pi\varepsilon|\xi|}^{2\pi R|\xi|}\frac{\sin t}{t}\,dt\to\frac{2}{\pi}\int_{0}^{\infty}\frac{\sin t}{t}\,dt=1$$

as $\varepsilon\to 0$ and $R\to\infty$. Therefore the truncated Hilbert transforms $H_{\varepsilon,R}$ on $L^2(\mathbb{R})$, given by the equivalent formulas

$$H_{\varepsilon,R}f=k_{\varepsilon,R}*f,\qquad (H_{\varepsilon,R}f)^\wedge(\xi)=(k_{\varepsilon,R})^\wedge(\xi)\widehat{f}(\xi),$$

can be seen to be uniformly bounded (in R and ε) for all f in $L^2(\mathbb{R})$, by Plancherel's Identity. Furthermore, in the limit as $R\to\infty$ and $\varepsilon\to 0$, the symbols $(k_{\varepsilon,R})^\wedge(\xi)$ converge pointwise to the symbol $m_H(\xi)=-i\,\operatorname{sgn}(\xi)$ of the Hilbert transform. Also, in $L^2(\mathbb{R})$, $H_{\varepsilon,R}f\to Hf$, where H is given by formula (11.3). By the continuity of the Fourier transform, it follows that

$$(H_{\varepsilon,R}f)^{\wedge} \to m_H\widehat{f}, \quad \text{and so} \quad (Hf)^{\wedge} = m_H\widehat{f}.$$

Thus the two definitions of the Hilbert transform of an L^2 function, via its actions in the frequency domain and in the time domain, agree.

Recall that by definition $\int_{\mathbb{R}} := \lim_{R\to\infty} \int_{R}^{-R}$, so the symmetry at infinity is built into the definition of the improper integral. The symmetry around zero in the truncation provides the correct cancellation necessary to carry out the calculation.

See Exercise 11.84 for a different proof of formula (11.3), using complex analysis.

Exercise 11.8. Show that the Hilbert transform of the characteristic function of the interval $I = [a, b]$ is given by

$$H\chi_I(x) = \frac{1}{\pi}\log\left|\frac{x-a}{x-b}\right|. \qquad \Diamond$$

Figure 11.2 on Page 320 shows the graph of $H\chi_I$ for $I = [1, 3]$.

11.3. Boundedness properties of the Hilbert transform

In this section we discuss the boundedness properties of the Hilbert transform on $L^p(\mathbb{R})$, for $1 < p < \infty$. The Hilbert transform is, by Definition 11.1, an isometry on $L^2(\mathbb{R})$. There are several alternative proofs of a less powerful statement: that there exists a positive constant C such that for all $f \in L^2(\mathbb{R})$,

$$\|Hf\|_2 \le C\|f\|_2.$$

This inequality means that the operator H is *bounded* on $L^2(\mathbb{R})$.

We explore in Chapter 12 alternative proofs of the boundedness on $L^2(\mathbb{R})$ of the Hilbert transform and its analogues. Project 12.11, introduces the discrete and finite analogues of the Hilbert transform and examines their boundedness, via the very useful *Cotlar's Lemma*[2].

[2]Named after the Ukranian-born, Argentinian/Venezuelan by adoption, mathematician Mischa Cotlar (1911–2007).

Project 12.16 explores how to write the Hilbert transform as an average of *Haar shift operators*, which are uniformly bounded on $L^2(\mathbb{R})$.

It is natural to replace $L^2(\mathbb{R})$ by $L^p(\mathbb{R})$, and wonder whether the Hilbert transform is bounded on $L^p(\mathbb{R})$ for $p \neq 2$.

11.3.1. Bounded operators.

In this section we define the operator norm and bounded operators, and we explore equivalent definitions of these concepts. We specialize to the case of bounded operators from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$. In this case we see that translation- or dilation-invariance of a given linear operator restricts the pairs (p, q) for which boundedness occurs.

Definition 11.9. Let X and Y be normed spaces. An operator T is *bounded from* X *into* Y if $T : X \to Y$ and if for each f in X, the Y-norm of Tf is no greater than a fixed constant multiple of the X-norm of f. In other words, there is a constant C independent of f such that

$$\|Tf\|_Y \leq C\|f\|_X$$

for all f in X. The smallest such C is called the *operator norm* of T, and is often denoted by $\|T\|_{X\to Y}$ or simply $\|T\|$. In the special case when T is bounded from X into X we say T is *bounded on* X, and we denote its operator norm by $\|T\|_X$ or $\|T\|$. $\Diamond$

Exercise 11.10. (*Equivalent Definitions of Operator Norm*) Let $T : X \to Y$ be a bounded operator from the normed space X into the normed space Y. Verify that the following three definitions of the operator norm $\|T\|_{X\to Y}$ of T are equivalent:

$$\begin{aligned}
\|T\|_{X\to Y} &= \inf\{C > 0 : \|Tf\|_Y \leq C\|f\|_X \text{ for all } x \in X\}, \\
\|T\|_{X\to Y} &= \sup\left\{\frac{\|Tf\|_Y}{\|f\|_X} : f \neq 0, f \in X\right\}, \\
\|T\|_{X\to Y} &= \sup\left\{\frac{\|Tf\|_Y}{\|f\|_X} : \|f\|_X = 1, f \in X\right\}.
\end{aligned}$$

$\Diamond$

The Fourier multiplier T_m with symbol $m \in L^\infty(\mathbb{R})$ is a bounded operator on $L^2(\mathbb{R})$. Indeed, using Plancherel's Identity twice,

$$\|T_m f\|_2 = \|\widehat{T_m f}\|_2 = \left(\int_{\mathbb{R}} |m(\xi)\widehat{f}(\xi)|^2\, d\xi\right)^{1/2}$$
$$\leq \|m\|_\infty \|\widehat{f}\|_2 = \|m\|_\infty \|f\|_2.$$

The converse is also true. It depends on the properties of essentially bounded functions; see [**Graf2**, Section 2.5], or [**Duo**, Chapter 3, Section 5]. Furthermore, the operator norm from $L^2(\mathbb{R})$ into itself of the Fourier multiplier T_m is the L^∞-norm of its symbol $m \in L^\infty(\mathbb{R})$. The corresponding result in the periodic setting is proved in Lemma 11.76.

Exercise 11.11. (*The Norms of an Operator and of its Symbol*) Verify that if $m \in L^\infty(\mathbb{R})$ then $\|T_m\|_{L^2 \to L^2} = \|m\|_\infty$. $\Diamond$

Remark 11.12. Every bounded linear operator T on $L^2(\mathbb{R})$ that commutes with translations is given by convolution with a tempered distribution whose Fourier transform is a bounded function. It follows that every bounded linear operator on $L^2(\mathbb{R})$ that commutes with translations is a Fourier multiplier with bounded multiplier. (For a hint see [**SS2**, Problem 5, p.260]; for a full proof see [**Graf2**, Section 2.5].) In the discrete setting, we explored the interplay between convolutions and translation invariance in Chapter 6. In that setting, the interplay between convolution and Fourier multipliers is an immediate consequence of the time–frequency dictionary: convolutions are transformed into multiplication on the Fourier side, and vice versa. $\Diamond$

Exercise 11.13. (*Commutativity, Take 3*) Let T be a bounded operator on $L^2(\mathbb{R})$ that commutes with translations and dilations.

(i) Show that if T commutes with reflections, meaning that

$$\widetilde{(Tf)} = T(\tilde{f})$$

for all f in $L^2(\mathbb{R})$, then T is a constant multiple of the identity operator: $T = cI$ for an appropriate constant c.

(ii) Show that if T anticommutes with reflections, meaning that

$$\widetilde{(Tf)} = -T(\tilde{f})$$

for all f in $L^2(\mathbb{R})$, then T is a constant multiple of the Hilbert transform: $T = cH$ for an appropriate constant c. See [**Ste**, Chapter 2]. $\Diamond$

We are especially interested in boundedness on L^p spaces.

Definition 11.14. For p and q with $1 \le p \le \infty$ and $1 \le q \le \infty$, we say that an *operator T is bounded from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$* if $T : L^p(\mathbb{R}) \to L^q(\mathbb{R})$, and if there is a constant C such that

$$\|Tf\|_q \le C\|f\|_p \qquad \text{for all } f \in L^p(\mathbb{R}). \qquad \Diamond$$

By Exercise 11.10, we know that

$$\|T\|_{L^p \to L^q} = \sup\left\{\frac{\|Tf\|_{L^q(\mathbb{R})}}{\|f\|_{L^p(\mathbb{R})}} : f \in L^p(\mathbb{R}),\ \|f\|_{L^p(\mathbb{R})} = 1\right\}.$$

For linear operators, boundedness is equivalent to continuity at zero. In other words, a linear operator T is bounded from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$ if and only if $\|Tf_n\|_q \to 0$ for all sequences $\{f_n\} \subset L^p(\mathbb{R})$ such that $\|f_n\|_p \to 0$ as $n \to \infty$.

Exercise 11.15. Let T be a linear operator that maps $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$. Show that T is a bounded operator if and only if T is continuous at $f = 0$. $\Diamond$

For the Fourier transform, we discussed boundedness properties on $L^p(\mathbb{R})$ in Section 8.6, for a snapshot see Table 8.2. In particular, the Fourier transform is a bounded mapping from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$, provided that $1 \le p \le 2$ and p, q are dual exponents: $p^{-1} + q^{-1} = 1$. Furthermore, the *Hausdorff–Young Inequality*[3] holds for all such p, q and for all $f \in L^p(\mathbb{R})$:

$$\|\widehat{f}\|_q \le \|f\|_p, \quad \text{for } \frac{1}{p} + \frac{1}{q} = 1, \quad 1 \le p \le 2.$$

In other words, for all such p and q the L^q-norm of the Fourier transform of f is bounded by the L^p-norm of f. For the proof of this inequality see Corollary 11.59.

What can we say about the boundedness properties of the Hilbert transform from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$? If a linear operator is bounded from

[3] Named after the German mathematician Felix Hausdorff (1858–1942), and the English mathematician William Henry Young (1863–1942).

$L^p(\mathbb{R})$ into $L^q(\mathbb{R})$ and is dilation-invariant, then q must be equal to p (Exercise 11.16). The Hilbert transform is a dilation-invariant linear operator. Therefore it can only be bounded from $L^p(\mathbb{R})$ into $L^p(\mathbb{R})$, and not from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$ for $p \neq q$.

Exercise 11.16. (*Dilation-Invariant Operators from L^p to L^q*) For f in $L^p(\mathbb{R})$, let $\delta_a f(x) := f(ax)$, for all $a > 0$. Show that $\|\delta_a f\|_p = a^{-1/p}\|f\|_p$. Let T be a dilation-invariant linear operator that is bounded from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$. Check that for all $a > 0$, $\|Tf\|_q \leq a^{1/q-1/p}C\|f\|_p$, where $C = \|T\|_{L^p(\mathbb{R})\to L^q(\mathbb{R})}$. Conclude that $p = q$. $\Diamond$

It is also known that if a translation-invariant linear operator T is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$, then p cannot be greater than q. See Exercise 11.17 and [**Duo**, Section 6.7, p.66].

Exercise 11.17. (*Translation-Invariant Operators from L^p to L^q*) For f in $L^p(\mathbb{R})$, show that $\lim_{h\to\infty} \|f + \tau_h f\|_p = 2^{1/p}\|f\|_p$. Let T be a translation-invariant linear operator that is bounded from L^p into L^q. Check that $\|Tf + \tau_h Tf\|_q \leq C\|f + \tau_h f\|_p$, where $C = \|T\|_{L^p\to L^q}$. Let $h \to \infty$ to obtain $\|Tf\|_q \leq 2^{1/p-1/q}C\|f\|_p$. Conclude that $p \leq q$. Hint: First do the exercise for functions f of compact support; then approximate. $\Diamond$

11.3.2. Boundedness in $L^p(\mathbb{R})$ of the Hilbert transform. In the previous section we learned that dilation-invariant linear operators can only be bounded from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$ when $p = q$. The Hilbert transform is a linear operator that is both dilation- and translation-invariant. Is the Hilbert transform bounded on $L^p(\mathbb{R})$? The answer is *yes* for $1 < p < \infty$, and *no* for $p = 1$ or ∞; see Exercise 11.21.

On the time domain, the Hilbert transform is given by convolution with the distributional kernel $k_H(x) = \text{p.v. } 1/(\pi x)$, which is not an integrable function. If the kernel were an integrable function $k \in L^1(\mathbb{R})$, then the *convolution operator* T given by convolution against the kernel k,

$$Tf(x) = k * f(x) = \int_{\mathbb{R}} k(x-y)f(y)\,dy$$

would automatically be bounded on $L^p(\mathbb{R})$ for all p such that $1 \le p \le \infty$, by *Young's Inequality*:

$$\|f * k\|_p \le \|k\|_1 \|f\|_p.$$

(We prove Young's Inequality as Corollary 11.69.) Unfortunately,

the Hilbert transform kernel k_H is not integrable,

for two reasons: the integral of k_H on any interval containing 0 is infinite, as is the integral of k_H near infinity.

Even so, the Hilbert transform is bounded on $L^p(\mathbb{R})$ for all p with $1 < p < \infty$. Although it is not bounded on either $L^1(\mathbb{R})$ or $L^\infty(\mathbb{R})$, it satisfies a weaker boundedness property on each of these two spaces, as we will see below for the case $p = 1$. We state the precise result in the case $1 < p < \infty$; we sketch a proof in Section 11.5.2, once certain interpolation results have been established.

Theorem 11.18. *The Hilbert transform is a bounded operator on $L^p(\mathbb{R})$ for all p such that $1 < p < \infty$: for each such p there is a constant $C_p > 0$ such that for all $f \in L^p(\mathbb{R})$,*

$$\|Hf\|_p \le C_p \|f\|_p.$$

The Hilbert transform is the prototypical example of a *Calderón–Zygmund singular integral operator*[4]. These are operators given by integration against a kernel $K(x, y)$,

$$Tf(x) = \text{p.v.} \int_{\mathbb{R}} K(x, y) f(y)\, dy.$$

Near the diagonal $y = x$, Calderón–Zygmund kernels $K(x, y)$ grow like the Hilbert kernel $k_H(x - y)$, and therefore the integral must be understood in some principal value sense. The cancellation encoded in the symmetry of p.v. $1/x$ is replaced by an appropriate estimate on the gradient of the kernel, or on differences in the values of the kernel at certain pairs of points in $\mathbb{R}^2$.

Aside 11.19. A one-dimensional *standard kernel* K is a function on $\mathbb{R}\backslash\{0\}$ such that $|K(x, y)| \le C/|x - y|$, and there exists $\delta > 0$ such that $|K(x, y) - K(x, z)| \le C|y - z|^\delta / |x - y|^{1+\delta}$ if $|x - y| > 2|y - z|$, and

[4]Named after the Argentinian mathematician Alberto Pedro Calderón (1920–1998), and the Polish born American mathematician Antoni Zygmund (1900–1992).

$|K(x,y) - K(w,y)| \leq C|x-w|^\delta/|x-y|^{1+\delta}$ if $|x-y| > 2|x-w|$. The following are excellent references for the reader interested in pursuing the study of singular integral operators: [**Duo**], [**Graf2**], [**Graf3**], [**Sad**], [**Ste**], and [**Tor**]. $\Diamond$

Exercise 11.20. Show that the kernel $K_H(x,y) := k_H(x-y)$ of the Hilbert transform is a standard kernel. $\Diamond$

11.4. Weak-$L^p(\mathbb{R})$ and weak boundedness

In this section we introduce the space *weak-*$L^p(\mathbb{R})$ which strictly contains the space $L^p(\mathbb{R})$, for each p with $1 \leq p < \infty$. We also introduce the notion of *weak boundedness*, weaker than the usual L^p boundedness. Some operators are weakly bounded on $L^1(\mathbb{R})$ but not bounded, such as the Hilbert transform and also the *Hardy–Littlewood maximal function*; we discuss both examples. Finally we establish the *layer cake representation* for the L^p norm of a function in terms of its distribution function.

11.4.1. Weak-type $(1,1)$ vs bounded on $L^1(\mathbb{R})$. When $p = 1$, one can get away with a weaker notion of boundedness, which requires some of the ideas of measure theory. There is nothing special about $p = 1$ except that for the particular case under study, the Hilbert transform is not bounded in $L^1(\mathbb{R})$ (Exercise 11.21) although it is an operator of weak-type $(1,1)$. In subsequent sections we define this notion for other values of p. In this section we just want to justify the appearance of certain level sets, and of the inequalities that will help us define both the spaces weak-$L^p(\mathbb{R})$, and the notion of weak boundedness on L^p spaces.

Suppose $f \in L^1(\mathbb{R})$, and let $E_\lambda(f)$ be the set of points $x \in \mathbb{R}$ for which the absolute value $|f(x)|$ is larger than $\lambda > 0$, so that

$$E_\lambda(f) := \{x \in \mathbb{R} : |f(x)| > \lambda\} \subset \mathbb{R}. \tag{11.3}$$

First of all, functions in $L^1(\mathbb{R})$ are *measurable*. This means that the sets $E_\lambda(f)$ are themselves *measurable*; meaning that we can assign a *length* or *measure* to each such set, and this length coincides with ordinary length when the set is an interval. This measure maps the measurable subsets of $\mathbb{R}$ to nonnegative numbers, is additive for finite collections of disjoint sets, and has other convenient properties. Here the word *measurable* hides many technical details; see any text on measure theory such as [**SS2**], [**Bart1**], [**Fol**], and [**Roy**].

Suppose T is a linear bounded operator from $L^1(\mathbb{R})$ to itself, that is, there is a constant $C > 0$ such that for all $f \in L^1(\mathbb{R})$,

$$\|Tf\|_1 = \int_{\mathbb{R}} |Tf(x)|\,dx \leq C\|f\|_1.$$

Even if we are not fully acquainted with the Lebesgue integral, our experience with Riemann integration tells us that the following inequality is plausible, because on the right-hand side we are integrating a positive function on a subset of $\mathbb{R}$:

$$\|Tf\|_1 = \int_{\mathbb{R}} |Tf(x)|\,dx \geq \int_{E_\lambda(Tf)} |Tf(x)|\,dx.$$

By definition, if $x \in E_\lambda(Tf)$ then $|Tf(x)| > \lambda$, and so we can also believe[5] that

$$\begin{aligned}\int_{E_\lambda(Tf)} |Tf(x)|\,dx \geq \int_{E_\lambda(Tf)} \lambda\,dx &= \lambda \int_{E_\lambda(Tf)} 1\,dx \\ &= \lambda \int_{\mathbb{R}} \chi_{E_\lambda(Tf)}(x)\,dx.\end{aligned}$$

Finally if $E_\lambda(Tf)$ is an interval, then this last integral is equal to the length of $E_\lambda(Tf)$. Since $Tf \in L^1(\mathbb{R})$, the set $E_\lambda(Tf)$ is a *measurable set*, and we can extend the notion of length to such sets, obtaining the so-called *Lebesgue measure*. The Lebesgue measure of a measurable set A is denoted by $m(A)$, and it coincides with the Lebesgue integral[6] of χ_A on $\mathbb{R}$, in other words

$$m(A) = \int_A 1\,dx = \int_{\mathbb{R}} \chi_A(x)\,dx.$$

The symbol $m(A)$ is read as *the measure of the set* A.

[5] The properties of the integral that we are using are: (i) if $f \geq 0$ and $A \subset B$ then $\int_A f \leq \int_B f$; (ii) if $g \leq f$ on A, then $\int_A g \leq \int_A f$; (iii) if $\lambda \in \mathbb{R}$ then $\int_A \lambda f = \lambda \int_A f$; (iv) $\int_A 1\,dx = \int \chi_A(x)\,dx$.

[6] We do not define the Lebesgue integral precisely in this book. However the reader can see some comments in Appendix B, and also in Chapter 2.

What we have shown is that if T is bounded on $L^1(\mathbb{R})$, then there is a constant $C \geq 1$ such that for all $f \in L^1(\mathbb{R})$,

$$m\big(E_\lambda(Tf)\big) = m(\{x \in \mathbb{R} : |Tf(x)| > \lambda\}) \leq \frac{C}{\lambda}\|f\|_1. \tag{11.4}$$

An operator that satisfies inequality (11.4) is said to be of *weak-type* $(1,1)$. We have just shown that *bounded on* $L^1(\mathbb{R})$ implies *weak-type* $(1,1)$, see Lemma 11.26. The converse is false, in general. The Hilbert transform is not bounded in $L^1(\mathbb{R})$, as the following example shows, however it is of weak-type $(1,1)$.

Exercise 11.21. Show that $H\chi_{[0,1]}$ is neither an integrable function nor a bounded function, despite the fact that the characteristic function $\chi_{[0,1]}$ is both. Hint: Exercise 11.8 gives an explicit formula for $H\chi_{[0,1]}$. ♢

In Section 11.4.3 we define operators of *weak-type* (p,q), for all $p > 1$, $q > 1$. This concept will be important for the interpolation theorems in Section 11.5 and their applications.

11.4.2. The function spaces weak-$L^p(\mathbb{R})$. We now introduce the notion of *weak-L^p spaces*. As the name suggests, these spaces contain the ordinary L^p spaces; this fact is encoded in the famous *Chebyshev's inequality*[7]. In fact, the weak-L^p spaces are strictly larger than the L^p spaces.

Definition 11.22. Suppose $1 \leq p < \infty$. A measurable function f is said to belong to *weak-*$L^p(\mathbb{R})$ if there is a constant $C > 0$ such that for all $\lambda > 0$,

$$m(\{x \in \mathbb{R} : |f(x)| > \lambda\}) \leq \left(\frac{C}{\lambda}\right)^p.$$

The smallest constant C for which the above inequality holds is denoted by $[f]_p$, and called the *weak L^p-norm of f*.

For $p = \infty$, the space *weak-*$L^\infty(\mathbb{R})$ is defined to be $L^\infty(\mathbb{R})$, with the usual L^∞-norm. ♢

[7]Named after the Russian mathematician Pafnuty Lvovich Chebyshev (1821–1894). His name can be alternatively transliterated as Chebychev, Chebyshov, Tchebycheff or Tschebyscheff (French and German transcriptions).

The spaces *weak*-$L^p(\mathbb{R})$ are often denoted by $L^{p,\infty}(\mathbb{R})$. They form part of the larger family of functional spaces known as the *Lorenz spaces* $L^{p,q}(\mathbb{R})$, which are defined for $p,\ q \in (0,\infty]$. See [**Graf2**, Section 1.4] for more on Lorenz spaces.

Exercise 11.23. Show that if $1 \le p < \infty$ and $f \in L^p(\mathbb{R})$, then

$$m(\{x \in \mathbb{R} : |f(x)| > \lambda\}) \le \left(\frac{\|f\|_p}{\lambda}\right)^p.$$

This result is known as *Chebyshev's inequality*. Some authors reserve that name for the special case when $p = 1$. ♢

Exercise 11.23 shows that $L^p(\mathbb{R})$ is a subset of weak-$L^p(\mathbb{R})$, and that $[f]_p \le \|f\|_p$. The following example shows that it is a proper subset.

Example 11.24. The function $f(x) = |x|^{-1/p}$ is not in $L^p(\mathbb{R})$, but it is in weak-$L^p(\mathbb{R})$. For this f, we can compute explicitly that the set $E_\lambda(f)$ is exactly the interval $[-\lambda^{-p}, \lambda^{-p}]$. See Figure 11.1. Therefore

$$m(E_\lambda(f)) = 2\lambda^{-p} = \left(\frac{2}{\lambda}\right)^p.$$

Hence f is in weak-$L^p(\mathbb{R})$. Furthermore, $[f]_p = 2$. ♢

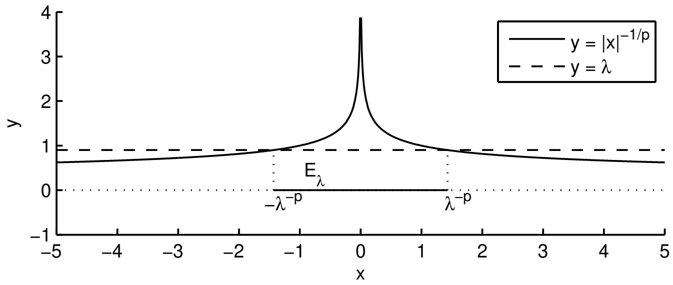

Figure 11.1. The function $f(x) = |x|^{-1/p}$ for $p = 3.4$, showing the set $E_\lambda = E_\lambda(f)$ for $\lambda = 0.9$.

11.4.3. Weak boundedness on $L^p(\mathbb{R})$. Some operators do not map $L^p(\mathbb{R})$ into another $L^q(\mathbb{R})$ space, but instead into the larger space of weak-$L^q(\mathbb{R})$. Here we define what it means for such operators to be bounded, in the sense of *weak-type* (p, q).

Definition 11.25. An operator T that maps $L^p(\mathbb{R})$ into weak-$L^q(\mathbb{R})$ is said to be of *weak-type* (p, q) if there is a constant $C > 0$ such that for all $\lambda > 0$,

$$m(\{x \in \mathbb{R} : |Tf(x)| > \lambda\}) \leq \left(\frac{C\|f\|_p}{\lambda}\right)^q. \tag{11.5}$$

The smallest constant $C > 0$ for which inequality (11.5) holds for all $f \in L^p(\mathbb{R})$ is called the *weak-(p, q) norm of T*, and it is denoted by $[T]_{L^p(\mathbb{R}) \to L^{q,\infty}(\mathbb{R})}$. ♢

In Section 11.4.1 we showed that if an operator is bounded in $L^1(\mathbb{R})$ then it is of weak-type $(1, 1)$. There is nothing special about the choice $p = q = 1$, as the next lemma reveals.

Lemma 11.26. *Suppose T is a linear bounded operator from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$. Then T is an operator of weak-type (p, q). Furthermore*

$$[T]_{L^p(\mathbb{R}) \to L^{q,\infty}(\mathbb{R})} \leq \|T\|_{L^p(\mathbb{R}) \to L^q(\mathbb{R})}.$$

Exercise 11.27. Prove Lemma 11.26, making appropriate modifications to the argument presented in Section 11.4.1 for the particular case $p = q = 1$. ♢

11.4.4. The Hilbert transform is of weak-type $(1, 1)$. We can now state precisely the boundedness result on $L^1(\mathbb{R})$ that holds for the Hilbert transform.

Theorem 11.28 (Kolmogorov's Inequality[8] (1927))**.** *The Hilbert transform is of weak-type $(1, 1)$, but is not bounded on $L^1(\mathbb{R})$.*

Exercise 11.21 shows that H cannot be bounded on $L^1(\mathbb{R})$, nor on $L^\infty(\mathbb{R})$. However, H is of weak-type $(1, 1)$. This can be seen to be a consequence of the *Calderón–Zygmund decomposition*, whose statement and discussion we defer to Project 12.14. The proof that this

[8] Named after the Russian mathematician Andrey Nikolaevich Kolmogorov (1903–1987).

decomposition exists uses another technique widely used in modern harmonic analysis, namely a *stopping-time argument*. Bounded functions are mapped into the larger space BMO of functions of *bounded mean oscillation*, which contains certain unbounded functions such as $\log|x|$. Project 12.14 deals with this decomposition and some of its consequences, while Project 12.15 discusses the John–Nirenberg Lemma for BMO functions, illustrating the power of stopping-time techniques. See [**Gar**, Chapter VI].

This behavior is shared by operators in the important class of *Calderón–Zygmund Singular Integral Operators* that we introduced in Section 11.3. The departure point of the Calderón–Zygmund theory is an *a priori* L^2-estimate; everything else unfolds from there. Having methods other than Fourier analysis to obtain such an L^2-estimate becomes important in other contexts such as the analysis of *singular integral operators of non-convolution type*; *analysis on curves, or on domains other than* $\mathbb{R}^n$; and *analysis on weighted* L^2 *spaces* (where the underlying measure is not Lebesgue measure), among others. One such method is *Cotlar's Lemma*, introduced in Project 12.11. Others include *Schur's Test*[9] and $T(1)$*-theorems*. You can find a discussion of these topics in the Lecture Notes [**Per**, Section 2] by the first author, and also in books such as [**Duo**, chapter 9] and [**Graf3**, Chapter 8].

Exercise 11.29. We know (Exercise 11.8) that the Hilbert transform of $\chi_{[0,1]}$ is given by $H\chi_{[0,1]}(x) = \pi^{-1}\log(|x|\,|x-1|^{-1})$. Find $C > 0$ such that for all $\lambda > 0$,

$$m\{x \in \mathbb{R} : |H\chi_{[0,1]}(x)| \geq \lambda\} \leq \frac{C}{\lambda}.$$

This fact alone does not prove that H is of weak-type $(1,1)$, but it shows that the function $H\chi_{[0,1]}$ is in weak-$L^1(\mathbb{R})$. Figure 11.2 may be useful here. See Project 12.13 for an exact calculation of the measure of the set $E_\lambda(H\chi_A)$, in the case when A is a finite disjoint union of bounded intervals. ◊

11.4.5. Maximal function vs weak-type $(1,1)$. The Lebesgue Differentiation Theorem (Exercise 9.44) holds for continuous functions, as a consequence of the Fundamental Theorem of Calculus. A

[9]Named after the German mathematician Issai Schur (1875–1941).

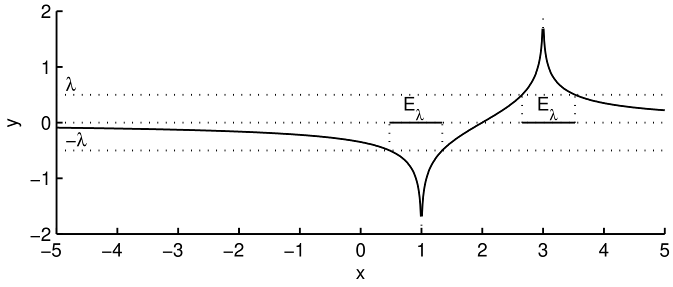

Figure 11.2. Graph of the Hilbert transform $H\chi_{[1,3]} = (1/\pi)\log|(x-1)/(x-3)|$ of the characteristic function of the interval $[1,3]$, showing the set $E_\lambda = E_\lambda(H\chi_{[1,3]})$ for $\lambda = 0.5$.

related operator, defined for all $x \in \mathbb{R}$ (possibly ∞), is the *Hardy–Littlewood maximal function*[10], denoted M, and defined by

$$Mf(x) = \sup_{x\in I} \frac{1}{|I|}\int_I |f(t)|\,dt. \tag{11.6}$$

The supremum in the definition of $Mf(x)$ is taken over all intervals containing x. Notice that the supremum is always defined, unlike the limit in the Lebesgue Differentiation Theorem as the intervals shrink to the point x, *a priori* that limit might not exist. The Lebesgue Differentiation Theorem guarantees both that the limit does exist, and that it equals the value of the function at the point x, for almost every x. Boundedness and weak-boundedness results for the maximal function can be used to deduce the Lebesgue Differentiation Theorem; see [**SS2**, Chapter 3, Section 1].

The maximal function is a bounded operator on $L^\infty(\mathbb{R})$, and is an operator of weak-type $(1,1)$.

Exercise 11.30. Verify that the Hardy–Littlewood maximal function is bounded on $L^\infty(\mathbb{R})$. Specifically, show that for all $f \in L^\infty(\mathbb{R})$ and each $x \in \mathbb{R}$,

$$|Mf(x)| \le \|f\|_\infty, \quad \text{and hence} \quad \|Mf\|_\infty \le \|f\|_\infty. \qquad \Diamond$$

[10]Named after the British mathematicians Godfrey Harold (G. H.) Hardy (1877–1947), and John Edensor Littlewood (1985–1977).

Theorem 11.31 (Hardy–Littlewood Maximal Theorem (1930))**.** *The Hardy–Littlewood maximal function is an operator of weak-type (1,1), meaning that there is a constant C such that*

$$m(\{x \in \mathbb{R} : |Mf(x)| > \lambda\}) \leq \frac{C\|f\|_1}{\lambda}$$

for each $f \in L^1(\mathbb{R})$ and for each $\lambda > 0$.

For a related operator, the *dyadic maximal function*, one can show that weak-boundedness in $L^1(\mathbb{R})$ is a direct consequence of the Calderón–Zygmund decomposition. This implies the result for the Hardy–Littlewood maximal function. For this and other arguments leading to the proof of Theorem 11.31, see Project 12.14.2. Maximal functions appear throughout harmonic analysis as controllers for other operators T, in the sense that $\|T\| \leq \|M\|$. It is important to understand their boundedness properties. For example, all convolution operators with good radial kernels can be bounded pointwise by a constant multiple of $Mf(x)$.

Exercise 11.32. Let $k_t \in L^1(\mathbb{R})$ be defined for $t > 0$ by $k_t(x) := t^{-1}K(|y|/t)$, where $K : [0,\infty) \to [0,\infty)$ is a monotone decreasing nonnegative function. This is a family of good kernels in $\mathbb{R}$; see Section 7.6. Show that the following pointwise estimate holds:

$$|K_t * f(x)| \leq Mf(x)\,\|k_1\|_1. \qquad \Diamond$$

As examples of such kernels, consider the *heat kernel* and the *Poisson kernel* on $\mathbb{R}$, introduced in Section 7.6. In these examples, $K(x) = e^{-|x|^2/4}/\sqrt{4\pi}$ for the heat kernel, and $K(x) = 1/(1 + |x|^2)$ for the *Poisson kernel*. A similar argument to the one that gives the Lebesgue Differentiation Theorem implies that if k_t is the heat kernel or the Poisson kernel, then $\lim_{t\to 0} k_t * f(x) = f(x)$ for almost every $x \in \mathbb{R}$, and for every $f \in L^1(\mathbb{R})$. See [**Pin**, Sections 3.5.1 and 3.5.2]. We already knew this fact for continuous functions, since in that case uniform convergence to f was proven in Chapter 7.

11.4.6. The distribution function and $L^p(\mathbb{R})$.

Suppose f is a measurable function, for example a step function. Then for each $\lambda \geq 0$ the set $E_\lambda(f)$ defined in formula (11.3) is measurable.

Definition 11.33. We define the *distribution function* d_f of f as follows: $d_f : [0,\infty) \to \mathbb{R}$ and

$$d_f(\lambda) := m\big(E_\lambda(f)\big) = m\{x \in \mathbb{R} : |f(x)| > \lambda\}, \qquad \text{for } \lambda \geq 0. \quad \Diamond$$

Thus the number $d_f(\lambda)$ answers the question 'How big is the set where $|f|$ is greater than λ?'

Exercise 11.34. Show that the distribution function d_f is a non-negative decreasing function of $\lambda \geq 0$. You will need to use two properties of Lebesgue measure that hold for all measurable subsets A and B of $\mathbb{R}$: (i) $m(A) \geq 0$ (positivity), (ii) If $A \subset B$ then $m(A) \leq m(B)$ (monotonicity). $\Diamond$

The distribution function is often useful in calculations, since integrals with respect to x involving the size $|f|$ of f can be converted to integrals with respect to λ involving $d_f(\lambda)$. Theorem 11.38 applies this idea in the important example of the L^p-norm $\|f\|_p$.

Exercise 11.35. Consider the step function

$$f(x) = 2\chi_{[-5,-2)]}(x) - 3\chi_{[-1,1)}(x) + 7\chi_{[6,7)}(x).$$

Show that the distribution function $d_f(\lambda)$ of $f(x)$ is the positive and decreasing function

$$d_f(\lambda) = 6\chi_{[0,2)}(\lambda) + 3\chi_{[2,3)}(\lambda) + \chi_{[3,7)}(\lambda).$$

The graphs of $f(x)$ and $d_f(\lambda)$ are shown in Figure 11.3. $\Diamond$

Exercise 11.36. Let $\{I_k : k = 1, \ldots, n\}$ be disjoint bounded intervals. Let f be the step function given by $f = \sum_{k=1}^n a_k \chi_{I_k}$. Find the distribution function d_f of f as a function of $\lambda \geq 0$. $\Diamond$

Exercise 11.37. Show that $d_f(\lambda) = d_{\tau_h f}(\lambda)$ for all $h \in \mathbb{R}$, and for all step functions $f = \sum_{k=1}^n a_k \chi_{I_k}$. In other words, the distribution function of f is identical to the distribution function of each translate $\tau_h f$ of f. As usual $\tau_h f(x) := f(x-h)$. This translation invariance holds for all measurable functions, not just for step functions. $\Diamond$

The previous exercise shows that the distribution function gives information about the size of f, but not about its behavior near any given point. However, it does give sufficient information to compute L^p-norms (which are translation invariant, see Exercise 11.39).

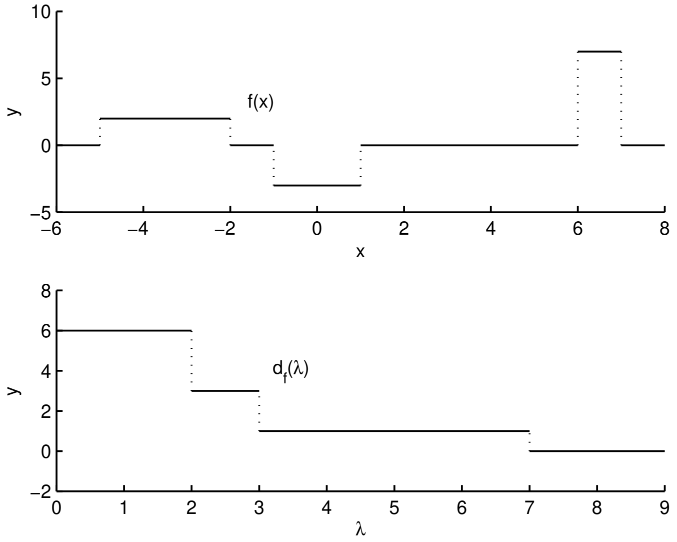

Figure 11.3. Graphs of the step function $f(x) = 2\chi_{[-5,-2)}(x) - 3\chi_{[-1,1)}(x) + 7\chi_{[6,7)}(x)$ in Exercise 11.35, and of its distribution function $d_f(\lambda) = 6\chi_{[0,2)}(\lambda) + 4\chi_{[2,3)}(\lambda) + \chi_{[3,7)}(\lambda)$.

Theorem 11.38. *For $f \in L^p(\mathbb{R})$ and $1 \leq p < \infty$ we have*

$$\|f\|_p^p = p \int_0^\infty \lambda^{p-1} d_f(\lambda)\, d\lambda.$$

Theorem 11.38 is a particular case of a theorem sometimes called the *layer cake representation*; see [**LL**, Section 1.13]. We state a more general version in Exercise 11.41.

Exercise 11.39. Use Theorem 11.38 and Exercise 11.37 to prove that the L^p-norm is invariant under translation; in other words,

$$\|f\|_p = \|\tau_h f\|_p \qquad \text{for all } h \in \mathbb{R},\ f \in L^p(\mathbb{R}). \tag{11.7}$$

(Exercise 7.40 asked you to prove this norm invariance for Schwartz functions.) ◊

Before we prove Theorem 11.38, let us calculate the L^p-norm and the λ integral for the step function in Exercise 11.35:

$$\|f\|_p^p = 2^p \int_{-5}^{-2} dx + 1 \int_{-1}^{1} dx + 7^p \int_{6}^{7} dx = 2^p \times 3 + 3^p \times 2 + 7^p \times 1.$$

On the other hand,

$$\begin{aligned} p \int_0^\infty & \lambda^{p-1} d_f(\lambda)\, d\lambda \\ &= 6 \int_0^2 p\lambda^{p-1}\, d\lambda + 3 \int_2^3 p\lambda^{p-1}\, d\lambda + \int_3^7 p\lambda^{p-1}\, d\lambda \\ &= 6\ \lambda^p\big|_0^2 + 4\ \lambda^p\big|_2^3 + \lambda^p\big|_3^7 \ = \ \|f\|_p^p. \end{aligned}$$

Exercise 11.40. Verify Theorem 11.38 for the general step function $f = \sum_{k=1}^n a_k \chi_{I_k}$. ♢

Proof of Theorem 11.38. By definition of d_f, we have

$$p \int_0^\infty \lambda^{p-1} d_f(\lambda)\, d\lambda = \int_0^\infty \int_{\{x:|f(x)|>\lambda\}} p\lambda^{p-1}\, dx\, d\lambda.$$

Figure 11.4 shows the region of integration for the double integral. By Fubini's Theorem (Theorem B.11) we may interchange the order of integration, obtaining

$$\int_{\mathbb{R}} \int_0^{|f(x)|} p\lambda^{p-1}\, d\lambda\, dx = \int_{\mathbb{R}} |f(x)|^p\, dx = \|f\|_p^p. \qquad \square$$

Exercise 11.41. (*Layer Cake Representation*) Show that for any continuously differentiable increasing function ϕ on $[0, \infty)$, we have

$$\int_{\mathbb{R}} \phi(|f|)\, dx = \int_0^\infty \phi'(\lambda) d_f(\lambda)\, d\lambda.$$

Theorem 11.38 corresponds to the choice $\phi(\lambda) = \lambda^p$. ♢

The distribution function is sublinear[11] with respect to f, in the sense described in Exercise 11.42.

Exercise 11.42. Show that for measurable functions f_1, $f_2 : \mathbb{R} \to \mathbb{R}$,

$$d_{f_1+f_2}(\lambda) \le d_{f_1}(\lambda/2) + d_{f_2}(\lambda/2). \qquad ♢$$

[11] A mapping from a vector space X into $\mathbb{R}$ is *sublinear* if $T(x+y) \le T(x) + T(y)$ for all x, $y \in X$.

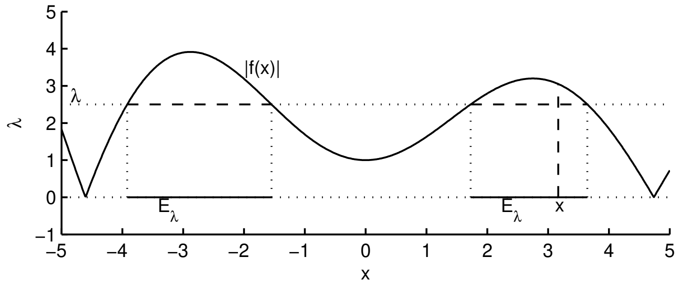

Figure 11.4. The domain of integration $-\infty < x < \infty$, $0 \leq \lambda \leq |f(x)|$ in the upper half plane, on which we use Fubini's Theorem in the proof of Theorem 11.38. Here $E_\lambda = E_\lambda(f) = \{x \in \mathbb{R} : |f(x)| > \lambda\}$. The dashed lines are those one might draw in a multivariable calculus class to explain interchanging the order of integration. We have used the function $f(x) = x\sin(0.72x)\exp(-x/20) + 1$.

Exercise 11.43. Show that the Hardy–Littlewood maximal operator M defined by (11.6) is a sublinear operator. That is, show that

$$M(f_1 + f_2)(x) \leq M(f_1)(x) + M(f_2)(x). \qquad \Diamond$$

11.5. Interpolation theorems

In this section we state some very useful theorems that are known as *interpolation theorems*. There are whole books devoted to such theorems. We highly recommend Cora Sadosky's book [**Sad**].

Essentially, these theorems tell us that if we know the boundedness properties of a linear operator at two endpoints, meaning two values of p, then we automatically obtain boundedness at many other intermediate points. This is a very powerful *convexity* result concerning bounded operators in L^p or weak-L^p spaces. We state the Marcinkiewicz and the Riesz–Thorin interpolation theorems, pointing out their similarities and differences. We outline the proof of the Marcinkiewicz interpolation theorem after presenting a convexity result pertaining to L^p spaces, and point to references for the proof of the Riesz–Thorin interpolation theorem. We present important applications of both theorems to the boundedness of the Hilbert transform and maximal functions in $L^p(\mathbb{R})$ (Section 11.5.2), and to prove a number of widely used inequalities from analysis (Section 11.6).

11.5.1. The Marcinkiewicz Interpolation Theorem. Our first interpolation theorem says that if an operator is of weak-type (p,p) and (q,q), where $1 \le p < q$, then it is of strong-type (r,r) (in other words, bounded on $L^r(\mathbb{R})$) for *all* r between p and q.

Theorem 11.44 (Marcinkiewicz Interpolation Theorem)**.** *Let T be a linear or sublinear operator defined on $L^p(\mathbb{R}) + L^q(\mathbb{R})$, where $1 \le p < q \le \infty$. Suppose T is of weak-type (p,p) and (q,q): there are positive constants A_p, A_q such that for all $\lambda > 0$ and for all $f \in L^p(\mathbb{R})$,*

$$d_{Tf}(\lambda) = m(\{x \in \mathbb{R} : |Tf(x)| > \lambda\}) \le \left(\frac{A_p\|f\|_p}{\lambda}\right)^p$$

and for all $f \in L^q(\mathbb{R})$,

$$d_{Tf}(\lambda) = m(\{x \in \mathbb{R} : |Tf(x)| > \lambda\}) \le \left(\frac{A_q\|f\|_q}{\lambda}\right)^q.$$

In the case $q = \infty$ the weak-type (q,q) hypothesis is replaced by the hypothesis that T is bounded on L^∞: namely $\|Tf\|_\infty \le A_\infty\|f\|_{L^\infty}$.

Then T is bounded on $L^r(\mathbb{R})$ for all r such that $p < r < q$. Moreover,

$$\|Tf\|_r \le 2\left(\frac{r}{r-p} + \frac{r}{q-r}\right)^{1/r} A_p^t A_q^{1-t}\|f\|_r,$$

where $t \in (0,1)$ satisfies $\dfrac{1}{r} = \dfrac{t}{p} + \dfrac{1-t}{q}$.

The proof of the Marcinkiewicz Interpolation Theorem can be found in many books on harmonic analysis or real analysis; for example [**Graf2**, Chapter 1], [**Sad**, Chapter 4], [**Duo**, Chapter 2], [**Tor**, Chapter IV]. The proof is outlined in Exercise 11.56.

Replacing the weak-type hypothesis by the stronger assumption of boundedness on $L^p(\mathbb{R})$ and on $L^q(\mathbb{R})$, one can get the same conclusion with a better constant: T is bounded on $L^r(\mathbb{R})$, and

$$\|Tf\|_r \le A_p^t A_q^{1-t}\|f\|_r.$$

This result can be obtained as a corollary of the Riesz–Thorin Interpolation Theorem (Section 11.5.4), which we state below, also without proof.

11.5.2. The maximal function and Hilbert transform are bounded on L^p. Knowing that the Hardy–Littlewood maximal function and the Hilbert transform are operators of weak-type (1,1), that the the Hilbert transform is bounded on $L^2(\mathbb{R})$, and that the maximal function is bounded on $L^\infty(\mathbb{R})$, we can deduce by interpolation the boundedness of these operators on $L^p(\mathbb{R})$ for $1 < p < \infty$.

For the Hardy–Littlewood maximal function we can prove boundedness on all L^p spaces for $1 < p < \infty$ in one stroke.

Corollary 11.45. *The maximal operator M is bounded on $L^p(\mathbb{R})$, for each p with $1 < p < \infty$.*

Proof. We have noted in the Hardy–Littlewood Maximal Theorem (Theorem 11.31) that M is of weak-type $(1, 1)$. (Project 12.14.2 outlines some different arguments leading to the proof of that result.) Also, M is bounded on $L^\infty(\mathbb{R})$ (Exercise 11.30). The Marcinkiewicz Interpolation Theorem then implies that M is bounded on $L^p(\mathbb{R})$ for all p with $1 < p < \infty$. □

For the Hilbert transform we use interpolation to capture the range $1 < p < 2$, and then a *duality argument* for the range $p > 2$.

Corollary 11.46. *The Hilbert transform H is bounded on $L^p(\mathbb{R})$ for $1 < p < 2$.*

Proof. We know from the definition that the Hilbert transform H is bounded on $L^2(\mathbb{R})$. We also know that boundedness in $L^2(\mathbb{R})$ implies that H is of weak-type $(2, 2)$. We mentioned without proof that H is of weak-type (1,1) (however see Project 12.14.3 for some hints on how to derive such a result from the Calderón–Zygmund decomposition). Then the Marcinkiewicz Interpolation Theorem implies that H is bounded on $L^p(\mathbb{R})$ for all p with $1 < p < 2$. □

For $p > 2$ we need to recall the definition of an an *adjoint* or *dual* operator in $L^p(\mathbb{R})$.

Definition 11.47. Let $p > 1$ and $p' > 1$ be dual exponents, so that $\frac{1}{p} + \frac{1}{p'} = 1$. The *adjoint* or *dual operator* T^* of a bounded operator T on $L^p(\mathbb{R})$ is the unique bounded operator[12] on the *dual space* $L^{p'}(\mathbb{R})$ such that for all $f \in L^p(\mathbb{R})$ and all $g \in L^{p'}(\mathbb{R})$,

$$\int_{\mathbb{R}} Tf(x)\,\overline{g(x)}\,dx = \int_{\mathbb{R}} f(x)\,\overline{T^*g(x)}\,dx. \qquad \Diamond$$

Example 11.48. The *adjoint* of H is $-H$, as the following formal calculation shows:

$$\begin{aligned}
\int Hf(x)g(x)\,dx &= \int (k_H * f)(x)g(x)\,dx \\
&= \int \left(\int k_H(x-y)f(y)\,dy \right) g(x)\,dx, \\
&= \int k_H(x-y)f(y)g(x)\,dx\,dy \qquad \text{(by Fubini)} \\
&= \iint f(y) \left(- \int k_H(y-x)g(x)\,dx \right) dy \\
&= \int f(y)\,(-Hg(y))\,dy.
\end{aligned}$$

It is well known that the operator norm on $L^{p'}(\mathbb{R})$ of the adjoint operator T^* of a bounded operator T on $L^p(\mathbb{R})$ coincides with the operator norm of T on $L^p(\mathbb{R})$.

Corollary 11.49. *The Hilbert transform H is bounded on $L^p(\mathbb{R})$ for $p > 2$.*

Proof. Given $p > 2$, its dual exponent p' lies between 1 and 2. By Corollary 11.46, H is bounded on $L^{p'}(\mathbb{R})$ for $1 < p' < 2$. Therefore its adjoint $-H$ is bounded on $L^p(\mathbb{R})$, where p is the dual exponent of p'. This duality argument now takes care of the range $2 < p < \infty$. $\square$

Overall, we have shown that the Hilbert transform is bounded on $L^p(\mathbb{R})$ for all $1 < p < \infty$, under the assumption that H is an operator of weak-type $(1, 1)$.

[12] Boundedness in $L^{p'}(R)$ of the dual operator is a consequence of the *Hahn–Banach Theorem* (see [**Sch**, Part II, Sections 2]), which is a standard result in functional analysis.

11.5.3. A result about L^p spaces. We prove a result similar in spirit to the Marcinkiewicz Interpolation Theorem. The result is a statement on the nature of the spaces $L^p(\mathbb{R})$. Note in particular that the constant appearing on the right-hand-side of the inequality in Proposition 11.50 is basically the same as the constant appearing in the interpolation theorem.

Proposition 11.50. *If $f \in L^p(\mathbb{R}) \cap L^q(\mathbb{R})$, where $1 \leq p < q \leq \infty$, then $f \in L^r(\mathbb{R})$ for all r with $p < r < q$. Moreover,*

$$\|f\|_r \leq \left(\frac{r}{r-p} + \frac{r}{q-r}\right)^{1/r} \|f\|_p^t \|f\|_q^{1-t},$$

where $t \in [0,1]$ is defined by $\dfrac{1}{r} = \dfrac{t}{p} + \dfrac{1-t}{q}$.

Proof. Suppose $q < \infty$. By Chebyshev's Inequality (Exercise 11.23), if $f \in L^p(\mathbb{R}) \cap L^q(\mathbb{R})$ then

$$d_f(\lambda) = m(\{x \in \mathbb{R} : |f(x)| > \lambda\}) \leq \min\left\{\frac{\|f\|_p^p}{\lambda^p}, \frac{\|f\|_q^q}{\lambda^q}\right\}.$$

Let $B = \left(\|f\|_q^q / \|f\|_p^p\right)^{1/(q-p)}$, and note that if $0 \leq \lambda \leq B$, then the above minimum is $\|f\|_p^p \lambda^{-p}$, and if $B \leq \lambda \leq \infty$, then the minimum is $\|f\|_q^q \lambda^{-q}$. We can now estimate the L^r-norm of f. By Theorem 11.38, we have

$$\begin{aligned}
\|f\|_r^r &= r\int_0^\infty \lambda^{r-1} d_f(\lambda)\, d\lambda \\
&\leq r\int_0^\infty \lambda^{r-1} \min\left\{\frac{\|f\|_p^p}{\lambda^p}, \frac{\|f\|_q^q}{\lambda^q}\right\} d\lambda \\
&= r\int_0^B \lambda^{r-1-p} \|f\|_p^p \, d\lambda + r\int_B^\infty \lambda^{r-1-q} \|f\|_q^q \, d\lambda \\
&= \frac{r}{r-p} \|f\|_p^p B^{r-p} + \frac{r}{q-r} \|f\|_q^q B^{r-q} \\
&= \left(\frac{r}{r-p} + \frac{r}{q-r}\right) \|f\|_p^{p\frac{(q-r)}{(q-p)}} \|f\|_q^{q\frac{(r-p)}{(q-p)}}.
\end{aligned}$$

Finally, if $t = \dfrac{p\,(q-r)}{r\,(q-p)}$ then

$$\frac{1}{r} = \frac{t}{p} + \frac{1-t}{q} \qquad \text{and} \qquad 1 - t = \frac{q\,(r-p)}{r\,(q-p)}.$$

The case when $q = \infty$ is left as an exercise for the reader. $\square$

The inequality in the proposition actually holds with constant 1 instead of $\left(\frac{r}{r-p} + \frac{r}{q-r}\right)$; see Exercise 11.62. We have presented the proof above because the same proof goes through if we replace L^p by weak-L^p, and then the resemblance to the Marcinkiewicz Interpolation Theorem is more noticeable. All we needed in this proof was Chebyshev's inequality, which always holds for functions in $L^p(\mathbb{R})$. But we could instead have required the function f to satisfy weak L^p- and L^q-estimates; in other words, replace $\|f\|_p$ by the quantity

$$[f]_p := \sup_{\lambda>0} \lambda(d_f(\lambda))^{1/p},$$

and $\|f\|_q$ by $[f]_q$. With those assumptions we could still conclude that $f \in L^r(\mathbb{R})$ for appropriate r. The following theorem states this idea more precisely.

Theorem 11.51. *If f lies in both weak-$L^p(\mathbb{R})$ and weak-$L^q(\mathbb{R})$, where $1 \le p < q$, then $f \in L^r(\mathbb{R})$ and*

$$\|f\|_r \le \left(\frac{r}{r-p} + \frac{r}{q-r}\right)^{1/r} [f]_p^t [f]_q^{1-t}, \qquad \textit{where } \frac{1}{r} = \frac{t}{p} + \frac{1-t}{q}.$$

Exercise 11.52. Following the proof of Proposition 11.50, write a complete proof of Theorem 11.51 for $q < \infty$. $\Diamond$

Exercise 11.53. Verify the following statements:

(i) If $f \in L^p(\mathbb{R}) \cap L^\infty(\mathbb{R})$ then $f \in L^r(\mathbb{R})$ for all $r > p$. Moreover

$$\|f\|_r \le \|f\|_p^t \|f\|_\infty^{1-t}, \quad \text{where } \frac{1}{r} = \frac{t}{p}.$$

(ii) If $f \in L^\infty(\mathbb{R})$ and f is in weak-$L^p(\mathbb{R})$, then $f \in L^r(\mathbb{R})$ for all $r > p$. Moreover

$$\|f\|_r \le \left(\frac{r}{r-p}\right)^{1/r} [f]_p^t \|f\|_\infty^{1-t}, \quad \text{where } \frac{1}{r} = \frac{t}{p}.$$ $\Diamond$

Now the reader may be able to devise a proof of the Marcinkiewicz interpolation theorem by herself. Here is a sequence of exercises leading to the proof. In the following exercises T is a linear or sublinear operator defined on $L^p(\mathbb{R}) + L^q(\mathbb{R})$, known to be of weak-type (p,p) and weak-type (q,q) for some p and q with $1 \le p < q$.

Exercise 11.54. Given $f \in L^r(\mathbb{R})$ and $\lambda > 0$, let $f_\lambda(x) = f(x)$ if $|f(x)| \leq \lambda$ and zero otherwise. Let $f^\lambda = f - f_\lambda$. Show that for all $p < r < q$, we have $f_\lambda \in L^p(\mathbb{R})$ and $f^\lambda \in L^q(\mathbb{R})$. See Exercise 8.51 for the case $p = 1$, $q = 2$. Show also that

$$\begin{aligned} \|f_\lambda\|_p^p &= \int_{\mathbb{R}} |f(t)|^p \chi_{\{x\in\mathbb{R}:|f(x)|\leq\lambda\}}(t)\,dt, \\ \|f^\lambda\|_q^q &= \int_{\mathbb{R}} |f(t)|^q \chi_{\{x\in\mathbb{R}:|f(x)|>\lambda\}}(t)\,dt. \end{aligned} \qquad \Diamond$$

Exercise 11.55. Given $\lambda > 0$, and $f \in L^r(\mathbb{R})$, decompose $f = f_p + f_q$, where $f_p = f_{\lambda/2}$, and $f_q = f^{\lambda/2}$. By Exercise 11.42,

$$d_{Tf}(\lambda) \leq d_{Tf_p}(\lambda/2) + d_{Tf_q}(\lambda/2).$$

Use Exercise 11.54, the weak-boundedness properties of T, and an argument similar to the one used in Proposition 11.50 to show that

$$\begin{aligned} \int_0^\infty d_{Tf_p}(\lambda/2)\lambda^{r-1}d\lambda &\leq \frac{2^{r-p}}{r-p}\|f\|_r^r, \\ \int_0^\infty d_{Tf_q}(\lambda/2)\lambda^{r-1}d\lambda &\leq \frac{2^{r-q}}{q-r}\|f\|_r^r. \end{aligned} \qquad \Diamond$$

Exercise 11.56. Combine Exercises 11.54 and 11.55 to give a proof of the Marcinkiewicz interpolation theorem for $q < \infty$. Make appropriate modifications to obtain a proof when $q = \infty$. $\Diamond$

11.5.4. The Riesz–Thorin Interpolation Theorem. The Riesz–Thorin Interpolation Theorem[13] is a variant of the interpolation theorem above, where the operator does not necessarily map a given space into itself. It is not exactly a generalization of the Marcinkiewicz Interpolation Theorem because it does not deal with operators of weak-type. However, it allows for different L^p spaces as inputs and outputs of the operator. The constant that appears in front of $A_1^t A_2^{1-t}$ is simpler than the one in the Marcinkiewicz Interpolation Theorem; it is just 1. This theorem has as almost immediate corollaries some of the most famous inequalities in analysis, as we show in Section 11.6.

[13] Named after the Hungarian mathematician Marcel Riesz (1886–1969), and the Swedish mathematician G. Olof Thorin (1912–2004), who was M. Riesz's student.

Theorem 11.57 (Riesz–Thorin Interpolation)**.** *Suppose p_1, p_2, q_1, and q_2 lie in $[1,\infty]$. For $0 \le t \le 1$ define p and q by*

$$\frac{1}{p} = \frac{t}{p_1} + \frac{1-t}{p_2}, \qquad \frac{1}{q} = \frac{t}{q_1} + \frac{1-t}{q_2}.$$

Suppose T is a linear operator from $L^{p_1}(\mathbb{R}) + L^{p_2}(\mathbb{R})$ into $L^{q_1}(\mathbb{R}) + L^{q_2}(\mathbb{R})$, such that T is bounded from $L^{p_i}(\mathbb{R})$ into $L^{q_i}(\mathbb{R})$ for $i = 1, 2$: there are constants A_1 and A_2 such that

$$\|Tf\|_{q_1} \le A_1\|f\|_{p_1} \qquad \text{for all } f \in L^{p_1}(\mathbb{R}),$$

and

$$\|Tf\|_{q_2} \le A_2\|f\|_{p_2} \qquad \text{for all } f \in L^{p_2}(\mathbb{R}).$$

Then T is bounded from $L^p(\mathbb{R})$ into $L^q(\mathbb{R})$. Moreover,

$$\|Tf\|_q \le A_1^t A_2^{1-t}\|f\|_p \qquad \text{for all } f \in L^p(\mathbb{R}).$$

A little experimentation with specific values of p_1, p_2, q_1, q_2, p, and q will help to build intuition about which combinations satisfy the hypotheses of the theorem.

See Figure 11.5 for a geometric interpretation of the interpolated points (p, q) as those points whose reciprocals $(1/p, 1/q)$ lie in the segment in $\mathbb{R}^2$ with endpoints $(1/p_1, 1/q_1)$ and $(1/p_2, 1/q_2)$.

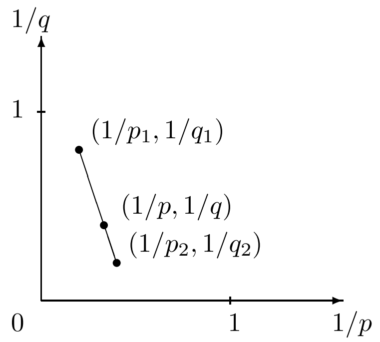

Figure 11.5. Riesz–Thorin Interpolation. Diagram in coordinates $(1/p, 1/q)$ showing the endpoint data p_1, q_1, p_2 and q_2, the interpolation segment, and the interpolated point at $\left(\frac{1}{p}, \frac{1}{q}\right) = \left(\frac{t}{p_1} + \frac{1-t}{p_2}, \frac{t}{q_1} + \frac{1-t}{q_2}\right)$.

Exercise 11.58. Sketch the geometric intepretation of the Marcinkiewicz Interpolation Theorem, in the same form as the diagram in Figure 11.5. ◊

The proof of Theorem 11.57 uses the so-called *three-lines theorem* for analytic functions, which can be found in many harmonic analysis textbooks. See for example [**SW2**, Chapter 5], [**Pin**, Section 3.2.1], or [**Graf2**, Chapter 1]. For readers with some background in complex analysis, this is a beautiful argument worth spending some time to understand.

11.6. A festival of inequalities

In this section we use Riesz–Thorin interpolation (Theorem 11.57) to prove several fundamental inequalities of analysis: the Hausdorff–Young Inequality; a generalized Hölder Inequality; a finite-dimensional version of Hölder's Inequality; Minkowski's Inequality and a finite-dimensional analogue; Minkowski's Integral Inequality; and Young's Inequality. The book by Elliot Lieb and Michael Loss [**LL**] is a wonderful resource for these and many other inequalities in analysis.

11.6.1. Hausdorff–Young Inequality.

The first inequality involves the Fourier transform; we mentioned it in Section 8.6 and Section 11.3.1.

Corollary 11.59 (Hausdorff–Young Inequality[14]). *Suppose* $1 \leq p \leq 2$ *and* $\dfrac{1}{p} + \dfrac{1}{q} = 1$. *If* f *is in* $L^p(\mathbb{R})$, *then its Fourier transform* $\widehat{f}$ *is in* $L^q(\mathbb{R})$ *and*

$$\|\widehat{f}\,\|_q \leq \|f\|_p.$$

Proof. We prove the Hausdorff–Young inequality by interpolating between Plancherel's Identity

$$\|\widehat{f}\,\|_2 = \|f\|_2$$

and the L^1–L^∞ boundedness of the Fourier transform, discussed in Chapter 8:

$$\|\widehat{f}\,\|_\infty \leq \|f\|_1.$$

[14]Named after the German mathematician Felix Hausdorff (1868–1942), and the British mathematician William Henry Young (1863–1942).

These are both strong L^p estimates, and one of them involves different domain and range, so we should use the Riesz–Thorin Interpolation Theorem with exponents

$$p_1 = 2, \quad q_1 = 2, \quad p_2 = 1, \quad q_2 = \infty$$

and constants $A_1 = 1$, $A_2 = 1$.

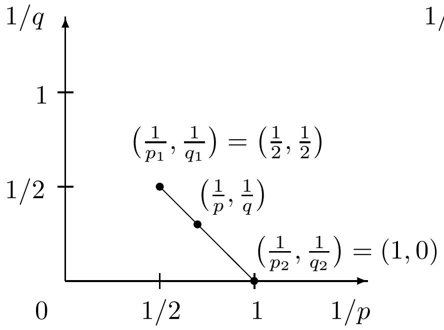

Figure 11.6. Interpolation diagram for the Hausdorff–Young Inequality. Endpoint data: $p_1 = 2$, $q_1 = 2$, $p_2 = 1$, $q_2 = \infty$.

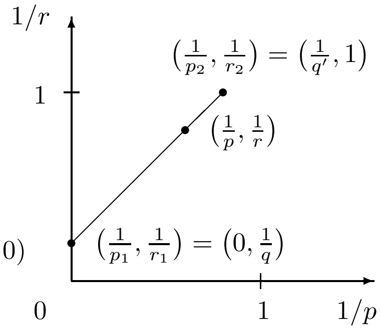

Figure 11.7. Interpolation diagram for the generalized Hölder inequality. Endpoint data: $p_1 = \infty$, $r_1 = q$, $p_2 = q'$, $r_2 = 1$. We have used $q = 5$, $q' = 5/4$.

As shown in Figure 11.6, the pair $(x, y) = (1/p, 1/q)$ reached by the interpolation theorem must lie on the line $x + y = 1$, for $x \in [1/2, 1]$. Therefore p and q are indeed dual exponents:

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad \text{and } 1 \le p \le 2.$$

Finally, the interpolation theorem gives us the desired inequality with constant $A_1^t A_2^{1-t} = 1$ for all $t \in [0, 1]$:

$$\|\widehat{f}\|_q \le \|f\|_p$$

for the above range of values of p and q, as claimed. □

The constant 1 is not optimal. The best possible constant in the Hausdorff–Young inequality was found by Bill Beckner in 1975 to be $\sqrt{p^{1/p}/q^{1/q}}$; see [**Bec**].

11.6.2. Hölder's Inequality. Our second inequality is a generalized version of the well known Hölder's Inequality[15]. It estimates the L^r-norm of the product of two functions in terms of the L^p- and L^q-norms of the factors, for appropriate r, p, and q.

Corollary 11.60 (Generalized Hölder Inequality)**.** *Suppose f is in $L^p(\mathbb{R})$, and g is in $L^q(\mathbb{R})$. Then their product fg lies in $L^r(\mathbb{R})$, and*

$$\|fg\|_r \leq \|f\|_p\|g\|_q, \qquad \textit{where} \qquad \frac{1}{r} = \frac{1}{p} + \frac{1}{q}.$$

Note the change of notation from that used above in the Riesz–Thorin Interpolation Theorem. Here the endpoint variables are called p and r, while q is an exponent with dual exponent q' so that $q^{-1} + q'^{-1} = 1$.

Proof. Fix q with $1 \leq q \leq \infty$, and take $g \in L^q$. First notice that if $p = \infty$, then $r = q$ and the following inequality holds:

$$\|fg\|_q \leq \|f\|_\infty\|g\|_q.$$

In fact,

$$\begin{aligned}\|fg\|_q &= \left(\int_{\mathbb{R}} |f(x)|^q|g(x)|^q\,dx\right)^{1/q}\\ &\leq \left(\|f\|_\infty^q \int_{\mathbb{R}} |g(x)|^q\,dx\right)^{1/q} \leq \|f\|_\infty\|g\|_q.\end{aligned}$$

Next recall *Hölder's Inequality*, which corresponds to the case $r = 1$, $p = q'$, where $\dfrac{1}{q} + \dfrac{1}{q'} = 1$:

$$\|fg\|_1 = \int_{\mathbb{R}} |f(x)g(x)|\,dx \leq \|f\|_{q'}\|g\|_q. \tag{11.8}$$

We indicate how to prove inequality (11.8) in Exercise 11.64, in case the reader has never done it before.

Apply the Riesz–Thorin Interpolation Theorem with $p_1 = \infty$, $q_1 = q$, $p_2 = q'$, $q_2 = 1$, and with the linear operator given by pointwise multiplication by g as follows: $(Tf)(x) = f(x)g(x)$ for all x. This time the constants are $A_1 = A_2 = \|g\|_q$.

[15]Named after the German mathematician Otto Ludwig Hölder (1859–1937).

As shown in Figure 11.7, the reciprocals of the targeted exponents (x, y) satisfy the equation of the line of slope 1 that passes through the point $(0, 1/q)$, that is the segment $y = x + 1/q$, $0 \le x \le 1/q'$. (In Figure 11.7 we have used the values $1/q = 1/5$ and $1/q' = 4/5$.) Therefore, denoting the exponents by $p = 1/x$ and $r = 1/y$, we conclude that they must satisfy

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q},$$

and the interpolation theorem tells us that

$$\|fg\|_r \le A_1^t A_2^{1-t} \|f\|_p = \|g\|_p^t \|g\|_p^{1-t} \|f\|_p = \|f\|_p \|g\|_q,$$

as required. □

Exercise 11.61. Verify that the constant 1 in the generalized Hölder inequality cannot be improved. Hint: consider the case $r = 1$, $p = q = 2$. ♢

Exercise 11.62. (*A Better Interpolation Constant*) Show that Proposition 11.50 holds with constant 1. In other words, show that if $1 \le p < r < q \le \infty$, and $f \in L^p(\mathbb{R}) \cap L^q(\mathbb{R})$, then $f \in L^r(\mathbb{R})$ and

$$\|f\|_r \le \|f\|_p^t \|f\|_q^{1-t}, \qquad \text{where } \frac{1}{r} = \frac{t}{p} + \frac{1-t}{q}.$$

Hint: For $q < \infty$, use Hölder's inequality (11.8) with dual exponents $p/(tr)$ and $q/((1-t)r)$, and let $f \to |f(x)|^{tr}$, $g \to |f(x)|^{(1-t)r}$. ♢

We have just encountered Hölder's inequality (inequality (11.8)). In the case $q = q' = 2$, it reduces to the Cauchy–Schwarz inequality, which we used in several arguments involving the Haar basis. The general case appeared when we discussed L^p-functions as distributions, in Section 8.6. Here is a finite-dimensional version.

Exercise 11.63. (*Finite-Dimensional Analogue of Hölder's Inequality*) Suppose a_i, $b_i \in \mathbb{R}$ for $i = 1, \ldots, n$. Prove that

$$\sum_{i=1}^n |a_i b_i| \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} \left(\sum_{i=1}^n |b_i|^{p'}\right)^{1/p'}, \qquad \text{where } \frac{1}{p} + \frac{1}{p'} = 1.$$

Show that the special case $n = 2$ and $p = p' = 2$ is a restatement of the geometric fact that $|\cos\theta| \le 1$. Hint: The following elementary inequality holds (why?) whenever a, $b \geq 0$ and $0 < \lambda < 1$:

$$a^{\lambda} b^{1-\lambda} \leq \lambda a + (1-\lambda) b.$$

For each $j = 1, \ldots, n$, set $a = |a_j|^p / \sum_{i=1}^n |a_i|^p$, $b = |b_j|^q / \sum_{i=1}^n |b_i|^q$, and $\lambda = 1/p$; then sum over j. ♢

Exercise 11.64. Use a limiting argument similar to the one we used in Chapter 8 to pass from the finite sums to Riemann sums, and taking limits, to integrals (to prove a particular version of the Minkowski's Integral Inequality), to obtain Hölder's inequality (11.8), at least for Riemann-integrable functions.

Alternatively, follow the argument in the proof of the finite-dimensional Holder inequality, except that now for each $x \in \mathbb{R}$ you set $a = |f(x)|^p / \|f\|_p^p$, $b = |g(x)|^q / \|g\|_q^q$, and $\lambda = 1/p$, and then integrate with respect to x. ♢

11.6.3. Minkowski's Inequality.

We can deduce the *triangle inequality* in $L^p(\mathbb{R})$ or $L^p(\mathbb{T})$, which is one of the defining properties of a norm, from Hölder's inequality. In this setting the triangle inequality bears the name of Minkowski.

Lemma 11.65 (Minkowski's Inequality[16]). *If $1 \leq p \leq \infty$, and f, $g \in L^p(\mathbb{R})$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. The result is clear if $p = 1$, if $p = \infty$, or if $f + g = 0$. Otherwise, we can estimate pointwise by means of the triangle inequality for complex numbers: $|f(x) + g(x)|^p \leq (|f(x)| + |g(x)|)\, |f(x) + g(x)|^{p-1}$. Now integrate over $\mathbb{R}$ to get

$$\|f+g\|_p^p \leq \int_{\mathbb{R}} |f(x)||f(x)+g(x)|^{p-1}\, dx + \int_{\mathbb{R}} |g(x)||f(x)+g(x)|^{p-1}\, dx,$$

[16] Named after the German mathematician Hermann Minkowski (1864–1909).

and apply Hölder's inequality, observing that if $\frac{1}{p}+\frac{1}{p'}=1$, then $(p-1)p'=p$. Thus $\|f+g\|_p^p$ is less than or equal to

$$\begin{aligned}\|f+g\|_p^p &\leq \Big(\int_{\mathbb{R}} |f(x)|^p\,dx\Big)^{1/p}\Big(\int_{\mathbb{R}} |f(x)+g(x)|^p\,dx\Big)^{1/p'}\\ &\qquad + \Big(\int_{\mathbb{R}} |g(x)|^p\,dx\Big)^{1/p}\Big(\int_{\mathbb{R}} |f(x)+g(x)|^p\,dx\Big)^{1/p'}\\ &= \big(\|f\|_p+\|g\|_p\big)\|f+g\|_p^{p/p'}.\end{aligned}$$

Dividing both sides by $\|f+g\|_p^{p/p'}$, which we know is not zero, we get exactly the desired inequality. □

One can have N summands, or even a continuum of summands (an integral!). That is the content of the following inequality, which we mentioned in the proof of Theorem 7.39.

Lemma 11.66 (Minkowski's Integral Inequality)**.** *Let $F(x,y)$ be a function of two variables such that for a.e. $y\in\mathbb{R}$, the function $F_y(x):=F(x,y)$ belongs to $L^p(\mathbb{R},dx)$, and the function $G(y)=\|F_y\|_p$ is in $L^1(\mathbb{R})$. Then the function $F_x(y):=F(x,y)$ is in $L^1(\mathbb{R},dy)$ for a.e. $x\in\mathbb{R}$, and the function $H(x):=\int_{\mathbb{R}}F(x,y)\,dy$ is in $L^p(\mathbb{R},dx)$. Moreover,*

$$\Big\|\int_{\mathbb{R}} F(x,y)\,dy\Big\|_{L^p(dx)} \leq \int_{\mathbb{R}} \|F(\cdot,y)\|_{L^p(dx)}\,dy.$$

This inequality is a standard topic in real analysis courses. It is helpful to keep in mind that it is the triangle inequality in $L^p(\mathbb{R})$ for integrals instead of sums.

Exercise 11.67. Verify that if f_1, f_2, …, $f_N\in L^p(\mathbb{R})$ then

$$\|f_1+f_2+\cdots+f_N\|_p \leq \|f_1\|_p+\|f_2\|_p+\cdots+\|f_N\|_p. \qquad \Diamond$$

Here is a finite-dimensional version.

Exercise 11.68. (*Finite-dimensional analogue of Minkowski's Inequality*) Let a_i, $b_i\in\mathbb{R}$, for $i=1,\ldots,n$. Prove that

$$\Big(\sum_{i=1}^n |a_i+b_i|^p\Big)^{1/p} \leq \Big(\sum_{i=1}^n |a_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^n |b_i|^p\Big)^{1/p}.$$

Hint: Follow the proof of Minkowski's inequality (Lemma 11.65); the integrals are replaced by finite sums. ♢

11.6.4. Young's inequality.

Our last inequality relates the L^r-norm of the convolution of an L^p function and an L^q function to their norms, for appropriate r, p, and q.

Corollary 11.69 (Young's Inequality[17]). *If $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, then $f * g \in L^r(\mathbb{R})$, where $\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}$, and*

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Proof. We prove Young's Inequality by interpolation. We can check directly that for fixed $g \in L^p(\mathbb{R})$, if $f \in L^1(\mathbb{R})$ then

$$\|f * g\|_p \leq \|g\|_p \|f\|_1, \tag{11.9}$$

and if $f \in L^{p'}(\mathbb{R})$ then

$$\|f * g\|_\infty \leq \|g\|_p \|f\|_{p'}, \tag{11.10}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Then use the Riesz–Thorin Interpolation Theorem. (The first inequality is an easy consequence of Minkowski's Integral Inequality, and the fact that the L^p-norm is invariant under translations. The second inequality is a direct application of Hölder's inequality and once again the translation invariance of the L^p-norm.) □

The constant 1 in Young's Inequality is not optimal. Bill Beckner found the optimal constant to be $(p^{1/p} q^{1/q})^{1/2}$; see [**Bec**].

Exercise 11.70. Sketch the interpolation diagram for the proof of Young's Inequality, in the same form as the diagrams in Figures 11.5, 11.6 and 11.7. ♢

Exercise 11.71. Prove inequality (11.9) when $r = p$, $q = 1$. ♢

Exercise 11.72. Prove inequality (11.10) when $r = \infty$, $q = p'$. ♢

[17] Named after the British mathematician William Henry Young (1863–1942).

11.6.5. Return to approximations of the identity. We prove an approximation-of-the-identity convergence theorem for L^p functions.

Theorem 11.73. *Suppose K is a positive function on $\mathbb{R}$ such that $\|K\|_1 = \int_{\mathbb{R}} K(x)\,dx = 1$. The functions $K_t(x) = t^{-1}K(t^{-1}x)$, for $t > 0$, define an approximation of the identity in $\mathbb{R}$. If f is in $L^p(\mathbb{R})$, then*

$$\lim_{t\to 0} \|K_t * f - f\|_p = 0, \qquad \textit{for } 1 \le p < \infty.$$

Proof. We have already proved this result for Schwartz functions (Theorem 7.39). We also know that $\mathcal{S}(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ (see Appendix B) for each p with $1 \le p < \infty$. We will use a standard $\varepsilon/3$ approximation argument to prove the theorem.

Define the linear operator $T_t : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ by convolution with the kernel K_t, $T_t(f) := K_t * f$. By Young's Inequality the operators T_t are uniformly bounded on $L^p(\mathbb{R})$:

$$\|T_t f\|_p = \|K_t * f\|_p \le \|K_t\|_1 \|f\|_p = \|K\|_1 \|f\|_p = \|f\|_p.$$

Given a function $f \in L^p(\mathbb{R})$ and a number $\varepsilon > 0$, there is a function $g \in \mathcal{S}(\mathbb{R})$ such that $\|g - f\|_p \le \varepsilon/3$. For $g \in \mathcal{S}(\mathbb{R})$ and $\varepsilon > 0$ there is a $t_0 > 0$ such that for all t with $0 < t < t_0$, $\|T_t g - g\|_p \le \varepsilon/3$. Finally the linearity and uniform boundedness of the operators T_t imply that $\|T_t f - T_t g\|_p = \|T_t(f - g)\| \le \|f - g\|_p \le \varepsilon/3$. Overall, we conclude that given $\varepsilon > 0$, there is some $t_0 > 0$ such that for all t with $0 < t < t_0$,

$$\|T_t f - f\|_p \le \|T_t f - T_t g\|_p + \|T_t g - g\|_p + \|g - f\|_p \le \varepsilon.$$

Therefore, for all $f \in L^p(\mathbb{R})$,

$$\|T_t f - f\|_p \to 0 \qquad \text{as } t \to 0. \qquad \square$$

The conclusion of Theorem 11.73 also holds for functions $f \in L^p(\mathbb{T})$. Here is an application. Since the Fejér and Poisson kernels on $\mathbb{T}$ are approximations of the identity, convergence in $L^p(\mathbb{T})$ to f of the Cesàro and Abel means follows:

$$\|\sigma_N f - f\|_{L^p(\mathbb{T})} \to 0, \qquad \text{as } N \to \infty; \tag{11.11}$$

$$\|A_r f - f\|_{L^p(\mathbb{T})} \to 0, \qquad \text{as } r \to 1. \tag{11.12}$$

This argument does not work for the partial Fourier sums $S_N f$, because the Dirichlet kernels do not define an approximation of the identity.

In the last section of this chapter we will show that the uniform (in N) L^p-boundedness on $\mathbb{T}$ of the partial Fourier sums $S_N f$ can be deduced from the L^p-boundedness on $\mathbb{T}$ of the periodic Hilbert transform. It will follows that the partial Fourier sums $S_N f$ converge in $L^p(\mathbb{T})$ to f, as N tends to infinity.

However, by Lemma 4.42, the fact that the Cesàro and Abel means of f converge in $L^p(\mathbb{T})$ to f implies that the same must be true for the partial Fourier sums of f, without any reference to the Hilbert transform, at least for twice-differentiable functions f. For those functions we proved that the partial Fourier sums converge uniformly on $\mathbb{T}$, and that implies boundedness in $L^p(\mathbb{T})$. This last statement is false in $\mathbb{R}$, as we have seen. The argument that we describe in Section 11.7 can be adapted to $\mathbb{R}$ to conclude that the partial Fourier integrals $S_R f$ converge to f in $L^p(\mathbb{R})$ for $1 < p < \infty$. See Exercise 11.83.

11.7. Some history to conclude our journey

Why did mathematicians get interested in the Hilbert transform? In this section we present two classical problems where the Hilbert transform appeared naturally. The first is in connection to Fourier series, where the convergence of the partial Fourier sums $S_N f$ to f in $L^p(\mathbb{T})$, for $1 < p < \infty$, follows from the boundedness of the periodic Hilbert transform. The second is in connection to complex analysis, where the Hilbert transform links the Poisson kernel and the conjugate Poisson kernel on the upper half-plane. These ideas bring us full circle back to where we started this journey: Fourier series.

11.7.1. Connection to Fourier series. For integrable functions on $\mathbb{T} = [0, 1]$, the n^{th} Fourier coefficient is defined by the formula

$$\widehat{f}(n) := \int_0^1 f(\theta)e^{-2\pi in\theta}\, d\theta.$$

(In this section we are identifying the unit circle $\mathbb{T}$ with the interval $[0, 1)$, instead of $[-\pi, \pi)$ as we did earlier in the book.) Since

$L^2(\mathbb{T}) \subset L^1(\mathbb{T})$, the above integral is also well defined for square-integrable functions f. We already know that the trigonometric system $\{e^{2\pi in\theta}\}_{n\in\mathbb{Z}}$ is an orthonormal complete system in $L^2(\mathbb{T})$. Therefore the following reconstruction and isometry formulas hold in $L^2(\mathbb{T})$:

$$f(\theta) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{2\pi in\theta}, \qquad \|f\|^2_{L^2(\mathbb{T})} = \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2.$$

The N^{th} partial sum is given by

$$S_N f(\theta) = \sum_{|n|\leq N} \widehat{f}(n)e^{2\pi in\theta}.$$

In the nineteenth century, mathematicians asked: For which 2π-periodic functions f is it true that

$$\lim_{N\to\infty} S_N f(\theta) = f(\theta)$$

at a given point $\theta \in \mathbb{T}$? As we saw in Section 3.3, some partial answers were found. Something more than continuity at the point θ was always required, for example Dini's condition. In 1889, Du Bois–Reymond found a continuous function whose partial Fourier sum diverges at one point, see Project 12.6. Summation methods, see Chapter 4, were introduced to counter these so-called pathologies. With the advent of measure theory and L^p spaces, new questions were formulated.

- **Is there convergence a.e.?** That is, does there exist an exceptional set E of measure zero such that for all $\theta \in \mathbb{T}$ not included in E,
$$\lim_{N\to\infty} S_N f(\theta) = f(\theta)\ ?$$
- **Is there convergence in the L^p-sense?** That is, does
$$\lim_{N\to\infty} \|S_N f - f\|_{L^p(\mathbb{T})} = 0\ ?$$

The answer to the second question is 'yes' for $1 < p < \infty$. For $p = 2$ it is Plancherel's Identity, and for $1 < p < \infty$ can be seen to be a consequence of the boundedness of the periodic analogue of the Hilbert transform in such L^p spaces, as we show below. The first question is much more difficult. For $p = 2$ the positive answer was given by L. Carleson in his 1965 paper [**Car**], settling the question for periodic continuous functions, which had remained open until then. We discussed some of this material in Section 3.3.

Define the *periodic Hilbert transform* or *conjugate function* on the Fourier side by

$$(H_P f)^\wedge(n) = -i \operatorname{sgn}(n)\widehat{f}(n).$$

This definition corresponds to the second periodization method introduced in Section 8.5.1 when discussing the Poisson summation formula.

Exercise 11.74. Given $f(\theta) = \sin(2\pi\theta)$ on $[0, 1]$, find $H_P f(\theta)$. $\Diamond$

Exercise 11.75. Let $f(\theta) = \sum_{|n|\leq N} c_n e^{in\theta}$ be an arbitrary trigonometric polynomial. Show that

$$H_P f(\theta) = -i\sum_{n=1}^{N} c_n e^{in\theta} + i\sum_{-N}^{n-1} c_n e^{in\theta}. \qquad \Diamond$$

The operator H_P is an example of a *Fourier multiplier* T_m on the circle. These operators are linear transformations on $L^2(\mathbb{T})$, defined on the Fourier side by multiplication of the Fourier coefficients $\widehat{f}(n)$ of a given $f \in L^2(\mathbb{T})$ by a complex number $m(n)$ depending on the frequency n. Specifically, the definition of a Fourier multiplier operator T_m is given by the formula

$$(T_m f)^\wedge(n) = m(n)\widehat{f}(n), \tag{11.13}$$

where the sequence $\{m(n)\}_{n\in\mathbb{Z}}$ is called the *symbol* of T_m.

The sequence $\{m(n)\}_{n\in\mathbb{Z}}$ is said to be a *bounded sequence*, written $\{m(n)\}_{n\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z})$, if there is a constant $M > 0$ such that $|m(n)| \leq M$ for all $n \in \mathbb{Z}$. For a bounded sequence $\{m(n)\}_{n\in\mathbb{Z}}$, we have

$$\begin{aligned}\|T_m f\|^2_{L^2(\mathbb{T})} &= \sum_{n\in\mathbb{Z}} |(T_m f)^\wedge(n)|^2 = \sum_{n\in\mathbb{Z}} |m(n)\widehat{f}(n)|^2 \\ &\leq M^2 \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 = M^2\|f\|^2_{L^2(\mathbb{T})},\end{aligned}$$

where the first and last equalities are applications of Parseval's Identity. We conclude that if the sequence $\{m(n)\}_{n\in\mathbb{Z}}$ is in $\ell^\infty(\mathbb{Z})$ then the multiplier T_m with symbol $\{m(n)\}_{n\in\mathbb{Z}}$ is bounded on $L^2(\mathbb{T})$. The converse is also true, and easy to prove in the periodic case.

Lemma 11.76. *A Fourier multiplier T_m on the circle $\mathbb{T}$ with symbol $\{m(n)\}_{n\in\mathbb{Z}}$ is bounded on $L^2(\mathbb{T})$ if and only if the sequence $\{m(n)\}_{n\in\mathbb{Z}}$ lies in $\ell^\infty(\mathbb{Z})$.*

Proof. We have just proved the 'if' direction. For the 'only if' direction, suppose that T_m is bounded on $L^2(\mathbb{T})$. Then there exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{T})$,

$$\|T_m f\|_{L^2(\mathbb{T})} \leq C\|f\|_{L^2(\mathbb{T})}.$$

In particular this is true for the exponential functions $e_n(\theta) = e^{2\pi in\theta}$, $n \in \mathbb{Z}$. The exponential functions form an orthonormal basis in $L^2(\mathbb{T})$; each of them has norm 1. Their Fourier coefficients have the simple form

$$\widehat{e_n}(k) = \begin{cases} 1, & \text{if } k = n; \\ 0, & \text{if } k \neq n. \end{cases}$$

Therefore, using Parseval's Identity twice, for each $n \in \mathbb{Z}$ we have

$$|m(n)|^2 = \sum_{k\in\mathbb{Z}} |m(k)\widehat{e_n}(k)|^2 = \|T_m e_n\|^2_{L^2(\mathbb{T})} \leq C^2\|e_n\|^2_{L^2(\mathbb{T})} = C^2.$$

Thus $|m(n)| \leq C$ for all $n \in \mathbb{Z}$, and the sequence $\{m(n)\}_{n\in\mathbb{Z}}$ is bounded, as required. □

The operators H and H_P share similar boundedness properties.

Theorem 11.77. *The periodic Hilbert transform H_P is bounded on $L^p(\mathbb{T})$ for $1 < p < \infty$: there is a constant $C_p > 0$ such that for all $f \in L^p(\mathbb{T})$,*

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

The operation of taking the N^{th} partial Fourier sum of a nice function is another important example of a Fourier multiplier on the unit circle $\mathbb{T}$. Specifically,

$$(S_N f)^\wedge(n) = m_N(n)\widehat{f}(n),$$

where the symbol $\{m_N(n)\}_{n\in\mathbb{Z}}$ of S_N is the sequence

$$m_N(n) = \begin{cases} 1, & \text{if } |n| \leq N; \\ 0, & \text{otherwise.} \end{cases}$$

The next two exercises show how to write S_N in terms of modulations of the periodic Hilbert transform, and that these modulation operators preserve the L_p-norm of a function.

Exercise 11.78. (*Partial Fourier sums as modulations of the periodic Hilbert transform*) Check that if $|n| \neq N$, then

$$m_N(n) = \frac{1}{2}\big(\operatorname{sgn}(n+N) - \operatorname{sgn}(n-N)\big),$$

If $|n| = N$ then $m_N(n) = \frac{1}{2}$.

Let M_N denote the modulation operator, defined by $M_N f(\theta) := f(\theta)e^{2\pi i\theta N}$. Show that for all $N \in \mathbb{Z}$,

$$i(M_N H_P M_{-N})^\wedge(n) = \operatorname{sgn}(n+N)\widehat{f}(n).$$

(Recall that the Fourier transform converts modulations into translations.) Deduce that

$$\begin{aligned} S_N f(\theta) &= -\frac{i}{2}(M_N H_P M_{-N} f(\theta) - M_{-N} H_P M_N f(\theta)) \\ &\quad + \frac{1}{2}(\widehat{f}(N)e^{iN\theta} + \widehat{f}(-N)e^{-iN\theta}), \end{aligned} \tag{11.14}$$

by applying the inverse Fourier transform. ♢

Exercise 11.79. (*Modulations preserve L^p-norms*) Show that

$$\|M_N f\|_{L^p(\mathbb{T})} = \|f\|_{L^p(\mathbb{T})} \qquad \text{for all } N \in \mathbb{Z}.$$

In other words, the modulation operators preserve the L^p-norm. ♢

We are ready to prove that $S_N f$ converges to f in $L^p(\mathbb{T})$.

Theorem 11.80. *The N^{th} partial Fourier sums $S_N f$ of a function $f \in L^p(\mathbb{T})$ converge to f in the L^p-norm as $N \to \infty$, for each p with $1 < p < \infty$.*

Proof. It follows from formula (11.14) that the partial sums S_N are uniformly (in N) bounded on $L^p(\mathbb{T})$, for each p with $1 < p < \infty$. To see this, take L^p-norms on both sides of (11.14), apply the triangle inequality in $L^p(\mathbb{T})$, and note that the last two terms in formula (11.14) tend to zero when $N \to \infty$ by the Riemann–Lebesgue Lemma. Furthermore, the Fourier coefficients of a function in $L^p(\mathbb{T})$ are uniformly bounded by the quantity $\|f\|_{L^p(\mathbb{T})}$. Next, use the boundedness properties of M_N and H_P as follows:

$$\begin{aligned}
&\|S_N f\|_{L^p(\mathbb{T})} \\
&\quad\le \frac{1}{2}\left[\|M_N H_P M_{-N} f\|_{L^p(\mathbb{T})} + \|M_{-N} H_P M_N f\|_{L^p(\mathbb{T})} + 2\|f\|_{L^p(\mathbb{T})}\right] \\
&\quad= \frac{1}{2}\left[\|H_P(M_{-N} f)\|_{L^p(\mathbb{T})} + \|H_P(M_N f)\|_{L^p(\mathbb{T})} + 2\|f\|_{L^p(\mathbb{T})}\right] \\
&\quad\le \frac{C_p}{2}\left[\|M_{-N} f\|_{L^p(\mathbb{T})} + \|M_N f\|_{L^p(\mathbb{T})} + 2\|f\|_{L^p(\mathbb{T})}\right] \\
&\quad= (C_p + 3/2)\|f\|_{L^p(\mathbb{T})}.
\end{aligned}$$

The bound $(C_p + 3/2)$ is independent of N, as claimed.

We know that the trigonometric polynomials are dense in $L^p(\mathbb{T})$. Given a trigonometric polynomial $f(\theta) = \sum_{|m|\le M} a_m e^{2\pi i m\theta}$, for $N \ge M$ we have $S_N f = f$. Thus, the partial sum operators S_N converge to the identity operator on a dense subset of $L^p(\mathbb{T})$. We can now deduce as we did to prove Theorem 9.41 in page 246, that

$$\lim_{N\to\infty} \|S_N f - f\|_p = 0. \qquad \square$$

Therefore, the convergence in $L^p(\mathbb{T})$ of the partial Fourier sums is a consequence of the boundedness of the periodic Hilbert transform on those spaces.

Exercise 11.81. Let $f \in L^p(\mathbb{R})$ (recall that $L^p(\mathbb{T}) \subset L^1(\mathbb{T})$). Use Hölder's inequality (11.8) to show that

$$|\widehat{f}(n)| \le (2\pi)^{-1/p}\|f\|_{L^p(\mathbb{T})}. \qquad \Diamond$$

Formula (11.14) also allows us to conclude that the periodic Hilbert transform cannot be bounded on $L^1(\mathbb{T})$, for if it were bounded that would imply the uniform boundedness on $L^1(\mathbb{T})$ of the partial Fourier sums $S_N f$, which is false.

Exercise 11.82. Show that there DOES NOT EXIST a constant $C > 0$ such that $\|S_N f\|_{L^1(\mathbb{T})} \le C\|f\|_{L^1(\mathbb{T})}$ holds for all $N > 0$. It may be useful to recall that the Dirichlet kernels D_N do not have uniformly bounded $L^1(\mathbb{T})$ norms, and moreover that $\|D_N\|_{L^1(\mathbb{T})} \sim \log N$; see Chapter 4, and Exercise 9.50. $\Diamond$

When M. Riesz proved the boundedness in $L^p(\mathbb{T})$ of the periodic Hilbert transform in 1927, the Marcinkiewicz interpolation theorem had not yet appeared, nor were the weak-type $(1,1)$ boundedness properties of the periodic Hilbert transform known. Using a complex-variables argument M. Riesz showed that the periodic Hilbert transform was bounded on $L^p(\mathbb{T})$ for even numbers $p = 2k$. Then by strong interpolation (Riesz–Thorin) which was already known, he obtained boundedness in the range $2 \leq p < \infty$. Finally duality took care of the range $1 < p < 2$. You can find this argument in [**Pin**, Section 3.3.1]. A variation of this argument was found by Cotlar; he showed how to obtain boundedness on $L^{2p}(\mathbb{T})$ given boundedness on $L^p(\mathbb{T})$. Therefore from the known boundedness on $L^2(\mathbb{T})$, one gets boundedness on $L^4(\mathbb{T})$, then on $L^8(\mathbb{T})$, $L^{16}(\mathbb{T})$, and so on. Again strong interpolation covers the range $2 \leq p < \infty$ and duality takes care of the rest. You can find this argument in [**Graf2**, Section 4.1.3].

Exercise 11.83. Show that if $f \in \mathcal{S}(\mathbb{R})$ then the partial Fourier integrals $S_R f$ can be decomposed in terms of the Hilbert transform H and the modulations $M_R f(x) := e^{2\pi i R x} f(x)$ defined on $\mathbb{R}$, as follows:

$$S_R = \frac{1}{2}[M_R H M_{-R} - M_{-R} H M_R].$$

This decomposition is analogous to formula (11.14), the one found for the partial Fourier sums. Use this decomposition and the boundedness on $L^p(\mathbb{R})$ of the Hilbert transform for $1 < p < \infty$, and the fact that for the modulation operators $\|M_R f\|_p = \|f\|_p$, to show that $S_R f$ converges to f in the L^p-norm. $\Diamond$

11.7.2. Connection to complex analysis. Consider a real-valued function $f \in L^2(\mathbb{R})$ and let $F(z)/2$ be its analytic extension to the upper half-plane $\mathbb{R}^2_+ = \{z = x + iy : y > 0\}$, suitably normalized. The analytic function $F(z)$ can be explicitly computed by means of the well known *Cauchy Integral Formula*[18]:

$$F(z) = \frac{1}{\pi i} \int_{\mathbb{R}} \frac{f(t)}{z - t}\, dt, \qquad z \in \mathbb{R}^2_+.$$

Notice the resemblance between the integral on the right-hand side and the Hilbert transform. No principal value is needed in the Cauchy

[18]Named after the French mathematician Augustin-Louis Cauchy (1789–1857).

Integral Formula, however, since the singularity is never achieved: $z \in \mathbb{R}^2_+$ cannot equal $t \in \mathbb{R}$. By separating the real and imaginary parts of the kernel, one can obtain explicit formulae for the real and imaginary parts of $F(z) = u(z) + iv(z)$ in terms of convolutions with the *Poisson* and *conjugate Poisson kernels* that we encountered in Section 7.6. Here are the formulae:

$$u(x+iy) = f * P_y(x), \qquad v(x+iy) = f * Q_y(x).$$

The function u is the *harmonic extension* of f to the upper half-plane, and the function v is the *harmonic conjugate* of u.

Exercise 11.84. (*The Poisson and Conjugate Poisson Kernels*) Show that the Poisson kernel $P_y(x)$ and the conjugate Poisson kernel $Q_y(x)$ are given by

$$P_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2}, \qquad Q_y(x) = \frac{1}{\pi}\frac{x}{x^2+y^2}.$$

You have already shown (Exercise 8.45) that

$$\widehat{Q_y}(\xi) = -i\ \mathrm{sgn}(\xi)\exp(-2\pi|y\xi|)$$

for each $y > 0$. Therefore, as $y \to 0$, $\widehat{Q_y}(\xi) \to -i\ \mathrm{sgn}(\xi)$. On the other hand, show that as $y \to 0$, $Q_y(x)$ approaches the principal value distribution p.v. $1/(\pi x)$, meaning that for all test functions $\phi \in \mathcal{S}(\mathbb{R})$,

$$\lim_{y\to 0}\int_{\mathbb{R}} Q_y(x)\phi(x)\,dx = \lim_{\varepsilon\to 0}\int_{|x|>\varepsilon}\frac{\phi(x)}{x}\,dx. \qquad \Diamond$$

The Poisson kernel is an approximation of the identity (Section 7.5). Therefore the limit as $y \to 0$ of $u = P_y * f$ is f, both in the L^2-sense and almost everywhere. On the other hand, as $y \to 0$, $v = Q_y * f$ approaches the Hilbert transform Hf in $L^2(\mathbb{R})$. Therefore, by continuity of the Fourier transform in $L^2(\mathbb{R})$, we conclude that

$$(Hf)^\wedge(\xi) = -i\,\mathrm{sgn}(\xi)\widehat{f}(\xi).$$

The connection between the Hilbert transform and the Poisson kernel is illustrated in Figure 11.8.

To sum up: the boundary values (on $\mathbb{R}$) of the harmonic conjugate $v(x,y) = Q_y * f(x)$ of the harmonic extension $u(x,y) = P_y * f(x)$ of f to the upper half-plane are given by the Hilbert transform Hf of f.

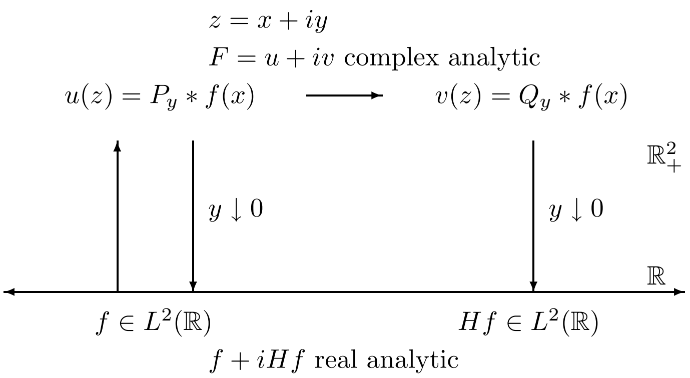

Figure 11.8. The Hilbert transform via the Poisson kernel.

An analogous result holds on the unit disk. By a limiting procedure on the unit disk, similar to the one described on the upper half-plane, one can conclude that the boundary values of the harmonic conjugate of the harmonic extension of a periodic, real-valued, continuously differentiable function f on $\mathbb{T}$ are given by

$$H_P f(\theta) = \text{p.v.}\frac{1}{\pi}\int_0^1 f(t)\cot\big(\pi(\theta - t)\big)\,dt. \tag{11.15}$$

Here we are identifying the function $f(\theta)$ with $f(z)$, for $z = e^{2\pi i\theta}$. The singularity at the diagonal is comparable to that of the Hilbert transform. This is the appropriate analogue of the Hilbert transform on the unit circle. The analogy is reinforced when we notice that on the Fourier side,

$$(H_P f)^{\wedge}(n) = -i\ \mathrm{sgn}(n)\widehat{f}(n). \tag{11.16}$$

This is how we defined the periodic Hilbert transform in Section 11.7.1.

Exercise 11.85. Verify that for $H_P f$ defined on the time-domain by equation (11.15), the Fourier coefficients are indeed given by equation (11.16). $\Diamond$

For more on the periodic Hilbert transform, and its connections to the conjugate function and functions analytic in the unit disc, see for example the books by Katznelson [**Kat**, Chapter III], Koosis [**Koo**, Chapter I.E], Torchinsky [**Tor**, Chapter III, Sections 5 and 6], and Pinsky [**Pin**, Section 3.3].

11.8. Looking further afield

In this book we have restricted ourselves to Fourier analysis and wavelets in one dimension (on $\mathbb{T}$, $\mathbb{R}$ or discrete one-dimensional versions of these spaces). There is also a rich theory in $\mathbb{R}^n$ of multiple Fourier series, Fourier transforms and wavelet transforms, with many fascinating applications. The reader may also like to investigate the Riesz transforms and their connections to partial differential equations. The Riesz transforms are natural extensions of the Hilbert transform to higher dimensions. See for example [**Graf2**, Chapter 4], or [**Duo**, Chapter 4].

We encourage the reader to start her/his excursion into multidimensional Fourier theory by reading the accounts in [**DM**, Sections 2.10, 2.11], [**Kör**, Chapter 79 or above], [**Pin**, Chapters 4 and 5], [**Pre**, Chapters 6, 7, 8], and [**SS1**, Chapter 6]. In these books you will find applications in many areas including physics, differential equations, probability, crystallography, random walks, Radon transforms, radioastronomy, and number theory.

Bibliography

[Abb] S. Abbott, *Understanding analysis*, Undergraduate Texts in Mathematics, Springer, 2001.

[AHMTT] P. Auscher, S. Hofmann, C. Muscalu, T. Tao and C. Thiele, *Carleson measures, trees, extrapolation and T(b) theorems*, Publ. Mat. **46** (2002), 257–325.

[Ari] J. Arias de Reyna, *Pointwise convergence of Fourier series*, Lecture Notes in Math. **1785**, Springer–Verlag, Berlin, 2002.

[Bae] A. Baernstein II, *Some sharp inequalities for conjugate functions*, Indiana Univ. Math. J. **27** (1978), 833–852.

[Bart1] R. G. Bartle, *The elements of integration and Lebesgue measure*, John Wiley & Sons, Inc., 1966.

[Bart2] R. G. Bartle, *A modern theory of integration*, Graduate Studies in Mathematics **32**, Amer. Math. Soc., Providence, RI, 2001.

[Bea] J. C. Bean, *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*, Jossey–Bass, San Francisco, CA, 2001.

[Bec] W. Beckner, *Inequalities in Fourier Analysis on* $\mathbb{R}^n$, Proceedings of the National Academy of Sciences **72** (1975), 638–641.

[BK] C. Bénéteau and D. Khavinson, *The isoperimetric inequality via approximation theory and free boundary problems*, Comput. Methods Funct. Theory **6** No. 2 (2006), 253–274.

[Bôc] M. Bôcher, *Introduction to the theory of Fourier's series*, Ann. of Math. **7** (1906), 81–152.

[Bre] D. M. Bressoud, *A radical approach to real analysis*, Math. Assoc. of Amer., Washington, DC, 1994.

[Bur] Frank E. Burris, *Matrix formulation of the discrete Hilbert transform*, IEEE Transactions on Circuits and Systems **22** No. 10 (1975), 836–838.

[Cal] A. Calderón, *Intermediate spaces and interpolation, the complex method*, Studia Math. **24** (1964), 113–190.

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

[Chu] F. R. K. Chung, *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics **92**, Amer. Math. Soc., Providence, RI, 1997.

[CMa] R. R. Coifman and M. Maggioni, *Diffusion wavelets*, Appl. Comput. Harmon. Anal. **21** (2006), 53–94.

[CMe] R. R. Coifman and Y. Meyer, *Remarques sur l'analyse de Fourier à fenêtre*, C. R. Acad. Sci. Paris, Série I **312** (1991), 259–261.

[CW] R. R. Coifman and V. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Trans. Information Theory **38** (1992), 713–718.

[CT] J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp. **19** (1965), 297–301.

[Cot] M. Cotlar, *A combinatorial inequality and its applications to L^2 spaces*, Rev. Mat. Cuyana **1** (1955), 41–55.

[CMP] D. Cruz-Uribe, J. M. Martell, C. Perez, *Sharp weighted estimates for approximating dyadic operators*, Electron. Res. Announc. Math. Sci. (to appear). Preprint, arXiv:1001.4724 (2010).

[Dav] B. Davis, *On the weak-type* $(1,1)$ *inequality for conjugate functions*, Proc. Amer. Math. Soc. **44** (1974), 307–311.

[Dau88] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure and Appl. Math. **41** (1988), 909–996.

[Dau92] I. Daubechies, *Ten lectures on wavelets*, CBMS–NSF Lecture Notes **61**, SIAM, Philadelphia, PA, 1992.

[DV] O. Dragičević and A. Volberg, *Sharp estimate of the Ahlfors–Beurling operator via averaging martingale transforms*, Mich. Math. J. **51** (2003), 415–435.

[Duo] J. Duoandikoetxea, *Fourier analysis*, English transl. by D. Cruz-Uribe, SFO, Graduate Studies in Mathematics **29**, Amer. Math. Soc., Providence, RI, 2001.

[DM] H. Dym and H. P. McKean, *Fourier series and integrals*, Probability and Mathematical Statistics **14**, Academic Press, New York–London, 1972.

[Eng] Shlomo Engelberg, *Edge detection using Fourier coefficients*, The American Mathematical Monthly **115** No. 6 (2008), 499–513.

[Fol] G. B. Folland, *Real analysis: Modern techniques and their applications*, Wiley, New York, NY, 1999.

[Fra] M. W. Frazier, *An introduction to wavelets through linear algebra*, Springer–Verlag, New York, NY, 1999.

[Gab] D. Gabor, *Theory of communication*, J. Inst. Electr. Eng., London **93** (III) (1946), 429–457.

[Gar] J. B. Garnett, *Bounded Analytic Functions*, Academic Press, Orlando, FL, 1981.

[Gib1] J. W. Gibbs, *Fourier's series*, letter in Nature **59** (1898), 200.

[Gib2] J. W. Gibbs, *Fourier's series*, letter in Nature **59** (1899), 606.

[GS] D. Gottlieb and C.-W. Shu, *On the Gibbs phenomenon and its resolution*, SIAM Rev. **39** No. 4 (1997), 644–668.

[Grab] J. V. Grabiner, *Who gave you the epsilon? Cauchy and the origins of rigorous calculus*, Amer. Math. Monthly **90** No. 3 (1983), 185–194.

[Graf1] L. Grafakos, *An elementary proof of the square summability of the discrete Hilbert transform*, Amer. Math. Monthly **101** No. 5 (1994), 456–458.

[Graf2] L. Grafakos, *Classical Fourier analysis, second edition*, Graduate Texts in Mathematics **249**, Springer, New York, 2008.

[Graf3] L. Grafakos, *Modern Fourier analysis, second edition*, Graduate Texts in Mathematics **250**, Springer, New York, 2009.

[GM] L. Grafakos and S. Montgomery-Smith, *Best constants for uncentered maximal functions*, Bull. London Math. Soc. **29** (1997), 60–64.

[Grat] I. Grattan-Guinness, *The development of the foundations of mathematical analysis from Euler to Riemann*, MIT Press, Cambridge, MA, 1970.

[Gre1] B. Green, *Long arithmetic progressions of primes*, in Analytic number theory, Clay Math. Proc. **7**, Amer. Math. Soc., Providence, RI, 2007, 149–167.

[Gre2] B. Green, *Long arithmetic progressions of primes*, Gaz. Math. No. 112 (2007), 26–27.

[Haa] A. Haar, *Zur Theorie der orthogonalen Funktionensysteme*, Math. Annal. **69** (1910), 331–371.

[HL] G. H. Hardy and J. E. Littlewood, *A maximal theorem with function theoretic applications*, Acta Math. **54** (1930), 81–116.

[HW] E. Hernandez and G. Weiss, *A first course on wavelets*, CRC Press, Boca Raton, FL, 1996.

[Hig] N. J. Higham, *Handbook of Writing for the Mathematical Sciences*, SIAM, Philadelphia, PA, 1998.

[HT] Holshneider and P. Tchamitchian, *Pointwise analysis of Riemann's nowhere differentiable function* Invent. Math. **105** (1991), 157–176.

[Hun] R. Hunt, *On the convergence of Fourier series*, Orthogonal expansions and their continuous analogues (Proc. Conf., Edwardsville, IL, 1967), Southern Illinois Univ. Press, Carbondale, IL, 1968, 235–255.

[Hyt] T. Hytönen, *On Petermichl's dyadic shift and the Hilbert transform*, Comptes Rendus Mathématique, Académie des Sciences, Paris **346** (21–22) (2008), 1133–1136.

[Jaf1] S. Jaffard, *Local behavior of Riemann's function*, Contemp. Mat. **189** (1995), 278–307.

[Jaf2] S. Jaffard, *Old friends revisited: The multifractal nature of some classical functions.* J. Fourier Anal. Appl. **3** (1997), 1–22.

[JMR] S. Jaffard, Y. Meyer, and R. D. Ryan, *Wavelets: Tools for science and technology*, revised edition, SIAM, Philadelphia, PA, 2001.

[Jer] A. J. Jerri, *The Gibbs phenomenon in Fourier analysis, Splines and wavelet approximations*, Kluwer Academic Publishers, Boston, MA, 1998.

[KK] J.-P. Kahane and Y. Katznelson, *Sur les ensembles de divergence des séries trigonométriques*, Studia Math. **26** (1966), 305–306.

[KL] J.-P. Kahane and P.-G. Lemarié-Rieusset, *Fourier series and wavelets*, Gordon and Breach Publishers, 1995.

[Kat] Y. Katznelson, *Introduction to harmonic analysis*, Dover Publications Inc., 1976.

[Kel] S. E. Kelly, *Gibbs phenomenon for wavelets*, Applied and Computational Harmonic Analysis **3** (1996), 72–81.

[Kol] A. Kolmogorov, *Une série de Fourier–Lebesgue divergente partout*, C. R. Acad. Sci. Paris **183** (1926), 1327–1328.

[Koo] P. Koosis, *Introduction to the H_p spaces*, Cambridge University Press, Cambridge, 1980.

[Kör] T. W. Körner, *Fourier analysis*, Cambridge University Press, Cambridge, 1988.

[Kra] B. Kra, *The Green–Tao theorem on arithmetic progressions in the primes: an ergodic point of view*, Bull. Amer. Math. Soc. (N.S.) **43** No. 1 (2006), 3–23.

[Kran] S. G. Krantz, *A panorama of harmonic analysis*, The Carus Mathematical Monographs **27**, Math. Assoc. Amer., Providence, RI, 1999.

[Lac] M. Lacey, *Haar Shifts, Commutators, and Hankel Operators*, posted at arXiv:0806.1650v1. To appear in Conference Proceedings, CIMPA–UNESCO conference on Real Analysis and its Applications, La Falda Argentina, June 2008.

[LPR] M. Lacey, S. Petermichl, and M. Reguera-Rodriguez, *Sharp A2 inequality for Haar shift operators*, Math. Ann. (to appear). Preprint, arXiv:0906.1941 (2009).

[Ler1] A. Lerner, *An elementary approach to several results on the Hardy–Littlewood maximal operator*, Proc. Amer. Math. Soc. **136** No. 8 (2008), 2829–2833.

[Ler2] A. Lerner, *A pointwise estimate for local sharp maximal function with applications to singular integrals*. To appear Bull. London Math. Soc.

[LL] E. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics **14**, Amer. Math. Soc., Providence, RI, 2001.

[LC] G.M. Livadas and A. G. Constantinides, *Image edge detection and segmentation based on the Hilbert transform*, in Proc. of ICASSP-88, vol. 2 (1988), 1152–1155.

[Mall89] S. Mallat, *Multiresolution approximations and wavelet orthonormal bases for $L^2(\mathbb{R})$*, Trans. Amer. Math. Soc. **315** (1989), 69–87.

[Mall98] S. Mallat, *A wavelet tour of signal processing*, Academic Press, San Diego, CA, 1998.

[Malv] H. Malvar, *Lapped transforms for efficient transform/subband coding*, IEEE Trans. Acoust. Speech Signal Process **38** (1990), 969–978.

[Mel] A. Melas, *The best constant for the centered Hardy–Littlewood maximal inequality*, Ann. of Math. **157** (2003), 647–688.

[Mic] A. A. Michelson, *Fourier's series*, Letter in Nature **58** (1898), 544–545.

[MS] A. A. Michelson and S. W. Stratton, *A new harmonic analyser*, Philosophical Magazine **45** (1898), 85–91.

[MP] M. Mohlenkamp and M. C. Pereyra, *Wavelets, their friends, and what they can do for you*, EMS Series of Lectures in Mathematics, European Mathematical Society Publishing House, Zurich, 2008.

[Mun] J. R. Munkres, *Topology: A first course*, Prentice Hall, Englewood Cliffs, NJ, 1975.

[Nat] M. Nathanson, *Elementary Methods in Number Theory*, Graduate Texts in Mathematics **195**, Springer, New York, NY, 2000.

[NOTh] F. Nazarov, R. Oberlin, C. Thiele, *A Calderon Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain*, to appear MRL, arxiv:0912.3010

[NTV] F. Nazarov, S. Treil, A. Volberg, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** No. 3 (2008), 583–597.

[Per] M. C. Pereyra, *Lecture notes on dyadic harmonic analysis*, Contemporary Mathematics **289** (2001), Amer. Math. Soc., Ch. I, 1–61.

[Pet] S. Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C.R. Acad. Sci. Paris, Sér. I Math. **330** (2000), 455–460.

[PTV] S. Petermichl, S. Treil, and A. Volberg, *Why the Riesz transforms are averages of the dyadic shifts*, Publ. Mat. Barc., Spec. Vol. (2002), 209–228.

[Pin] M. A. Pinsky, *Introduction to Fourier analysis and wavelets*, The Brooks/Cole Series in Advanced Mathematics (Paul J. Sally, Jr, ed.), Brooks/Cole, 2001.

[Pre] E. Prestini, *The evolution of applied harmonic analysis: Models of the real world*, Birkhäuser, Boston, MA, 2004.

[RS] M. Reed and B. Simon, *Methods of modern mathematical physics, Vol. II: Fourier analysis, self-adjointness*, Academic Press, 1975.

[Roy] H. L. Royden, *Real analysis*, 2nd ed., MacMillan, New York, NY, 1968.

[Rud] W. Rudin, *Principles of mathematical analysis*, McGraw–Hill, New York, NY, 1976.

[Sad] C. Sadosky, *Interpolation of operators and singular integrals*, Marcel Dekker, Inc., New York, NY, 1979.

[Sch] M. Schechter, *Principles of functional analysis*, Graduate Studies in Mathematics **36**, 2nd ed., Amer. Math. Soc., Providence, RI, 2002.

[Sha] C. E. Shannon, *Communications in the presence of noise*, Proc. Inst. Radio Eng. **37** (1949), 10–21.

[SV] H.-T. Shim and H. Volkmer, *On the Gibbs phenomenon for wavelet expansions*, Journal of Approximation Theory **84** (1996), 74–95.

[Spi] M. Spiegel, *Schaum's outline of Fourier analysis with applications to boundary value problems*, McGraw–Hill, New York, NY, 1974.

[Ste] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, NJ, 1970.

[SS1] E. M. Stein and R. Shakarchi, *Fourier analysis: an introduction*, Princeton Lectures in Analysis I, Princeton Univ. Press, Princeton, NJ, 2003.

[SS2] E. M. Stein and R. Shakarchi, *Real analysis: Measure theory, integration, and Hilbert spaces*, Princeton Lectures in Analysis III, Princeton Univ. Press, Princeton, NJ, 2005.

[SW1] E. M. Stein and G. Weiss, *An extension of a theorem of Marcinkiewicz and some of its applications*, J. Math. Mech. **8** (1959), 263–264.

[SW2] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Univ. Press, Princeton, NJ, 1971.

[SN] G. Strang and T. Nguyen, *Wavelets and filter banks*, Wellesley–Cambridge Press, Wellesley, MA, 1996.

[Stra] G. Strang, *Linear algebra and its applications*, Saunders College Publishing, third edition, 1988.

[Str] R. S. Strichartz, *The way of analysis*, Jones & Bartlett Publishers, revised edition, 2000.

[Str2] R. S. Strichartz, *A guide to distribution theory and Fourier transforms*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1994.

[Str3] R. S. Strichartz, *Gibbs' Phenomenon and Arclength*, J. Fourier Anal. and Appl. **6** Issue 5 (2000), 533–536.

[Tao1] T. Tao, *Analysis I*, Text and Readings in Mathematics **37**, Hindustan Book Agency, India, 2006.

[Tao2] T. Tao, *Analysis II*, Text and Readings in Mathematics **37**, Hindustan Book Agency, India, 2006.

[Tao3] T. Tao, *An uncertainty principle for cyclic groups of prime order*, Math. Research Letters **12** (2005), 121–127.

[Tao4] T. Tao, *Arithmetic progressions and the primes*, Collect. Math., Vol. Extra (2006), 37–88.

[Ter] A. Terras, *Fourier analysis on finite groups and applications*, London Math. Soc. Student Texts **43**, Cambridge University Press, Cambridge, 1999.

[Tor] A. Torchinsky, *Real-variable methods in harmonic analysis*, Pure and Applied Mathematics **123**, Academic press, Inc., San Diego, CA, 1986.

[Tre] A. Treibergs, *Inequalities that imply the isoperimetric inequality*, online at `http://www.math.utah.edu/$\sim$treiberg/isoperim/isop.pdf`.

[Vag] A. Vagharshakyan, *Recovering singular integrals from Haar shifts*, Preprint (2009) posted at: arXiv:0911.4968v1.

[VK] M. Vetterli and J. Kovačević, *Wavelets and subband coding*, Prentice Hall, Upper Saddle River, NJ, 1995.

[Wil] H. Wilbraham, *On a certain periodic function*, Cambridge & Dublin Math. J. **3** (1848), 198–201.

[Wic] V. Wickerhauser, *Adapted wavelet analysis from theory to software*, A. K. Peters, Wellesley, 1994.

[Wie] N. Wiener, *The ergodic theorem*, Duke Math. J. **5** (1939), 1–18.

[Wieg] J. Wiegert, *The Sagacity of Circles: A History of the Isoperimetric Problem*, online at `http://mathdl.maa.org/mathDL/23/` (write keyword: sagacity)

[Woj1] P. Wojtaszczyk, *A mathematical introduction to wavelets*, London Math. Soc. Student Texts **37**, Cambridge University Press, Cambridge, 2003.

[Woj2] P. Wojtaszczyk, *Banach spaces for analysts*, Cambridge Studies in Advanced Mathematics **25**, Cambridge University Press, Cambridge, 1991.

[Zyg1] A. Zygmund, *Trigonometric Series*, 2nd ed., Cambridge University Press, Cambridge, 1959.

[Zyg2] A. Zygmund, *Intégrales Singuliéres*, Lecture Notes in Math. **204**, Springer–Verlag, Berlin, 1971.