Harmonic Analysis

W. Schlag

Contents

CHAPTER 1

Fourier Series: Convergence and Summability

We begin this book with one of the most basic objects in analysis, namely the Fourier series associated with a function or a measure on the circle. To be specific, let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the one-dimensional torus (in other words, the circle). We consider various function spaces on the torus $\mathbb{T}$, namely the space of continuous functions $C(\mathbb{T})$, the space of Hölder continuous functions $C^\alpha(\mathbb{T})$ where $0 < \alpha \leqslant 1$, and the Lebesgue spaces $L^p(\mathbb{T})$ where $1 \leqslant p \leqslant \infty$. The space of complex Borel measures on $\mathbb{T}$ will be denoted by $\mathcal{M}(\mathbb{T})$. Any $\mu \in \mathcal{M}(\mathbb{T})$ has associated with it a *Fourier series*

$$\mu \sim \sum_{n=-\infty}^{\infty} \hat{\mu}(n) e(nx) \tag{1.1}$$

where we let $e(x) := e^{2\pi i x}$ and

$$\hat{\mu}(n) := \int_0^1 e(-nx)\,\mu(dx) = \int_{\mathbb{T}} e(-nx)\,\mu(dx)$$

The symbol $\sim$ in (1.1) is formal, and simply means that the series on the right-hand side is associated with μ. If $\mu(dx) = f(x)\,dx$ where $f \in L^1(\mathbb{T})$, then we also write $\hat{f}(n)$ instead of $\hat{\mu}(n)$.

Of course, the central question which we wish to explore in this chapter is when μ equals the right-hand side in (1.1) or represents f in a suitable sense. Note that if we start from a *trigonometric polynomial*

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e(nx)$$

where all but finitely many a_n are zero, then we see that

$$\hat{f}(n) = a_n \qquad \forall n \in \mathbb{Z} \tag{1.2}$$

In other words, we have the pointwise equality

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e_n(x)$$

with $e_n(x) := e(nx)$. What enters (1.2) is of course the basic *orthogonality relation*

$$\int_{\mathbb{T}} e_n(x)\overline{e_m(x)}\,dx = \delta_0(n-m) \tag{1.3}$$

where $\delta_0(j) = 1$ if $j = 0$ and $\delta_0(j) = 0$ otherwise.

It is therefore natural to explore the question of *convergence* of the Fourier series for more general functions. Of course, the sense of convergence of the infinite series needs to be specified before this fundamental question can be answered. It is fair to say that much of modern analysis (including functional analysis) arose out of the struggle with this question. For example, the notion of Lebesgue integral was developed in order to overcome deficiencies of the older Riemannian definition of the integral which had been revealed in the study of Fourier series. The reader will also note the recurring theme of convergence of Fourier series throughout this book.

It is natural to start from the most classical notion of convergence, namely that of *pointwise convergence* in case the measure μ is of the form $\mu(dx) = f(x)\,dx$ with $f(x)$ continuous or better. The partial sums of $f \in L^1(\mathbb{T})$ are defined as

$$\begin{aligned} S_N f(x) &= \sum_{n=-N}^{N} \hat{f}(n) e(nx) = \sum_{n=-N}^{N} \int_{\mathbb{T}} e(-ny) f(y)\,dy\; e(nx) \\ &= \int_{\mathbb{T}} \sum_{n=-N}^{N} e(n(x-y)) f(y)\,dy = \int_{\mathbb{T}} D_N(x-y) f(y)\,dy \end{aligned}$$

where $D_N(x) = \sum_{n=-N}^{N} e(nx)$ is the *Dirichlet kernel*. In other words, we have shown that the partial sum operator S_N is given by convolution with the *Dirichlet kernel D_N*:

$$S_N f(x) = (D_N * f)(x) \tag{1.4}$$

In order to understand basic properties of this convolution, we first sum the geometric series defining $D_N(x)$ so as to have an explicit expression for the Dirichlet kernel.

Exercise 1.1. Verify that for each integer $N \geqslant 0$

$$D_N(x) = \frac{\sin((2N+1)\pi x)}{\sin(\pi x)} \tag{1.5}$$

and draw the graph of D_N for several different values of N, say $N = 2$ and $N = 5$. Prove the bound

$$|D_N(x)| \leqslant C \min\Big(N, \frac{1}{|x|}\Big) \tag{1.6}$$

for all $N \geqslant 1$ and some absolute constant C. Finally, prove the bound

$$C^{-1} \log N \leqslant \|D_N\|_{L^1(\mathbb{T})} \leqslant C \log N \tag{1.7}$$

for all $N \geqslant 2$ where C is another absolute constant.

The growth of the bound in (1.7) as well as the oscillatory nature of D_N as given by (1.5) may indicate that it is in general very delicate to understand pointwise or almost everywhere convergence properties of $S_N f$. This will become more clear as we develop the theory.

In order to study (1.4) we need to establish some basic properties of the convolution of two functions f, g on $\mathbb{T}$. If f and g are continuous, say, then define

$$(f * g)(x) := \int_{\mathbb{T}} f(x-y)g(y)\,dy = \int_{\mathbb{T}} g(x-y)f(y)\,dy \tag{1.8}$$

It is helpful to think of $f * g$ as an average of translates of f by the measure $g(y)\,dy$ or vice versa. In particular, convolution commutes with the translation operator τ_z which is defined for any $z \in \mathbb{T}$ by the action on functions: $(\tau_z f)(x) = f(x-z)$. Indeed, one immediately verifies that

$$\tau_z(f * g) = (\tau_z f) * g = f * (\tau_z g) \tag{1.9}$$

In passing, we mention the important relation between the Fourier transform and translations:

$$\widehat{(\tau_z \mu)}(n) = e(-zn)\hat{\mu}(n) \quad \forall\, n \in \mathbb{Z}$$

In what follows, we abbreviate *almost everywhere* or *almost every* by *a.e.*

Lemma 1.1. *The operation of convolution as defined above satisfies the following properties:*

(1) *Let $f, g \in L^1(\mathbb{T})$. Then for a.e. $x \in \mathbb{T}$ one has that $f(x-y)g(y)$ is L^1 in y. Thus, the integral in* (1.8) *is well-defined for a.e. $x \in \mathbb{T}$ (but not necessarily for every x), and the bound $\|f * g\|_1 \leqslant \|f\|_1\|g\|_1$ holds.*
(2) *More generally, $\|f * g\|_p \leqslant \|f\|_p\|g\|_1$ for all $1 \leqslant p \leqslant \infty$, $f \in L^p$, $g \in L^1$. This is called Young's inequality.*
(3) *If $f \in C(\mathbb{T})$, $\mu \in \mathcal{M}(\mathbb{T})$ then $f * \mu$ is well-defined. Show that, for $1 \leqslant p \leqslant \infty$,*

$$\|f * \mu\|_p \leqslant \|f\|_p\|\mu\|$$

which allows one to extend $f * \mu$ to arbitrary $f \in L^p$.
(4) *If $f \in L^p(\mathbb{T})$ and $g \in L^{p'}(\mathbb{T})$ where $1 \leqslant p \leqslant \infty$, and $\frac{1}{p} + \frac{1}{p'} = 1$ then $f * g$, originally defined only a.e., extends to a continuous function on $\mathbb{T}$ and*

$$\|f * g\|_\infty \leqslant \|f\|_p\|g\|_{p'} \tag{1.10}$$

(5) *For $f, g \in L^1(\mathbb{T})$ show that for all $n \in \mathbb{Z}$*

$$\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n)$$

Proof. (1) is an immediate consequence of Fubini's theorem since $f(x-y)g(y)$ is jointly measurable on $\mathbb{T} \times \mathbb{T}$ and belongs to $L^1(\mathbb{T} \times \mathbb{T})$. (2) can be obtained by interpolating between the cases $p = 1$ just convered and the easy bound for $p = \infty$. Alternatively, one can use Minkowski's inequality:

$$\|f * g\|_p \leqslant \int_{\mathbb{T}} \|f(\cdot - y)\|_p |g(y)|\,dy \leqslant \|f\|_p\|g\|_1$$

which also implies (3). The bound (1.10) is just Hölder's inequality. Part (4) follows from the density of $C(\mathbb{T})$ in $L^p(\mathbb{T})$ for $1 \leqslant p < \infty$ and the translation invariance (1.9). Indeed, one verifies from the uniform continuity of functions in $C(\mathbb{T})$ and (1.9) that $f * \mu \in C(\mathbb{T})$ for any $f \in C(\mathbb{T})$ and $\mu \in \mathcal{M}(\mathbb{T})$. Since uniform limits of continuous functions are continuous, (4) now follows from the aforementioned density of $C(\mathbb{T})$ and (1.10).

Finally, (5) is a consequence of Fubini's theorem and the character property of the exponentials $e(n(x+y)) = e(nx)e(ny)$. □

The following exercise introduces the convolution on the Fourier side in the context of the largest class of functions where the respective series are absolutely convergent. This class of functions is necessarily a subalgebra of $C(\mathbb{T})$ called the *Wiener algebra*.

Exercise 1.2. Let $\mu \in \mathcal{M}(\mathbb{T})$ have the property that

$$\sum_{n\in\mathbb{Z}} |\hat{\mu}(n)| < \infty \tag{1.11}$$

Show that $\mu(dx) = f(x)\,dx$ where $f \in C(\mathbb{T})$. Denote the space of all measures with this property by $\mathbb{A}(\mathbb{T})$ and identify these measure with their respective densities. Show that $\mathbb{A}(\mathbb{T})$ is an algebra under multiplication, and that

$$\widehat{fg}(n) = \sum_{m\in\mathbb{Z}} \hat{f}(m)\hat{g}(n-m) \qquad \forall\, n \in \mathbb{Z}$$

where the sum on the right-hand side is absolutely convergent for every $n \in \mathbb{Z}$, and itself is absolutely convergent over all n. Moreover, $\|f * g\|_{\mathbb{A}} \leqslant \|f\|_{\mathbb{A}}\|g\|_{\mathbb{A}}$ where $\|f\|_{\mathbb{A}} := \|\hat{f}\|_{\ell^1}$. Finally, verify that if $f, g \in L^2(\mathbb{T})$, then $f * g \in \mathbb{A}(\mathbb{T})$.

Note that the Wiener algebra has a unit, namely the constant function 1. It is clear that if f has an inverse in $\mathbb{A}(\mathbb{T})$, then it is $\frac{1}{f}$ which in particular require that $f \neq 0$ everywhere on $\mathbb{T}$. It is a remarkable theorem due to Norbert Wiener that the converse here holds, too. I.e., if $f \in \mathbb{A}(\mathbb{T})$ does not vanish anywhere on $\mathbb{T}$, then $\frac{1}{f} \in \mathbb{A}(\mathbb{T})$.

One of the most basic as well as oldest results on the pointwise convergence of Fourier series is the following theorem. We shall see later that if *fails* for functions which are merely continuous.

Theorem 1.2. *If $f \in C^\alpha(\mathbb{T})$ with $0 < \alpha \leqslant 1$, then $\|S_N f - f\|_\infty \to 0$ as $N \to \infty$.*

Proof. One has, with $\delta \in (0, \frac{1}{2})$ to be determined,

$$\begin{aligned} S_N f(x) - f(x) &= \int_0^1 (f(x-y) - f(x))D_N(y)\,dy \\ &= \int_{|y|\leqslant\delta} (f(x-y) - f(x))D_N(y)\,dy \\ &\quad + \int_{\frac{1}{2}>|y|>\delta} (f(x-y) - f(x))D_N(y)\,dy \end{aligned} \tag{1.12}$$

Here one exploits the fact that

$$\int_{\mathbb{T}} D_N(y)\,dy = 1$$

We now use the bound from (1.6), i.e.,

$$|D_N(y)| \leqslant C \min\Big(N, \frac{1}{|y|}\Big)$$

Here and in what follows, C will denote a numerical constant that can change from line to line. The first integral in (1.12) can be estimated as follows

$$\int_{|y|\leqslant\delta} |f(x)-f(x-y)|\frac{1}{|y|}\,dy \leqslant [f]_\alpha \int_{|y|\leqslant\delta} |y|^{\alpha-1}\,dy \leqslant C[f]_\alpha\delta^\alpha \tag{1.13}$$

with the usual C^α semi-norm:

$$[f]_\alpha = \sup_{x,y}\frac{|f(x)-f(x-y)|}{|y|^\alpha}$$

To bound the second term in (1.12) one needs to invoke the oscillation of $D_N(y)$. In fact,

$$\begin{aligned}
B &:= \int_{\frac12>|y|>\delta} (f(x-y)-f(x))D_N(y)\,dy = \\
&= \int_{\frac12>|y|>\delta} \frac{f(x-y)-f(x)}{\sin(\pi y)}\sin((2N+1)\pi y)\,dy \\
&= -\int_{\frac12>|y|>\delta} h_x(y)\sin\Big((2N+1)\pi(y+\frac{1}{2N+1})\Big)\,dy
\end{aligned}$$

where $h_x(y) := \frac{f(x-y)-f(x)}{\sin(\pi y)}$.

Therefore, with all integrals being understood to be in the interval $(-\frac12,\frac12)$,

$$\begin{aligned}
2B &= \int_{|y|>\delta} h_x(y)\sin((2N+1)\pi y)\,dy \\
&\quad - \int_{|y-\frac{1}{2N+1}|>\delta} h_x\big(y-\frac{1}{2N+1}\big)\sin((2N+1)\pi y)\,dy \\
&= \int_{|y|>\delta}\Big(h_x(y)-h_x(y-\frac{1}{2N+1})\Big)\sin((2N+1)\pi y)\,dy \\
&\quad - \int_{[-\delta,-\delta+\frac{1}{2N+1}]} h_x\big(y-\frac{1}{2N+1}\big)\sin((2N+1)\pi y)\,dy \\
&\quad + \int_{[\delta,\delta+\frac{1}{2N+1}]} h_x\big(y-\frac{1}{2N+1}\big)\sin((2N+1)\pi y)\,dy
\end{aligned}$$

These integrals are estimated by putting absolute values inside. To do so we use the bounds

$$|h_x(y)| < C\frac{\|f\|_\infty}{\delta}$$

$$|h_x(y)-h_x(y+\tau)| < C\left(\frac{|\tau|^\alpha[f]_\alpha}{\delta}+\frac{\|f\|_\infty}{\delta^2}|\tau|\right)$$

if $|y|>\delta>2\tau$.

In view of the preceding one checks

$$|B| \leqslant C\left(\frac{N^{-\alpha}[f]_\alpha}{\delta} + \frac{N^{-1}\|f\|_\infty}{\delta^2}\right) \tag{1.14}$$

provided $\delta > \frac{1}{N}$. Choosing $\delta = N^{-\alpha/2}$ one concludes from (1.12), (1.13), and (1.14) that

$$|(S_N f)(x) - f(x)| \leqslant C\left(N^{-\alpha^2/2} + N^{-\alpha/2} + N^{-1+\alpha}\right) \tag{1.15}$$

for any function with $\|f\|_\infty + [f]_\alpha \leqslant 1$, which proves the theorem. □

The reader is invited to optimize the rate of decay that was derived in (1.15). In other words, the challenge is to obtain the largest $\beta > 0$ in terms of α so that the bound in (1.15) becomes $CN^{-\beta}$ for any f with $\|f\|_\infty + [f]_\alpha \leqslant 1$.

The difficulty with the Dirichlet kernel, such as its slow $\frac{1}{|x|}$-decay, can be regarded as a result of the "discontinuity" of $\widehat{D_N} = \chi_{[-N,N]}$; in fact, this indicator function on the lattice $\mathbb{Z}$ jumps at $\pm N$. Therefore, we may hope to obtain a kernel which is easier to analyze — in a sense that will be made precise by means of the notion of *approximate identity* below — by substituting D_N with a suitable average whose Fourier transform does not exhibit such jumps.

One elementary way of carrying this out is given by the Cesàro means, i.e.,

$$\sigma_N f := \frac{1}{N}\sum_{n=0}^{N-1} S_n f$$

Setting $K_N := \frac{1}{N}\sum_{n=0}^{N-1} D_n$, which is called the *Fejér kernel*, one therefore has $\sigma_N f = K_N * f$.

EXERCISE 1.3. Let K_N be the Fejér kernel with N a positive integer.

- Verify that $\widehat{K}_N$ looks like a triangle, i.e., for all $n \in \mathbb{Z}$

$$\widehat{K}_N(n) = \left(1 - \frac{|n|}{N}\right)^+ \tag{1.16}$$

- Show that

$$K_N(x) = \frac{1}{N}\left(\frac{\sin(N\pi x)}{\sin(\pi x)}\right)^2 \tag{1.17}$$

- Conclude that

$$0 \leqslant K_N(x) \leqslant C\,N^{-1}\min\left(N^2, x^{-2}\right) \tag{1.18}$$

We remark that the square and thus the positivity in (1.17) are not entirely surprising, since the triangle in (1.16) can be written as the convolution of two rectangles (convolution is now on the level of the Fourier coefficients on the lattice $\mathbb{Z}$). Therefore, we should expect K_N to look like the square of a version of D_M with M about half the size of N.

The properties established in Exercise 1.3 ensure that K_N are what one calls an *approximate identity*.

Definition 1.3. $\{\Phi_N\}_{N=1}^{\infty} \subset L^{\infty}(\mathbb{T})$ form an approximate identity provided

A1) $\int_0^1 \Phi_N(x)\,dx = 1$ for all N
A2) $\sup_N \int_0^1 |\Phi_N(x)|\,dx < \infty$
A3) for all $\delta > 0$ one has $\int_{|x|>\delta} |\Phi_N(x)|\,dx \to 0$ as $N \to \infty$.

The name here derives from the fact that $\Phi_N * f \to f$ as $N \to \infty$ in any reasonable sense, see Proposition 1.5. In other words, $\Phi_N \rightharpoonup \delta_0$ in the weak-$*$ sense of measures. Clearly, the so-called *box kernels*

$$\Phi_N(x) = N\chi_{[|x|N<\frac{1}{2}]}, \quad N \geqslant 1$$

satisfy A1)–A3) and can be considered as the most basic example of an approximate identity. Note that $\{D_N\}_{N\geqslant 1}$ are *not* an approximate identity. Finally, we remark that Definition 1.3 has nothing to do with the torus $\mathbb{T}$, it applies equally well to the line $\mathbb{R}$, tori $\mathbb{T}^d$, or Euclidean spaces $\mathbb{R}^d$.

Next, we verify that K_N belong to this class.

Lemma 1.4. *The Fejér kernels $\{K_N\}_{N=1}^{\infty}$ form an approximate identity.*

Proof. We clearly have $\int_0^1 K_N(x)\,dx = 1$ and $K_N(x) \geqslant 0$ so that A1) and A2) hold. A3) follows from the bound (1.18). ◻

Now we establish the basic convergence property of such families.

Proposition 1.5. *For any approximate identity $\{\Phi_N\}_{N=1}^{\infty}$ one has*

(1) *If $f \in C(\mathbb{T})$, then $\|\Phi_N * f - f\|_\infty \to 0$ as $N \to \infty$*
(2) *If $f \in L^p(\mathbb{T})$ where $1 \leqslant p < \infty$, then $\|\Phi_N * f - f\|_p \to 0$ as $N \to \infty$*
(3) *For any measure $\mu \in \mathcal{M}(\mathbb{T})$, one has*

$$\Phi_N * \mu \rightharpoonup \mu \qquad N \to \infty$$

in the weak-$*$ sense.

Proof. We begin with the uniform convergence. Since $\mathbb{T}$ is compact, f is uniformly continuous. Given $\varepsilon > 0$, let $\delta > 0$ be such that

$$\sup_x \sup_{|y|<\delta} |f(x-y) - f(x)| < \varepsilon$$

Then, by A1)–A3),

$$\begin{aligned}
|(\Phi_N * f)(x) - f(x)| &= \Big| \int_{\mathbb{T}} (f(x-y) - f(x))\Phi_N(y)\,dy \Big| \\
&\leqslant \sup_{x\in\mathbb{T}} \sup_{|y|<\delta} |f(x-y) - f(x)| \int_{\mathbb{T}} |\Phi_N(t)|\,dt + \int_{|y|\geqslant\delta} |\Phi_N(y)| 2\|f\|_\infty\,dy \\
&< C\varepsilon
\end{aligned}$$

provided N is large.

Fix any $f \in L^p(\mathbb{T})$ and let $g \in C(\mathbb{T})$ be such that $\|f - g\|_p < \varepsilon$. Then

$$\begin{aligned}\|\Phi_N * f - f\|_p &\leqslant \|\Phi_N * (f-g)\|_p + \|f-g\|_p + \|\Phi_N * g - g\|_p \\ &\leqslant \big(\sup_N \|\Phi_N\|_1 + 1\big)\|f-g\|_p + \|\Phi_N * g - g\|_\infty\end{aligned}$$

where we have used Young's inequality, see Lemma 1.1, to obtain the first term on the right-hand side. Using A2), the assumption on g, and Young's inequality finishes the proof.

The final statement is immediate from (1) and duality. □

This simple convergence result applied to the Fejér kernels implies some basic analytical properties. A *trigonometric polynomial* is a series $\sum_n a_n e(nx)$ in which only finitely many a_n are non-zero. Clearly, this class forms a vector space over the complex numbers. In fact, they form an algebra under both multiplication and convolution.

Corollary 1.6. *The exponential family $\{e(nx)\}_{n\in\mathbb{Z}}$ satisfies the following properties:*

(1) *The trigonometric polynomials are dense in $C(\mathbb{T})$ in the uniform topology, and in $L^p(\mathbb{T})$ for any $1 \leqslant p < \infty$.*
(2) *For any $f \in L^2(\mathbb{T})$*

$$\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$$

This is called the Plancherel theorem.
(3) *The exponentials $\{e(nx)\}_{n\in\mathbb{Z}}$ form a complete orthonormal basis in $L^2(\mathbb{T})$.*
(4) *For all $f, g \in L^2(\mathbb{T})$ one has* Parseval's identity

$$\int_{\mathbb{T}} f(x)\bar{g}(x)\,dx = \sum_{n\in\mathbb{Z}} \hat{f}(n)\overline{\hat{g}(n)}$$

Proof. By Lemma 1.4, $\{K_N\}_{N=1}^\infty$ form an approximate identity and Proposition 1.5 applies. Since $\sigma_N f = K_N * f$ is a trigonometric polynomial, we are done with (1).

Properties (2)–(4) are equivalent by basic Hilbert space theory. We begin from the orthogonality relations (1.3). One thus has Bessel's inequality

$$\sum |\hat{f}(n)|^2 \leqslant \|f\|_2^2$$

and equality is equivalent to the linear span of $\{e_n\}$ being dense in $L^2(\mathbb{T})$. That, however, is guaranteed by part (1). □

We remark that the previous corollary is only possible with the Lebesgue integral, since unlike the Riemann integral it guarantees that $L^2(\mathbb{T})$ is a complete space and thus a Hilbert space. We record two further basic facts which are immediate consequences of the approximate identity property of the Fejér kernels.

Corollary 1.7. *One has the following uniqueness property: If $f \in L^1(\mathbb{T})$ and $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$. More generally, if $\mu \in \mathcal{M}(\mathbb{T})$ satisfies $\hat{\mu}(n) = 0$ for all $n \in \mathbb{Z}$, then $\mu = 0$.*

Proof. One has $\sigma_N f = 0$ for all N by assumption and now apply the convergence property of Proposition 1.5. $\square$

The following is the *Riemann-Lebesgue* lemma.

Corollary 1.8. *If $f \in L^1(\mathbb{T})$, then $\hat{f}(n) \to 0$ as $n \to \infty$*

Proof. Given $\varepsilon > 0$, let N be such that $\|\sigma_N f - f\|_1 < \varepsilon$. Then $|\hat{f}(n)| = |\widehat{\sigma_N f}(n) - \hat{f}(n)| \leqslant \|\sigma_N f - f\|_1 < \varepsilon$ for all $|n| > N$. $\square$

We now turn our attention to the issue of convergence of the partial sums $S_N f$ in the sense of $L^p(\mathbb{T})$ or $C(\mathbb{T})$. Observe that it makes no sense to ask about uniform convergence of $S_N f$ for general $f \in L^\infty(\mathbb{T})$ because uniform limits of continuous functions are continuous.

Proposition 1.9. *The following statements are equivalent for any $1 \leqslant p \leqslant \infty$:*

a) *For every $f \in L^p(\mathbb{T})$ (or $f \in C(\mathbb{T})$ if $p = \infty$) one has*

$$\|S_N f - f\|_p \to 0 \ \text{ as } N \to \infty$$

b) $\sup_N \|S_N\|_{p\to p} < \infty$

Proof. The implication $b) \Longrightarrow a)$ follows from the fact that trigonometric polynomials are dense in the respective norms, see Corollary 1.6. The implication $a) \Longrightarrow b)$ can be deduced immediately from the uniform boundedness principle of functional analysis. Alternatively, can also prove this directly by the method of the "gliding hump": suppose $\sup_N \|S_N\|_{p\to p} = \infty$. For every positive integer ℓ one can therefore find a large integer N_ℓ such that

$$\|S_{N_\ell} f_\ell\|_p > 2^\ell$$

where f_ℓ is a trigonometric polynomial with $\|f_\ell\|_p = 1$. Now let

$$f(x) = \sum_{\ell=1}^{\infty} \frac{1}{\ell^2} e(M_\ell x) f_\ell(x)$$

with some integers $\{M_\ell\}$ to be specified. Notice that

$$\|f\|_p \leqslant \sum_{\ell=1}^{\infty} \frac{1}{\ell^2} \|f_\ell\|_p < \infty$$

Now choose $\{M_\ell\}$ tending to infinity so rapidly that the Fourier support of

$$e(M_j x) f_j(x)$$

lies to the right of the Fourier support of

$$\sum_{\ell=1}^{j-1} \frac{1}{\ell^2} e(M_\ell x) f_\ell(x)$$

for every $j \geqslant 2$ (here Fourier support means those integers for which the corresponding Fourier coefficients are non-zero). We also demand that $N_\ell - M_\ell \to \infty$ as $\ell \to \infty$. Then

$$\|(S_{M_\ell+N_\ell} - S_{M_\ell-N_\ell-1})f\|_p = \frac{1}{\ell^2}\|S_{N_\ell} f_\ell\|_p > \frac{2^\ell}{\ell^2}$$

which $\to \infty$ as $\ell \to \infty$. On the other hand, since $N_\ell + M_\ell \to \infty$ and $M_\ell - N_\ell - 1 \to \infty$, the left-hand side $\to 0$ as $\ell \to \infty$. This contradiction finishes the proof. □

Exercise 1.4. Let $\{c_n\}_{n\in\mathbb{Z}}$ be an arbitrary sequence of complex numbers and associate to it formally a Fourier series

$$f(x) \sim \sum_n c_n\, e(nx)$$

Show that there exists $\mu \in \mathcal{M}(\mathbb{T})$ with the property that $\hat{\mu}(n) = c_n$ for all $n \in \mathbb{Z}$ if and only if $\{\sigma_n f\}_{n\geqslant 1}$ is bounded in $\mathcal{M}(\mathbb{T})$. Discuss the case of $L^p(\mathbb{T})$ with $1 \leqslant p < \infty$ and $C(\mathbb{T})$ as well.

We can now settle the question of uniform convergence of S_N on functions in $C(\mathbb{T})$.

Corollary 1.10. *Fourier series do not converge on $C(\mathbb{T})$ and $L^1(\mathbb{T})$, i.e., there exists $f \in C(\mathbb{T})$ so that $\|S_N f - f\|_\infty \nrightarrow 0$ and $g \in L^1(\mathbb{T})$ so that $\|S_N g - g\|_1 \nrightarrow 0$ as $n \to \infty$.*

Proof. By Proposition 1.9 it suffices to verify the limits

$$\begin{aligned} \sup_N \|S_N\|_{\infty\to\infty} &= \infty \\ \sup_N \|S_N\|_{1\to 1} &= \infty \end{aligned} \tag{1.19}$$

Both properties follow from the fact that

$$\|D_N\|_1 \to \infty$$

as $N \to \infty$, see (1.7). To deduce (1.19) from this, notice that

$$\begin{aligned} \|S_N\|_{\infty\to\infty} &= \sup_{\|f\|_\infty=1} \|D_N * f\|_\infty \\ &\geqslant \sup_{\|f\|_\infty=1} |(D_N * f)(0)| = \|D_N\|_1 \end{aligned}$$

We remark that the inequality sign here can be replaced with an equality in view of the translation invariance (1.9). Furthermore, with $\{K_M\}_{M=1}^\infty$ being the Fejér kernels,

$$\|S_N\|_{1\to 1} \geqslant \|D_N * K_M\|_1 \to \|D_N\|_1$$

as $M \to \infty$. □

Exercise 1.5. The previous proof was indirect. Construct $f \in C(\mathbb{T})$ and $g \in L^1(\mathbb{T})$ so that $S_N f$ does not converge to f uniformly as $N \to \infty$, and such that $S_N g$ does not converge to g in $L^1(\mathbb{T})$.

In the positive direction, we shall see below that for $1 < p < \infty$

$$\sup_N \|S_N\|_{p \to p} < \infty$$

so that by Corollary 1.10 for any $f \in L^p(\mathbb{T})$ with $1 < p < \infty$

$$\|S_N f - f\|_p \to 0 \text{ as } N \to \infty \tag{1.20}$$

The case $p = 2$ is clear, see Corollary 1.6, but $p \neq 2$ is a somewhat deeper result. We will develop the theory of the conjugate function to obtain it, see Chapter 4. Note that unlike Theorem 1.2 there is no explicit rate of convergence here in terms of some expression involving N. Clearly, this cannot be expected given only the size of $\|f\|_p$ since $g(x) := e(-mx)f(x)$ has the same L^p-norm as f but $\hat{g}(n) = \hat{f}(n+m)$ for all $n \in \mathbb{Z}$. This suggests that we may hope to obtain such a rate provided we remove this freedom of translation on the Fourier side. One way of accomplishing this is by imposing a little regularity, as expressed for example in terms of the standard *Sobolev spaces*. Since we do not yet have the L^p-convergence in (1.20) for general p at our disposal, we need to restrict ourselves to $p = 2$ which is particularly simple.

Exercise 1.6. For any $s \in \mathbb{R}$ define the Hilbert space $H^s(\mathbb{T})$ by means of the norm

$$\|f\|_{H^s}^2 := |\hat{f}(0)|^2 + \sum_{n \in \mathbb{Z}} |n|^{2s} |\hat{f}(n)|^2 \tag{1.21}$$

Derive a rate of convergence for $\|S_N f - f\|_2$ in terms of N alone assuming that $\|f\|_{H^s} \leqslant 1$ where $s > 0$ is fixed.

We now investigate the connection between regularity of a function and its associated Fourier series somewhat further. The following estimate, which builds upon the fact that $\frac{d}{dx} e(nx) = 2\pi i n e(nx)$, goes by the name of *Bernstein's inequality*.

Proposition 1.11. *Let f be a trigonometric polynomial with $\hat{f}(k) = 0$ for all $|k| > n$. Then*

$$\|f'\|_p \leqslant Cn \|f\|_p$$

for any $1 \leqslant p \leqslant \infty$. The constant C is absolute.

Proof. Let

$$V_n(x) := (1 + e(nx) + e(-nx)) K_n(x) \tag{1.22}$$

be de la Vallée Poussin's kernel. We leave it to the reader to check that

$$\widehat{V}_n(j) = 1 \text{ if } |j| \leqslant n$$

as well as

$$\|V_n'\|_1 \leqslant Cn$$

Then $f = V_n * f$ and thus $f' = V_n' * f$ so that by Young's inequality

$$\|f'\|_p \leqslant \|V_n'\|_1 \|f\|_p \leqslant Cn\|f\|_p$$

as claimed. □

It is interesting to note that the following converse holds, due to Bohr.

EXERCISE 1.7. Suppose that $f \in C^1(\mathbb{T})$ satisfies $\hat{f}(j) = 0$ for all j with $|j| < n$. Then

$$\|f'\|_p \geqslant Cn\|f\|_p$$

for all $1 \leqslant p \leqslant \infty$, where C is independent of $n \in \mathbb{Z}^+$, the choice of f and p.

Now we will turn to the question how smoothness of a function reflects itself in the decay of its Fourier coefficients. We begin with the easy observation, based on integration by parts, that for any $f \in C^1(\mathbb{T})$ one has

$$\widehat{f'}(n) = 2\pi i n \hat{f}(n) \qquad \forall n \in \mathbb{Z} \tag{1.23}$$

This not only shows that $\hat{f}(n) = O(n^{-1})$ but also that $C^1(\mathbb{T}) \hookrightarrow H^1(\mathbb{T})$ where H^1 is the Sobolev space defined in (1.21). If f possesses more derivatives, say $k \geqslant 2$, then we may iterate this relation to obtain decay of the form $O(n^{-2})$. The following exercise establishes the connection between rapid decay of the Fourier coefficients and infinite regularity of the function.

EXERCISE 1.8. Let $f \in L^1(\mathbb{T})$.

- Show that $f \in C^\infty(\mathbb{T})$ if and only if $\hat{f}$ decays rapidly, i.e., for every $M \geqslant 1$ one has $\hat{f}(n) = O(|n|^{-M})$ as $|n| \to \infty$.
- Show that $f(x) = F(e(x))$ where F is analytic on some neighborhood of $\{|z| = 1\}$ if and only if $\hat{f}(n)$ decays exponentially, i.e., $\hat{f}(n) = O(e^{-\varepsilon|n|})$ as $|n| \to \infty$ for some $\varepsilon > 0$.

As far as less than one derivative is concerned, i.e., for Hölder continuous functions, one has the following facts. Starting from

$$\hat{f}(n) = -\int_{\mathbb{T}} e(-n(y + \frac{1}{2n}))f(y)\,dy = -\int_{\mathbb{T}} e(-ny)f\big(y - \frac{1}{2n}\big)\,dy$$

whence

$$\hat{f}(n) = \frac{1}{2}\int_{\mathbb{T}} e(-ny)\big(f(y) - f(y - \frac{1}{2n})\big)\,dy$$

which implies that if $f \in C^\alpha(\mathbb{T})$, then

$$\hat{f}(n) = O(n^{-\alpha}) \;\text{ as } |n| \to \infty \tag{1.24}$$

Note that (1.23) is strictly stronger than this bound, since it allows for square summation and thus the embedding into $H^1(\mathbb{T})$, whereas (1.24) does not. However, since we already observed that $C^1(\mathbb{T}) \hookrightarrow H^1(\mathbb{T})$ and since clearly $C^0(\mathbb{T}) \hookrightarrow H^0(\mathbb{T}) = L^2(\mathbb{T})$, one can invoke some interpolation machinery at this point to conclude that $C^\alpha(\mathbb{T}) \hookrightarrow H^\alpha(\mathbb{T})$ for all $0 \leqslant \alpha \leqslant 1$; however, we shall not make use of this fact.

Next, we ask how much regularity on f it takes so that

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)| < \infty \tag{1.25}$$

In other words, we ask for a sufficient condition for f to lie in the Wiener algebra $\mathbb{A}(\mathbb{T})$ of Exercise 1.2. There are a number of ways to interpret the requirement of "regularity". It is easy to settle this imbedding question on the level of the spaces $H^s(\mathbb{T})$. In fact, applying Cauchy-Schwarz yields

$$\sum_{n\neq 0} |\hat{f}(n)| \leqslant \Big(\sum_{n\neq 0} |\hat{f}(n)|^2 |n|^{1+\varepsilon}\Big)^{\frac{1}{2}} \Big(\sum_{n\neq 0} |n|^{-1-\varepsilon}\Big)^{\frac{1}{2}}$$

so that

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 |n|^{1+\varepsilon} < \infty \tag{1.26}$$

is a sufficient condition for (1.25) to hold. In other words, we have shown that $H^s(\mathbb{T}) \hookrightarrow \mathbb{A}(\mathbb{T})$ for any $s > \frac{1}{2}$. We leave it to the reader to check that this fails for $s = \frac{1}{2}$.

Next, we would like to address the more challenging question as to the minimal $\alpha > 0$ so that $C^\alpha(\mathbb{T}) \hookrightarrow \mathbb{A}(\mathbb{T})$. In fact, we first ask which α have the property that $C^\alpha(\mathbb{T}) \hookrightarrow H^s(\mathbb{T})$ for some $s > \frac{1}{2}$. We already observed that $\alpha > \frac{1}{2}$ suffices for this, but it required some interpolation theory.

Instead, we prefer to give a direct proof of this fact in the following theorem. It introduces an important idea that we shall see repeatedly throughout this book, namely the grouping of the Fourier coefficients $\hat{f}(n)$ into blocks of the same size of n; more precisely, we introduce the partial sums

$$(P_j f)(x) := \sum_{2^{j-1}\leqslant |n| < 2^j} \hat{f}(n) e(nx) \tag{1.27}$$

for all $j \geqslant 1$.

Theorem 1.12. *For any* $1 \geqslant \alpha > 0$ *one has* $C^\alpha(\mathbb{T}) \hookrightarrow H^\beta(\mathbb{T})$ *for arbitrary* $0 < \beta < \alpha$. *In particular, for any* $f \in C^\alpha(\mathbb{T})$ *with* $\alpha > \frac{1}{2}$ *one has*

$$\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty$$

and thus $C^\alpha(\mathbb{T}) \hookrightarrow \mathbb{A}(\mathbb{T})$ *for any* $\alpha > \frac{1}{2}$.

Proof. Let $[f]_\alpha \leqslant 1$. We claim that for every $j \geqslant 0$,

$$\sum_{2^j \leqslant |n| < 2^{j+1}} |\hat{f}(n)|^2 \leqslant C 2^{-2j\alpha} \tag{1.28}$$

The reader should note the similarity between this bound and the estimate (1.24), with (1.28) being of course strictly stronger. What allows for this improvement is the use of orthogonality or in other words, the Plancherel theorem. If (1.28) is true, then for any $\beta < \alpha$,

$$\sum_{n\neq 0} |\hat{f}(n)|^2 |n|^{2\beta} \leqslant C \sum_{j=0}^{\infty} 2^{2j\beta} \sum_{2^j \leqslant |n| < 2^{j+1}} |\hat{f}(n)|^2 \leqslant C \sum_{j=0}^{\infty} 2^{-2j(\alpha-\beta)} < \infty$$

To prove (1.28) we choose a kernel φ_j so that

$$\widehat{\varphi}_j(n) = 1 \text{ if } 2^j \leqslant |n| \leqslant 2^{j+1} \tag{1.29}$$

and

$$\widehat{\varphi}_j(n) = 0 \text{ if } |n| \ll 2^j \text{ or } |n| \gg 2^{j+1} \tag{1.30}$$

where $\ll$ and $\gg$ mean "much smaller" and "much bigger", respectively. The point is of course that (1.29) implies that

$$\sum_{2^j \leqslant |n| < 2^{j+1}} |\hat{f}(n)|^2 \leqslant \|\varphi_j * f\|_2^2 \tag{1.31}$$

so that it remains to bound the right-hand side for which (1.30) will be decisive — at least implicitly. The explicit properties of φ_j that will allow us to bound the right-hand side are *cancellation* meaning that $\widehat{\varphi}_j(0) = 0$ and bounds on the kernel $\varphi_j(x)$ which resemble those of the Fejér kernel K_N with $N = 2^j$.

There are various ways to construct φ_j. We use de la Vallée Poussin's kernel from above for this purpose. Set

$$\varphi_j(x) := V_{2^{j-1}}(x) \cdot (e((3 \cdot 2^{j-1} - 1)x) + e(-(3 \cdot 2^{j-1} - 1)x)) \tag{1.32}$$

We leave it to the reader to check that

$$\widehat{\varphi}_j(n) = 1 \text{ for } 2^j \leqslant |n| \leqslant 2^{j+1}$$

and that $\widehat{\varphi}_j(0) = 0$ which is the same as

$$\int_{\mathbb{T}} \varphi_j(x)\, dx = 0$$

Moreover, since the φ_j are constructed from Fejér kernels one has the bounds

$$|\varphi_j(x)| \leqslant C\, 2^{-j} \min\left(2^{2j},\, |x|^{-2}\right)$$

Therefore,

$$\begin{aligned}
|(\varphi_j * f)(x)| &= \Big| \int_{\mathbb{T}} \varphi_j(y)(f(x-y) - f(x))\, dy \Big| \\
&\leqslant \int_{\mathbb{T}} |\varphi_j(y)||f(x-y) - f(x)|\, dy \\
&\leqslant C \int_0^1 |\varphi_j(y)||y|^{\alpha}\, dy \\
&\leqslant C 2^{-j} \int_{|y| > 2^{-j}} |y|^{\alpha-2}\, dy + C 2^j \int_{|y| \leqslant 2^{-j}} |y|^{\alpha}\, dy \\
&\leqslant C 2^{-\alpha j}
\end{aligned}$$

as claimed. In summary, we see that for any $0 \leqslant \beta < \alpha \leqslant 1$

$$\|f\|_{H^\beta} \leqslant C(\alpha,\beta)\|f\|_{C^\alpha}$$

and the theorem follows. □

Next, we show that Theorem 1.12 is optimal up to the fact that one can take $\alpha = \beta$ in the first statement.

Exercise 1.9. Consider *lacunary series* of the form

$$f(x) := \sum_{k \geqslant 1} k^{-2} 2^{-\alpha k} e(2^k x)$$

to show that $C^\alpha(\mathbb{T})$ does not embed into H^β for any $\beta > \alpha$.

The second statement of Theorem 1.12 concerning $\mathbb{A}(\mathbb{T})$ is more difficult.

Proposition 1.13. *There is no embedding from $C^{\frac{1}{2}}(\mathbb{T})$ into $\mathbb{A}(\mathbb{T})$. In fact, there exists a function $f \in C^{\frac{1}{2}}(\mathbb{T}) \backslash \mathbb{A}(\mathbb{T})$.*

Proof. We claim that there exists a sequence of trigonometric polynomials $P_n(x) = \sum_{\ell=0}^{2^n-1} a_{n,\ell}\, e(\ell x)$ such that with $N = 2^n$,

$$\begin{aligned} \|P_n\|_\infty &\simeq \sqrt{N} \\ \|\widehat{P_n}\|_{\ell^1} &\simeq N \end{aligned} \tag{1.33}$$

for each $n \geqslant 1$. Here $a \simeq b$ means $C^{-1}a \leqslant b \leqslant Ca$ where C is an absolute constant, and $\widehat{P_n}$ refers to the sequence of Fourier coefficients.

Assuming such a sequence for now, we set

$$T_n(x) := 2^{-n} e(2^n x) P_n(x)$$
$$f := \sum_{n=1}^{\infty} T_n$$

Note that this series converges uniformly to $f \in C(\mathbb{T})$ since $\|T_n\|_\infty \leqslant C 2^{-\frac{n}{2}}$. Moreover, the Fourier supports of the T_n are pairwise disjoint for distinct n. Thus, $\|f\|_{\mathbb{A}(\mathbb{T})} = \infty$ and $f \notin \mathbb{A}(\mathbb{T})$. Finally, let $h \simeq 2^{-m}$ for some positive integer m. Then

$$\begin{aligned} |f(x+h) - f(x)| &\leqslant \sum_{n=1}^{m} |T_n(x+h) - T_n(x)| + \sum_{n>m} \|T_n\|_\infty \\ &\leqslant \sum_{n=1}^{m} C|h| \|T_n'\|_\infty + \sum_{n>m} C\, 2^{-\frac{n}{2}} \\ &\leqslant C\Big(\sum_{n=1}^{m} 2^{\frac{n}{2}}\, |h| + 2^{-\frac{m}{2}}\Big) \leqslant C|h|^{\frac{1}{2}} \end{aligned}$$

and thus $f \in C^{\frac{1}{2}}(\mathbb{T})$ as desired. To pass to the last line, we used Bernstein's inequality, i.e., Proposition 1.11.

It thus remains to establish the existence of polynomials as in (1.33). Define them inductively by $P_0(x) = Q_0(x) = 1$ and

$$
\begin{aligned}
P_{n+1}(x) &= P_n(x) + e(2^n x) Q_n(x) \\
Q_{n+1}(x) &= P_n(x) - e(2^n x) Q_n(x)
\end{aligned}
$$

for each $n \geqslant 0$. Since

$$|P_{n+1}(x)|^2 + |Q_{n+1}(x)|^2 = 2(|P_n(x)|^2 + |Q_n(x)|^2) = 2^{n+1}$$

we see that $\|P_{n+1}\|_\infty$ is of the desired size. Furthermore, all coefficients of both P_n and Q_n are either $+1$ or -1, and the only exponentials with nonzero coefficients are $e(\ell x)$ with $1 \leqslant \ell \leqslant 2^n - 1$. Hence $\{P_n\}_{n=0}^\infty$ is the desired family of polynomials. □

EXERCISE 1.10. Show that any trigonometric polynomial P with $\|\hat{P}\|_{\ell^1} = N$ and the property that the cardinality of its Fourier support is at most N satisfies $\sqrt{N} \leqslant \|P\|_2 \leqslant \|P\|_\infty \leqslant N$. Hence, the polynomials constructed in the previous proof are "extremal" for the lower bound on $\|P\|_\infty$, whereas the Dirichlet kernel (or Fejér kernel) are extremal for the upper bound.

We conclude this chapter with a brief discussion of Fourier series associated with functions on $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ with $d \geqslant 2$. The exponential basis in this case is $\{e(\nu \cdot x)\}_{\nu \in \mathbb{Z}^d}$, and this is an orthonormal family in the usual $L^2(\mathbb{T}^d)$ sense. Thus, to every measure $\mu \in \mathcal{M}(\mathbb{T}^d)$ we associate a Fourier series

$$\sum_{\nu \in \mathbb{Z}^d} \hat{\mu}(\nu)\, e(\nu \cdot x), \qquad \hat{\mu}(\nu) := \int_{\mathbb{T}^d} e(-\nu \cdot x)\, \mu(dx)$$

As before, a special role is played by the trigonometric polynomials

$$\sum_{\nu \in \mathbb{Z}^d} a_\nu\, e(\nu \cdot x)$$

where all but finitely many a_ν vanish. In contrast to the one-dimensional torus, in higher dimensions we face a very nontrivial ambiguity in the definition of partial sums. In fact, we can pose the convergence problem relative to any exhaustion of $\mathbb{Z}^d$ by finite sets A_k, which are increasing and whose union is all of $\mathbb{Z}^d$. Possible choices here are of course the squares $[-k, k]^d$, but one can also take more general rectangles, or balls relative to the Euclidean metric or other shapes. Even though this distinction may seem innocuous, it gave rise to very important developments in harmonic such as Fefferman's ball multiplier theorem, and the still unresolved Bochner-Riesz conjecture. We will return to these matters in a later chapter.

For now, we merely content ourselves with some very basic results.

PROPOSITION 1.14. *The space of trigonometric polynomials is dense in* $C(\mathbb{T}^d)$, *and one has the Plancherel theorem for* $L^2(\mathbb{T}^d)$, *i.e.,*

$$\|f\|_2^2 = \sum_{\nu \in \mathbb{Z}^d} |\hat{f}(\nu)|^2 \qquad \forall\, f \in L^2(\mathbb{T}^d)$$

If $f \in C^\infty(\mathbb{T}^d)$, then the Fourier series associated to f converges uniformly to f irrespective of the way in which the partial sums are formed.

Proof. We base the proof on the fact that the products of Fejér kernels with respect to the individual coordinate axes form an approximate identity. In other words, the family

$$K_{N,d}(x) := \prod_{j=1}^{d} K_N(x_j) \qquad \forall\ N \geqslant 1$$

forms an approximate identity on $\mathbb{T}^d$. Here $x = (x_1, \ldots, x_d)$. This follows immediately from the one-dimensional analysis above. Hence, for any $f \in C(\mathbb{T}^d)$ one has

$$\|K_{N,d} * f - f\|_\infty \to 0 \ \text{ as } N \to \infty$$

By inspection, $K_{N,d} * f$ is a trigonometric polynomial which implies the claimed density. This implies all assertions about the L^2 case.

Suppose that $f \in C^\infty(\mathbb{T}^d)$. Repeated integration by parts yields the estimate $\hat{f}(\nu) = O(|\nu|^{-m})$ as $|\nu| \to \infty$ for arbitrary but fixed $m \geqslant 1$. The convergence statement now follows from $K_{N,d} * f \to f$ in $C(\mathbb{T}^d)$ as $N \to \infty$, and the rapid decay of the coefficients. □

A useful corollary to the previous result is the density of tensor functions in $C(\mathbb{T}^d)$. A *tensor function* on $\mathbb{T}^d$ is a linear combination of functions of the form $\prod_{j=1}^{d} f_j(x_j)$ where $f_j \in C(\mathbb{T})$. In particular, trigonometric polynomials are tensor functions, whence the claim.

Notes

An encyclopedic treatment of Fourier series is the classic reference by Zygmund [**58**]. A less formidable but still comprehensive account of many classical results is Katznelson [**30**]. A standard reference for interpolation theory is the book by Bergh and Löfström [**4**]. For the construction in Proposition 1.13, see [**30**, page 36, Exercise 6.6]. For gap series see the timeless paper by Rudin [**41**] which introduced the Λ_p-set problem, later solved by Bourgain [**5**].

Problems

Problem 1.1. Suppose that $f \in L^1(\mathbb{T})$ and that $\{S_n f\}_{n=1}^\infty$ (the sequence of partial sums of the Fourier series) converges in $L^p(\mathbb{T})$ to g for some $p \in [1, \infty]$ and some $g \in L^p$. Prove that $f = g$. If $p = \infty$ conclude that f is continuous.

Problem 1.2. Let $T(x) = \sum_{n=0}^{N} [a_n \cos(2\pi n x) + b_n \sin(2\pi n x)]$ be an arbitrary trigonometric polynomial with real coefficients $a_0, \ldots, a_N, b_0, \ldots, b_N$. Show that there is a polynomial $P(z) = \sum_{\ell=0}^{2N} c_\ell z^\ell \in \mathbb{C}[z]$ so that

$$T(x) = e^{-2\pi i N x} P(e^{2\pi i x})$$

and such that $P(z) = z^{2N}\, \overline{P(\bar{z}^{-1})}$. How are the zeros of P distributed in the complex plane?

Problem 1.3. Suppose $T(x) = \sum_{n=0}^{N} \left[a_n \cos(2\pi nx) + b_n \sin(2\pi nx)\right]$ is such that $T \geqslant 0$ everywhere on $\mathbb{T}$ and $a_n, b_n \in \mathbb{R}$ for all $n = 0, 1, \ldots, N$. Show that there are $c_0, \ldots, c_N \in \mathbb{C}$ such that

$$T(x) = \Big| \sum_{n=0}^{N} c_n e^{2\pi i n x} \Big|^2 \tag{1.34}$$

Find the c_n for the Fejér kernel.

Problem 1.4. Suppose that $T(x) = a_0 + \sum_{h=1}^{H} a_h \cos(2\pi hx)$ satisfies $T(x) \geqslant 0$ for all $x \in \mathbb{T}$ and $T(0) = 1$. Show that for any complex numbers $y_1, y_2, \ldots, y_N$,

$$\Big| \sum_{n=1}^{N} y_n \Big|^2 \leqslant (N+H)\Big(a_0 \sum_{n=1}^{N} |y_n|^2 + \sum_{h=1}^{H} |a_h| \Big| \sum_{n=1}^{N-h} y_{n+h} \bar{y}_n \Big| \Big)$$

Hint: Write (1.34) with $\sum_n c_n = 1$. The apply Cauchy-Schwartz to $\sum_n y_n = \sum_{m,n} y_n\, c_m$.

Problem 1.5. Suppose $\sum_{n=1}^{\infty} n\, |a_n|^2 < \infty$ and $\sum_{1}^{\infty} a_n$ is Cesàro summable. Show that $\sum_{n=1}^{\infty} a_n$ converges. Use this to prove that any $f \in C(\mathbb{T}) \cap H^{\frac{1}{2}}(\mathbb{T})$ satisfies $S_n f \to f$ uniformly. Note that $H^{\frac{1}{2}}(\mathbb{T})$ does not embed into $\mathbb{A}(\mathbb{T})$, so this convergence does not follow by trivial means.

Problem 1.6. Show that there exists an absolute constant C so that

$$C^{-1} \sum_{n \neq 0} |n|\, |\hat{f}(n)|^2 \leqslant \int_{\mathbb{T}^2} \frac{|f(x) - f(y)|^2}{\sin^2(\pi(x-y))}\, dxdy \leqslant C \sum_{n \neq 0} |n|\, |\hat{f}(n)|^2$$

for any $f \in H^{\frac{1}{2}}(\mathbb{T})$.

Problem 1.7. Use the previous two problems to prove the following theorem of Pal-Bohr: For any real function $f \in C(\mathbb{T})$ there exists a homeomorphism $\varphi : \mathbb{T} \to \mathbb{T}$ such that

$$S_n(f \circ \varphi) \longrightarrow f \circ \varphi$$

uniformly. *Hint:* Without loss of generality let $f > 0$. Consider the domain defined in terms of polar coordinates by means of $r(\theta) = f(\theta/2\pi)$. Then apply the Riemann mapping theorem to the unit disc.

Problem 1.8. Show that

$$\|f * g\|^2_{L^2(\mathbb{T})} \leqslant \|f * f\|_{L^2(\mathbb{T})} \|g * g\|_{L^2(\mathbb{T})}$$

for all $f, g \in L^2(\mathbb{T})$.

Problem 1.9. Given N disjoint arcs $\{I_\alpha\}_{\alpha=1}^{N} \subset \mathbb{T}$, set $f = \sum_{\alpha=1}^{N} \chi_{I_\alpha}$. Show that

$$\sum_{|\nu| > k} |\hat{f}(\nu)|^2 \lesssim \frac{N}{k}.$$

Problem 1.10. Given any function $\psi : \mathbb{Z}^+ \to \mathbb{R}^+$ so that $\psi(n) \to 0$ as $n \to \infty$, show that you can find a measurable set $E \subset \mathbb{T}$ for which

$$\limsup_{n \to \infty} \frac{|\widehat{\chi_E}(n)|}{\psi(n)} = \infty.$$

Problem 1.11. The problem discusses some well-known partial differential equations on the level of Fourier series.

- Solve the heat equation $u_t - u_{\theta\theta} = 0$, $u(0) = u_0$ (the data at time $t = 0$) on $\mathbb{T}$ via Fourier series. Show that if $u_0 \in L^2(\mathbb{T})$, then $u(t, \theta)$ is analytic in θ for every $t > 0$ and solves the heat equation. Prove that $\|u(t) - u_0\|_2 \to 0$ as $t \to 0$. Write $u(t) = G_t * u_0$ and show that $G(t)$ is an approximate identity for $t > 0$. Conclude that $u(t) \to 0$ as $t \to 0$ in the L^p or $C(\mathbb{T})$ sense. Repeat for higher-dimensional tori.
- Solve the Schrödinger equation $iu_t - u_{\theta\theta} = 0$, $u(0) = u_0$ on $\mathbb{T}$ with $u_0 \in L^2(\mathbb{T})$. In what sense can you say that this Fourier series "solves" the equation? Show that $\|u(t)\|_2 = \|u_0\|_2$ for all t. Discuss the limit $u(t)$ as $t \to 0$. Repeat for higher-dimensional tori.
- Solve the wave equation $u_{tt} - u_{\theta\theta} = 0$ on $\mathbb{T}$ by Fourier series. Discuss the Cauchy problem as in the previous items. Show that if $u_t(0) = 0$, then with $u(0) = f$,
$$u(t, \theta) = \frac{1}{2}(f(\theta + t) + f(\theta - t))$$

CHAPTER 2

Harmonic Functions on $\mathbb{D}$ and the Poisson Kernel

There is a close connection between Fourier series and analytic or harmonic functions on the disc $\mathbb{D} := \{z \in \mathbb{C} \mid |z| \leqslant 1\}$. In fact, at least formally, Fourier series can be viewed as the "boundary values" of a Laurent series

$$\sum_{n=-\infty}^{\infty} a_n z^n \tag{2.1}$$

which can be seen by setting $z = x + iy = e(\theta)$. Alternatively, suppose we are given a continuous function f on $\mathbb{T}$ and wish to find the harmonic extension u of f into $\mathbb{D}$, i.e., a solution to

$$\begin{aligned} \triangle u &= 0 \text{ in } \mathbb{D} \\ u &= f \text{ on } \partial\mathbb{D} = \mathbb{T} \end{aligned} \tag{2.2}$$

"Solution" here refers to a classical one, i.e., a function $f \in C^2(\mathbb{D}) \cap C(\bar{\mathbb{D}})$. However, as we shall see it is also very important to investigate other notions of solutions of (2.2) with rougher f.

Note that we cannot use negative powers of z as in (2.1) in order to write an ansatz for u. However, we can use complex conjugates instead. Indeed, since $\triangle z^n = 0$ and $\triangle \bar{z}^n = 0$ for every integer $n \geqslant 0$, we are lead to defining

$$u(z) = \sum_{n=0}^{\infty} \hat{f}(n) z^n + \sum_{n=-\infty}^{-1} \hat{f}(n) \bar{z}^{|n|} \tag{2.3}$$

which at least formally satisfies $u(e(\theta)) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e(n\theta) = f(\theta)$. Inserting $z = re(\theta)$ and

$$\hat{f}(n) = \int_{\mathbb{T}} e(-n\varphi) f(\varphi)\, d\varphi$$

into (2.3) yields

$$u(re(\theta)) = \int_{\mathbb{T}} \sum_{n \in \mathbb{Z}} r^{|n|} e(n(\theta - \varphi)) f(\varphi)\, d\varphi$$

This resembles the derivation of the Dirichlet kernel in the previous chapter, and we now ask the reader to find a closed form for the sum.

Exercise 2.1. Check that, for $0 \leqslant r < 1$,

$$P_r(\theta) := \sum_{n \in \mathbb{Z}} r^{|n|} e(n\theta) = \frac{1 - r^2}{1 - 2r\cos(2\pi\theta) + r^2}$$

This is the *Poisson kernel*.

Based on our formal calculation above, we therefore expect to obtain the harmonic extension of a "nice enough" function f on $\mathbb{T}$ by means of the convolution

$$u(re(\theta)) = \int_{\mathbb{T}} P_r(\theta - \varphi) f(\varphi)\, dy = (P_r * f)(\theta)$$

for $0 \leqslant r < 1$.

Note that $P_r(\theta)$, for $0 \leqslant r < 1$, is a harmonic function of the variables $x + iy = re(\theta)$. Moreover, for any finite measure $\mu \in \mathcal{M}(\mathbb{T})$ the expression $(P_r * \mu)(\theta)$ is not only well-defined, but defines a harmonic function on $\mathbb{D}$.

The remainder of this chapter will therefore be devoted to analyzing the boundary behavior of $P_r * \mu$. Clearly, Proposition 1.5 will play an important role in this investigation, but the fact that we are dealing with harmonic functions will enter in a crucial way (such as through the maximum principle).

In the following, we use the notion of *approximate identity* in a more general form than in the previous chapter. However, the reader will have no difficulty transferring this notion to the present context, including Proposition 1.5.

EXERCISE 2.2. Check that $\{P_r\}_{0<r<1}$ is an approximate identity. The role of $N \in \mathbb{Z}^+$ in Definition 1.3 is played here by $0 < r < 1$ and $N \to \infty$ is replaced with $r \to 1$.

An important role is played by the kernel $Q_r(\theta)$ which is the *harmonic conjugate* of $P_r(\theta)$. Recall that this means that $P_r(\theta) + iQ_r(\theta)$ is analytic in $z = re(\theta)$ and $Q_0 = 0$. In this case it is easy to find $Q_r(\theta)$ since

$$P_r(\theta) = \mathrm{Re}\left(\frac{1+z}{1-z}\right)$$

and therefore

$$Q_r(\theta) = \mathrm{Im}\left(\frac{1+z}{1-z}\right) = \frac{2r\sin(2\pi\theta)}{1 - 2r\cos(2\pi\theta) + r^2}\ .$$

EXERCISE 2.3.

(1) Show that $\{Q_r\}_{0<r<1}$ is *not* an approximate identity.
(2) Check that $Q_1(\theta) = \cot(\pi\theta)$. Draw the graph of $Q_1(\theta)$. What is the asymptotic behavior of $Q_1(\theta)$ for θ close to zero?

We will study conjugate harmonic functions later. First, we clarify in what sense the harmonic extension $P_r * f$ of f attains f as its boundary values.

DEFINITION 2.1. For any $1 \leqslant p \leqslant \infty$ define

$$h^p(\mathbb{D}) := \left\{ u : \mathbb{D} \to \mathbb{C} \text{ harmonic} \;\middle|\; \sup_{0<r<1} \int_{\mathbb{T}} |u(re(\theta))|^p \, d\theta < \infty \right\}$$

These are the "little" Hardy spaces with norm

$$\|u\|_p := \sup_{0<r<1} \|u(re(\cdot))\|_{L^p(\mathbb{T})}$$

It is important to observe that $P_r(\theta) \in h^1(\mathbb{D})$. This function has "boundary values" δ_0 (the Dirac mass at $\theta = 0$) since $P_r = P_r * \delta_0$.

Theorem 2.2. *There is a one-to-one correspondence between $h^1(\mathbb{D})$ and $\mathcal{M}(\mathbb{T})$, given by $\mu \in \mathcal{M}(\mathbb{T}) \mapsto F_r(\theta) := (P_r * \mu)(\theta)$. Furthermore,*

$$\|\mu\| = \sup_{0<r<1} \|F_r\|_1 = \lim_{r\to 1} \|F_r\|_1 , \tag{2.4}$$

and

(1) *μ is absolutely continuous with respect to Lebesgue measure ($\mu \ll d\theta$) if and only if $\{F_r\}$ converges in $L^1(\mathbb{T})$. If so, then $d\mu = f\,d\theta$ where $f = L^1$-limit of F_r.*
(2) *The following are equivalent for $1 < p \leqslant \infty$: $d\mu = f\,d\theta$ with $f \in L^p(\mathbb{T})$*

$\Longleftrightarrow$ *$\{F_r\}_{0<r<1}$ is L^p- bounded*

$\Longleftrightarrow$ *$\{F_r\}$ converges in L^p if $1 < p < \infty$ and in the weak-$*$ sense in L^∞ if $p = \infty$ as $r \to 1$*

(3) *f is continuous $\Leftrightarrow$ F extends to a continuous function on $\overline{\mathbb{D}}$ $\Leftrightarrow$ F_r converges uniformly as $r \to 1-$.*

This theorem identifies $h^1(\mathbb{D})$ with $\mathcal{M}(\mathbb{T})$, and $h^p(\mathbb{D})$ with $L^p(\mathbb{T})$ for $1 < p \leqslant \infty$. Moreover, $h^\infty(\mathbb{D})$ contains the subclass of harmonic functions that can be extended continuously onto $\overline{\mathbb{D}}$; this subclass is the same as $C(\mathbb{T})$. Before proving the theorem we present two simple lemmas. In what follows we use the notation $F_r(\theta) := F\big(re(\theta)\big)$.

Lemma 2.3.

a) *If $F \in C(\overline{\mathbb{D}})$ and $\triangle F = 0$ in $\mathbb{D}$, then $F_r = P_r * F_1$ for any $0 \leqslant r < 1$.*
b) *If $\triangle F = 0$ in $\mathbb{D}$, then $F_{rs} = P_r * F_s$ for any $0 \leqslant r, s < 1$.*
c) *As a function of $r \in (0,1)$ the norms $\|F_r\|_p$ are non-decreasing for any $1 \leqslant p \leqslant \infty$.*

Proof.

a) Let $u(re(\theta)) := (P_r * F_1)(\theta)$ for any $0 \leqslant r < 1$, θ. Then $\triangle u = 0$ in $\mathbb{D}$. By Proposition 1.5 and Exercise 2.2, $\|u_r - F_1\|_\infty \to 0$ as $r \to 1$. Hence, u extends to a continuous function on $\overline{\mathbb{D}}$ with the same boundary values as F. By the maximum principle, $u = F$ as claimed.
b) Rescaling the disc $s\mathbb{D}$ to $\mathbb{D}$ reduces *b*) to *a*).
c) By b) and Young's inequality

$$\|F_{rs}\|_p \leqslant \|P_r\|_1 \|F_s\|_p = \|F_s\|_p$$

as claimed. □

Lemma 2.4. *Let $F \in h^1(\mathbb{D})$. Then there exists a unique measure $\mu \in \mathcal{M}(\mathbb{T})$ such that $F_r = P_r * \mu$.*

Proof. Since the unit ball of $\mathcal{M}(\mathbb{T})$ is weak-$*$ compact there exists a subsequence $r_j \to 1$ with $F_{r_j} \to \mu$ in weak-$*$ sense to some $\mu \in \mathcal{M}(\mathbb{T})$. Then, for any $0 < r < 1$,

$$P_r * \mu = \lim_{j\to\infty} (F_{r_j} * P_r) = \lim_{j\to\infty} F_{rr_j} = F_r$$

by Lemma 2.3, *b*). Let $f \in C(\mathbb{T})$. Then $\langle F_r, f\rangle = \langle P_r * \mu, f\rangle = \langle \mu, P_r * f\rangle \to \langle \mu, f\rangle$ as $r \to 1$ where we again use Proposition 1.5. This shows that in the weak-$*$ sense

$$\mu = \lim_{r\to 1} F_r \tag{2.5}$$

which implies uniqueness of μ. □

Proof of Theorem 2.2. If $\mu \in \mathcal{M}(\mathbb{T})$, then $P_r * \mu \in h^1(\mathbb{D})$. Conversely, given $F \in h^1(\mathbb{D})$ then by Lemma 2.4 there is a unique μ so that $F_r = P_r * \mu$. This gives the one-to-one correspondence. Moreover, (2.5) and Lemma 2.3 *c*) show that

$$\|\mu\| \leqslant \limsup_{r\to 1} \|F_r\|_1 = \sup_{0<r<1} \|F_r\|_1 = \lim_{r\to 1} \|F_r\|_1$$

Since clearly also

$$\sup_{0<r<1} \|F_r\|_1 \leqslant \sup_{0<r<1} \|P_r\|_1 \|\mu\| = \|\mu\|$$

(2.4) follows. If $f \in L^1(\mathbb{T})$ and $\mu(d\theta) = f(\theta)\, d\theta$, then Proposition 1.5 shows that $F_r \to f$ in $L^1(\mathbb{T})$. Conversely, if $F_r \to f$ in the sense of $L^1(\mathbb{T})$, then because of (2.5) necessarily $\mu(d\theta) = f(\theta)\, d\theta$ which proves (1). (2) and (3) are equally easy and we skip the details (simply invoke Proposition 1.5 again). □

Next, we turn to the issue of almost everywhere convergence of $P_r * f$ to f as $r \to 1$. By the previous theorem, we have pointwise convergence if f is continuous. If $f \in L^1(\mathbb{T})$, then we write $f = \lim_n g_n$ as a limit in L^1 with $g_n \in C(\mathbb{T})$. But then we face two limits, namely then one as $n \to \infty$ and that as $r \to 1-$, and we wish to interchange them. It is a common feature that such an interchange requires some form of *uniform control*. To be specific, we will use the *Hardy-Littlewood maximal function Mf* associated to f in order to dominate the convolution with the Poisson kernel $P_r * f$ *uniformly in* $0 < r < 1$. This turns out to be an instance of the general fact that *radially bounded approximate identities* are dominated by the Hardy-Littlewood maximal function, see condition A4) below.

The Hardy-Littlewood maximal function is defined as

$$Mf(\theta) = \sup_{\theta\in I\subset\mathbb{T}} \frac{1}{|I|} \int_I |f(\varphi)|\, d\varphi$$

where $I \subset \mathbb{T}$ is an open interval and $|I|$ denotes the length of I. The basic fact about Mf that we shall need in this context is given by the following result. We say that g belongs to weak-L^1 if

$$|\{\theta \in \mathbb{T} \mid |g(\theta)| > \lambda\}| \leqslant C\lambda^{-1}\|g\|_1$$

for all $\lambda > 0$. Henceforth, $|\cdot|$ applied to sets denotes Lebesgue measure. Note that any $g \in L^1$ belongs to weak-L^1. This is an instance of *Markov's inequality*

$$|\{\theta \in \mathbb{T} \mid |f(\theta)| > \lambda]\leqslant \lambda^{-p}\|f\|_p^p$$

for any $1 \leqslant p < \infty$ and any $\lambda > 0$.

Proposition 2.5. *The Hardy-Littlewood maximal function satisfies the following bounds:*

a) *M is bounded from L^1 to weak-L^1, i.e.,*

$$|\{\theta \in \mathbb{T} \mid Mf(\theta) > \lambda]| \leqslant \frac{3}{\lambda}\|f\|_1$$

for all $\lambda > 0$.

b) *For any $1 < p \leqslant \infty$, M is bounded on L^p.*

Proof. Fix some $\lambda > 0$ and any compact

$$K \subset \{\theta \in \mathbb{T} \mid Mf(\theta) > \lambda\} \tag{2.6}$$

There exists a finite cover $\{I_j\}_{j=1}^N$ of $\mathbb{T}$ by open arcs I_j such that

$$\int_{I_j} |f(\varphi)|\, d\varphi > \lambda |I_j| \tag{2.7}$$

for each j. We now apply Wiener's covering lemma to pass to a more convenient sub-cover: Select an arc of maximal length from $\{I_j\}$; call it J_1. Observe that any I_j such that $I_j \cap J_1 \neq \emptyset$ satisfies $I_j \subset 3 \cdot J_1$ where $3 \cdot J_1$ is the arc with the same center as J_1 and three times the length (if $3 \cdot J_1$ has length larger than 1, then set $3 \cdot J_1 = \mathbb{T}$). Now remove all arcs from $\{I_j\}_{j=1}^N$ that intersect J_1. Let J_2 be one of the remaining ones with maximal length. Continuing in this fashion we obtain arcs $\{J_\ell\}_{\ell=1}^L$ which are pair-wise disjoint and so that

$$\bigcup_{j=1}^N I_j \subset \bigcup_{\ell=1}^L 3 \cdot J_\ell$$

In view of (2.6) and (2.7) therefore

$$\begin{aligned} |K| &\leqslant \Big|\bigcup_{\ell=1}^L 3 \cdot J_\ell\Big| \leqslant 3 \sum_{\ell=1}^L |J_\ell| \\ &\leqslant \frac{3}{\lambda} \sum_{\ell=1}^L \int_{J_\ell} |f(\varphi)|\, d\varphi \leqslant \frac{3}{\lambda}\|f\|_1 \end{aligned}$$

as claimed.

To prove part b), one interpolates the bound from a) with the trivial L^∞ bound

$$\|Mf\|_\infty \leqslant \|f\|_\infty$$

by means of Marcinkiewicz's interpolation theorem. ∎

We remark that the same argument based on Wiener's covering lemma yields the corresponding statement for $M\mu$ when $\mu \in \mathcal{M}(\mathbb{T})$. In other words, if we define

$$(M\mu)(\theta) = \sup_{\theta \in I \subset \mathbb{T}} \frac{|\mu|(I)}{|I|}$$

then for all $\lambda > 0$

$$|\{\theta \in \mathbb{T} \mid (M\mu)(\theta) > \lambda\}| \leqslant \frac{3}{\lambda}\|\mu\|$$

where $\|\mu\|$ is the total variation of μ.

Exercise 2.4. Find within a multiplicative constant $M\delta_0$, where δ_0 is the Dirac measure at 0. Use this to prove that the weak-L^1 bound on $M\mu$ cannot be improved in general, even when μ is absolutely continuous.

For a considerable strengthening of the previous exercise, the reader should consult Problem 3.4 in the next chapter. We now introduce the class of approximate identities which are dominated by the maximal function.

Definition 2.6. Let $\{\Phi_n\}_{n=1}^{\infty}$ be an approximate identity as in Definition 1.3. We say that it is *radially bounded* if there exist functions $\{\Psi_n\}_{n=1}^{\infty}$ on $\mathbb{T}$ so that the following additional property holds:

A4) $|\Phi_n| \leqslant \Psi_n$, Ψ_n is even and decreasing, i.e., $\Psi_n(\theta) \leqslant \Psi_n(\varphi)$ for $0 \leqslant \varphi \leqslant \theta \leqslant \frac{1}{2}$, for all $n \geqslant 1$. Finally, we require that $\sup_n \|\Psi_n\|_1 < \infty$.

All basic examples of approximate identities which we have encountered so far (Fejér, Poisson, and box kernels) are of this type. The following lemma establishes the aforementioned uniform control of the convolution with radially bounded approximate identities in terms of the Hardy-Littlewood maximal function.

Lemma 2.7. *If* $\{\Phi_n\}_{n=1}^{\infty}$ *satisfies* A4)*, then for any* $f \in L^1(\mathbb{T})$ *one has*

$$\sup_n |(\Phi_n * f)(\theta)| \leqslant \sup_n \|\Psi_n\|_1 \; Mf(\theta)$$

for all $\theta \in \mathbb{T}$.

Proof. It clearly suffices to show the following statement: let

$$K : [-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R}^+ \cup \{0\}$$

be even and decreasing. Then for any $f \in L^1(\mathbb{T})$ we have

$$|(K * f)(\theta)| \leqslant \|K\|_1 Mf(\theta) \tag{2.8}$$

Indeed, assume that (2.8) holds. Then

$$\sup_n |(\Phi_n * f)(\theta)| \leqslant \sup_n (\Psi_n * |f|)(\theta) \leqslant \sup_n \|\Psi_n\|_1 \; Mf(\theta)$$

and the lemma follows. The idea behind (2.8) is to show that K can be written as an average of box kernels, i.e., for some positive measure μ

$$K(\theta) = \int_0^{\frac{1}{2}} \chi_{[-\varphi,\varphi]}(\theta)\, \mu(d\varphi) \tag{2.9}$$

We leave it to the reader to check that $d\mu = -dK + K\left(\frac{1}{2}\right)\delta_{\frac{1}{2}}$ is a suitable choice. Notice that (2.9) implies that

$$\int_0^1 K(\theta)\, d\theta = \int_0^{\frac{1}{2}} 2\varphi \, d\mu(\varphi)$$

Moreover, by (2.9),

$$
\begin{aligned}
|(K * f)(\theta)| &= \Big| \int_0^{\frac{1}{2}} \big(\frac{1}{2\varphi}\chi_{[-\varphi,\varphi]} * f\big)(\theta)\, 2\varphi\, \mu(d\varphi)\Big| \\
&\leqslant Mf(\theta) \int_0^{\frac{1}{2}} 2\varphi\, \mu(d\varphi) \\
&= Mf(\theta)\, \|K\|_1
\end{aligned}
$$

which is (2.8). □

The following theorem establishes the main almost everywhere convergence result of this chapter. The reader should note how the maximal function enters precisely in order to interchange the aforementioned double limit.

Theorem 2.8. *If $\{\Phi_n\}_{n=1}^{\infty}$ satisfies A1)–A4), then for any $f \in L^1(\mathbb{T})$ one has $\Phi_n * f \to f$ almost everywhere as $n \to \infty$.*

Proof. Pick $\varepsilon > 0$ and let $g \in C(\mathbb{T})$ with $\|f - g\|_1 < \varepsilon$. By Proposition 1.5, with $h = f - g$ one has

$$
\begin{aligned}
&|\{\theta \in \mathbb{T} \mid \limsup_{n\to\infty} |(\Phi_n * f)(\theta) - f(\theta)| > \sqrt{\varepsilon}\}| \\
\leqslant &|\{\theta \in \mathbb{T} \mid \limsup_{n\to\infty} |(\Phi_n * h)(\theta)| > \sqrt{\varepsilon}/2\}| + |\{\theta \in \mathbb{T} \mid |h(\theta)| > \sqrt{\varepsilon}/2\}| \\
\leqslant &|\{\theta \in \mathbb{T} \mid \sup_n |(\Phi_n * h)(\theta)| > \sqrt{\varepsilon}/2\}| + |\{\theta \in \mathbb{T} \mid |h(\theta)| > \sqrt{\varepsilon}/2\}| \\
\leqslant &|\{\theta \in \mathbb{T} \mid CMh(\theta) > \sqrt{\varepsilon}/2\}| + |\{\theta \in \mathbb{T}| \mid h(\theta)| > \sqrt{\varepsilon}/2\}| \\
\leqslant &C\sqrt{\varepsilon}
\end{aligned}
$$

To pass to the final inequality we used Proposition 2.5 as well as Markov's inequality (recall $\|h\|_1 < \varepsilon$). □

As a corollary we not only obtain the classical Lebesgue differentiation theorem (by considering the box kernel), but also almost everywhere convergence of the Cesáro means $\sigma_N f$ (via the Fejér kernel), as well as of the Poisson integrals $P_r * f$ to f for any $f \in L^1(\mathbb{T})$. A theorem of Kolmogoroff states that this *fails* for the partial sums $S_N f$ on $L^1(\mathbb{T})$. We will present this example in Chapter 5.

Exercise 2.5. It is natural to ask whether there is an analogue of Theorem 2.8 for measures $\mu \in \mathcal{M}(\mathbb{T})$. Prove the following:

a) If $\mu \in \mathcal{M}(\mathbb{T})$ is a positive measure which is singular with respect to Lebesgue measure m (in symbols, $\mu \perp m$), then for a.e. $\theta \in \mathbb{T}$ with respect to Lebesgue measure we have

$$\frac{\mu([\theta - \varepsilon, \theta + \varepsilon])}{2\varepsilon} \to 0 \text{ as } \varepsilon \to 0$$

b) Let $\{\Phi_n\}_{n=1}^{\infty}$ satisfy A1)–A4), and assume that the $\{\Psi_n\}_{n=1}^{\infty}$ from Definition 2.1 also satisfy

$$\sup_{\delta<|\theta|<\frac{1}{2}} |\Psi_n(\theta)| \to 0 \text{ as } n \to \infty$$

for all $\delta > 0$. Under these assumptions show that for any $\mu \in \mathcal{M}(\mathbb{T})$

$$\Phi_n * \mu \to f \text{ } a.e. \text{ as } n \to \infty$$

where $\mu(d\theta) = f(\theta)\, d\theta + \nu_s(d\theta)$ is the Lebesgue decomposition, i.e., $f \in L^1(\mathbb{T})$ and $\nu_s \perp m$.

Notes

Standard references on everything here are Hoffman [**25**], Garnett [**20**], and Koosis [**34**].

Problems

Problem 2.1. Let (X, μ) be a general measure space. We say a sequence $\{f_n\}_{n=1}^{\infty} \in L^1(\mu)$ is *uniformly integrable* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\mu(E) < \delta \Longrightarrow \sup_n \left| \int_E f_n \, du \right| < \varepsilon$$

Suppose μ is a finite measure. Let $\phi : [0, \infty) \to [0, \infty)$ be a continuous increasing function with $\lim_{t\to\infty} \frac{\phi(t)}{t} = +\infty$. Prove that

$$\sup_n \int \phi(|f_n(x)|)\, \mu(dx) < \infty$$

implies that $\{f_n\}$ is uniformly integrable.

Problem 2.2. Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence in $L^1(\mathbb{T})$. Show that there is a subsequence $\{f_{n_j}\}_{j=1}^{\infty}$ and a measure μ with $f_{n_j} \xrightarrow{\sigma*} \mu$ provided $\sup_n \|f_n\|_1 < \infty$. Here $\sigma*$ is the weak-star convergence of measures. Show that in general $\mu \notin L^1(\mathbb{T})$. However, if we assume that, in addition, $\{f_n\}_1^{\infty}$ is uniformly integrable, then $\mu(d\theta) = f(\theta)\, d\theta$ for some $f \in L^1(\mathbb{T})$. Can we conclude anything about strong convergence (i.e., in the L^1-norm) of $\{f_n\}$? Consider the analogous question on $L^p(\mathbb{T})$, $p > 1$.

Problem 2.3. Let μ be a positive finite Borel measure on $\mathbb{R}^d$ or $\mathbb{T}^d$. Set

$$M\mu(x) := \sup_{r>0} \frac{\mu(B(x,r))}{m(B(x,r))}$$

where m is the Lebesgue measure. This problem examines the behavior of $M\mu$ when μ is a singular measure, cf. Problem 3.4 for more on this case.

(1) Show that $\mu \perp m$ implies $\mu(\{x \mid M\mu(x) < \infty\}) = 0$
(2) Show that if $\mu \perp m$, then

$$\limsup_{r\to 0} \frac{\mu(B(x,r))}{m(B(x,r))} = \infty$$

for $\mu - a.e.\ x$. Also show that this limit vanishes m-a.e.

PROBLEM 2.4. For any $f \in L^1(\mathbb{R}^d)$ and $1 \leqslant k \leqslant d$ let

$$M_k f(x) = \sup_{r>0} r^{-k} \int_{B(x,r)} |f(y)|\, dy$$

Show that for every $\lambda > 0$

$$m_L(\{x \in L \mid M_k f(x) > \lambda\}) \leqslant \frac{C}{\lambda} \|f\|_{L^1(\mathbb{R}^d)}$$

where L is an arbitrary affine k-dimensional subspace and "m_L" stands for Lebesgue measure (i.e., k-dimensional measure) on this subspace. C is an absolute constant.

PROBLEM 2.5. Prove the Besicovitch covering lemma on $\mathbb{T}$: Suppose $\{I_j\}$ are finitely many arcs with $|I_j| < 1$. Then there is a sub-collection $\{I_{j_k}\}$ such that the following properties hold:

(1) $\cup_k I_{j_k} = \cup_j I_j$
(2) No point belongs to more than C of the I_{j_k}'s where C is an *absolute* constant. What is the best value of C?

What can you say about higher dimensions?

PROBLEM 2.6. Let $F \geqslant 0$ be a harmonic function on $\mathbb{D}$. Show that there exists a positive measure μ on $\mathbb{T}$ with $F(re(\theta)) = (P_r * \mu)(\theta)$ for $0 \leqslant r < 1$ and with $\|\mu\| = F(0)$.

PROBLEM 2.7. A sequence of complex numbers $\{a_n\}_{n\in\mathbb{Z}}$ is called positive definite if

- $\overline{a_n} = a_{-n}$ for all $n \in \mathbb{Z}$
- $\sum_{n,k} a_{n-k}\, z_n \overline{z_k} \geqslant 0$ for all complex sequences $\{z_n\}_{n\in\mathbb{Z}}$

Show that any positive definite sequence satisfies $|a_n| \leqslant a_0$ for all $n \in \mathbb{Z}$. Show that if μ is a positive measure, then $a_n = \int_{\mathbb{T}} e(-n\theta)\mu(d\theta)$ is such a sequence. Now prove that every positive definite sequence is of this form. *Hint:* apply the previous problem to the harmonic function $\sum_n a_n\, r^n e(n\theta)$.

CHAPTER 3

Analytic $h^1(\mathbb{D})$ Functions, F. & M. Riesz Theorem

In the previous chapter, we mainly dealt with functions harmonic in the disk satisfying various boundedness properties. We now turn to functions $F = u + iv \in h^1(\mathbb{D})$ which are *analytic* in $\mathbb{D}$. This is well-defined since analytic functions are complex-valued harmonic ones. These functions form the class $\mathbb{H}^1(\mathbb{D})$, the "big" Hardy space. We have shown there that $F_r = P_r * \mu$ for some $\mu \in \mathcal{M}(\mathbb{T})$. It is important to note that by analyticity $\hat{\mu}(n) = 0$ if $n < 0$. A theorem by F. & M. Riesz asserts that such measures are absolutely continuous. From the example

$$F(z) := \frac{1+z}{1-z} = P_r(\theta) + iQ_r(\theta), \qquad z = re(\theta)$$

one sees an important difference between the analytic and the harmonic cases. Indeed, while $P_r \in h^1(\mathbb{D})$, clearly $F \notin h^1(\mathbb{D})$. The boundary measure associated with P_r is δ_0, whereas F is not associated with any boundary measure in the sense of the previous chapter.

An important technical device that will allow us to obtain more information in the analytic case is given by *subharmonic functions*. Loosely speaking, this device will allow us to exploit algebraic properties of analytic functions that harmonic ones do not have (such as the fact that F^2 is again analytic if F is analytic, which does not hold in the harmonic category).

Definition 3.1. Let $\Omega \subset \mathbb{R}^2$ be a region (i.e., open and connected) and let $f : \Omega \to \mathbb{R} \cup \{-\infty\}$ where we extend the topology to $\mathbb{R} \cup \{-\infty\}$ in the obvious way. We say that f is *subharmonic* if

(1) f is continuous
(2) for all $z \in \Omega$ there exists $r_z > 0$ so that

$$f(z) \leqslant \int_0^1 f\big(z + re(\theta)\big)\, d\theta$$

for all $0 < r < r_z$. We refer to this as the (local) "submean value property".

It is helpful to keep in mind that in one dimension "harmonic = linear" and "subharmonic = convex". Subharmonic functions derive their name from the fact that they lie below harmonic ones, in the same way that convex functions lie below linear ones. We will make this precise by means of the important *maximum principle* which subharmonic functions obey. We begin by deriving some basic properties of this class.

Lemma 3.2. *Subharmonic functions satisfy the following properties:*

1) *If f and g are subharmonic, then $f \vee g := \max(f, g)$ is subharmonic.*
2) *If $f \in C^2(\Omega)$ then f is subharmonic $\Longleftrightarrow \triangle f \geqslant 0$ in Ω*
3) *F analytic $\Longrightarrow \log |F|$ and $|F|^\alpha$ with $\alpha > 0$ are subharmonic*
4) *If f is subharmonic and φ is increasing and convex, then $\varphi \circ f$ is subharmonic (we set $\varphi(-\infty) := \lim_{x\to-\infty} \varphi(x)$).*

Proof. 1) is immediate. For 2) use Jensen's formula

$$\int_{\mathbb{T}} f(z + re(\theta))\, d\theta - f(z) = \iint_{D(z,r)} \log \frac{r}{|w - z|} \triangle f(w)\, m(dw) \tag{3.1}$$

where m stands for two-dimensional Lebesgue measure and

$$D(z, r) = \{w \in \mathbb{C} \mid |w - z| < r\}$$

The reader is asked to verify this formula in the following exercise. If $\triangle f \geqslant 0$, then the submean value property holds. If $\triangle f(z_0) < 0$, then let $r_0 > 0$ be sufficiently small so that $\triangle f < 0$ on $D(z_0, r_0)$ Since $\log \frac{r_0}{|w-z_0|} > 0$ on this disk, Jensen's formula implies that the submean value property fails. Next, we verify 4) by means of Jensen's inequality for convex functions:

$$\varphi(f(z)) \leqslant \varphi\Big(\int_{\mathbb{T}} f(z + re(\theta))\, d\theta\Big) \leqslant \int_{\mathbb{T}} \varphi(f(z + re(\theta))\, d\theta$$

The first inequality sign uses that φ is increasing, whereas the second uses convexity of φ (this second inequality is called Jensen's inequality). If F is analytic, then $\log |F|$ is continuous with values in $\mathbb{R} \cup \{-\infty\}$. If $F(z_0) \neq 0$, then $\log |F(z)|$ is harmonic on some disk $D(z_0, r_0)$. Thus, one has the stronger mean value property on this disk. If $F(z_0) = 0$, then $\log |F(z_0)| = -\infty$, and there is nothing to prove. To see that $|F|^\alpha$ is subharmonic, apply 4) to $\log |F(z)|$ with $\varphi(x) = \exp(\alpha x)$. □

Exercise 3.1. Prove Jensen's formula (3.1) for C^2 functions.

Now we can derive the aforementioned domination of subharmonic functions by harmonic ones.

Lemma 3.3. *Let Ω be a bounded region. Suppose f is subharmonic on Ω, $f \in C(\bar{\Omega})$ and let u be harmonic on Ω, $u \in C(\bar{\Omega})$. If $f \leqslant u$ on $\partial\Omega$, then $f \leqslant u$ on Ω.*

Proof. We may take $u = 0$, so $f \leqslant 0$ on $\partial\Omega$. Let $M = \max_{\bar{\Omega}} f$ and assume that $M > 0$. Set

$$S = \{z \in \bar{\Omega} \mid f(z) = M\}$$

Then $S \subset \Omega$ and S is closed in Ω. If $z \in S$, then by the submean value property there exists $r_z > 0$ so that $D(z, r_z) \subset \Omega$. Hence S is also open. Since Ω is assumed to be connected, one obtains $S = \Omega$. This is a contradiction. □

The following lemma shows that the submean value property holds on any disk in Ω. The point here is that we upgrade the local submean value property to a true submean value property using the largest possible disks.

Lemma 3.4. *Let f be subharmonic in Ω, $z_0 \in \Omega$, $\overline{D(z_0, r)} \subset \Omega$. Then*

$$f(z_0) \leqslant \int_{\mathbb{T}} f(z_0 + re(\theta))\, d\theta$$

Proof. Let $g_n = \max(f, -n)$, where $n \geqslant 1$. Without loss of generality $z_0 = 0$. Define $u_n(z)$ to be the harmonic extension of g_n restricted to $\partial D(z_0, r)$ where $r > 0$ is as in the statement of the lemma. By the previous lemma,

$$f(0) \leqslant g_n(0) \leqslant u_n(0) = \int_{\mathbb{T}} u_n(re(\theta))\, d\theta$$

the last equality being the mean value property of harmonic functions. Since

$$\max_{|z| \leqslant r} u_n(z) \leqslant \max_{|z| \leqslant r} f(z)$$

the monotone convergence theorem for decreasing sequences yields

$$f(0) \leqslant \int_{\mathbb{T}} f(re(\theta))\, d\theta$$

as claimed. □

Corollary 3.5. *If g is subharmonic on $\mathbb{D}$, then for all θ*

$$g(rse(\theta)) \leqslant (P_r * g_s)(\theta)$$

for any $0 < r, s < 1$.

Proof. If $g > -\infty$ everywhere on $\mathbb{D}$, then this follows from Lemma 3.3. If not, then set $g_n = g \vee n$. Thus

$$g(rse(\theta)) \leqslant g_n(rse(\theta)) \leqslant (P_r * (g_n)_s)(\theta)$$

and consequently

$$g(rse(\theta)) \leqslant \limsup_{n \longrightarrow \infty} (P_r * (g_n)_s)(\theta) \leqslant (P_r * g_s)(\theta)$$

where the final inequality follows from Fatou's lemma (which can be applied in the "reverse form" here since the g_n have a uniform upper bound). □

Note that if $g_s \notin L^1(\mathbb{T})$, then $g \equiv -\infty$ on $D(0, s)$ and so $g \equiv -\infty$ on $D(0, 1)$. We now introduce the radial maximal function associated with any function on the disk. It, and the more general *nontangential maximal function* where the supremum is taken over a cone in $\mathbb{D}$, are of central importance in the analysis of this chapter.

Definition 3.6. Let F be any complex-valued function on $\mathbb{D}$ then $F^* : \mathbb{T} \to \mathbb{R}$ is defined as

$$F^*(\theta) = \sup_{0<r<1} |F(re(\theta))|$$

We showed in the previous chapter that any $u \in h^1(\mathbb{D})$ satisfies $u^* \leqslant CM\mu$ where μ is the boundary measure of u, i.e., $u_r = P_r * \mu$. In particular, one has

$$|\{\theta \in \mathbb{T} \mid u^*(\theta) > \lambda\}| \leqslant C\lambda^{-1}\|u\|_1$$

We shall prove the analogous bound for subharmonic functions which are L^1-bounded.

PROPOSITION 3.7. *Suppose g is subharmonic on $\mathbb{D}$, $g \geqslant 0$ and g is L^1-bounded, i.e.,*

$$\|g\|_1 := \sup_{0<r<1} \int_{\mathbb{T}} g(re(\theta))\, d\theta < \infty$$

Then

(1) $|\{\theta \in \mathbb{T} \mid g^*(\theta) > \lambda\}| \leqslant \frac{3}{\lambda}\|g\|_1$ for $\forall \lambda > 0$.
(2) *If g is L^p bounded, i.e.,*

$$\|g\|_p^p := \sup_{0<r<1} \int_{\mathbb{T}} g(re(\theta))^p\, d\theta < \infty$$

with $1 < p \leqslant \infty$, then

$$\|g^*\|_{L^p(\mathbb{T})} \leqslant C_p\|g\|_p$$

PROOF.

(1) Let $g_{r_n} \rightharpoonup \mu \in \mathcal{M}(\mathbb{T})$ in the weak-$*$ sense. Then $\|\mu\| \leqslant \|g\|_1$ and

$$g_s \longleftarrow g_{r_n s} \leqslant g_{r_n} * P_s \longrightarrow P_s * \mu$$

Thus, by Lemma 2.7,

$$g^* \leqslant \sup_{0<s<1} P_s * \mu \leqslant M\mu$$

and the desired bound now follows from Proposition 2.5.
(2) If $\|g\|_p < \infty$, then $\frac{d\mu}{d\theta} \in L^p(\mathbb{T})$ with $\|\frac{d\mu}{d\theta}\|_p \leqslant \|g\|_p$ and thus

$$g^* \leqslant CM\Big(\frac{d\mu}{d\theta}\Big) \in L^p(\mathbb{T})$$

by Proposition 2.5, as claimed.

□

We now present three versions of a theorem due to F. & M. Riesz. It is important to note that the following result fails without analyticity.

THEOREM 3.8 (First Version of the F. & M. Riesz Theorem). *Suppose $F \in h^1(\mathbb{D})$ is analytic. Then $F^* \in L^1(\mathbb{T})$.*

PROOF. $|F|^{\frac{1}{2}}$ is subharmonic and L^2-bounded. By Proposition 3.7 therefore $|F|^{\frac{1}{2}*} \in L^2(\mathbb{T})$. But $|F|^{\frac{1}{2}*} = |F^*|^{\frac{1}{2}}$ and thus $F^* \in L^1(\mathbb{T})$. □

Let $F \in h^1(\mathbb{D})$. By Theorem 2.2, $F_r = P_r * \mu$ where $\mu \in \mathcal{M}(\mathbb{T})$ has a Lebesgue decomposition $\mu(d\theta) = f(\theta)\, d\theta + \nu_s(d\theta)$, ν_s singular and $f \in L^1(\mathbb{T})$. By Exercise 2.5 *b*) one has $P_r * \mu \to f$ a.e. as $r \to 1$. Thus, $\lim_{r\to 1} F(re(\theta)) = f(\theta)$ exists for a.e. $\theta \in \mathbb{T}$. This justifies the statement of the following theorem.

Theorem 3.8 (Second Version). *Assume $F \in h^1(\mathbb{D})$ and F analytic. Let $f(\theta) = \lim_{r\to 1} F\big(re(\theta)\big)$. Then $F_r = P_r * f$ for all $0 < r < 1$.*

Proof. We have $F_r \to f$ a.e. and $|F_r| \leqslant F^* \in L^1$ by the previous theorem. Therefore, $F_r \to f$ in $L^1(\mathbb{T})$ and Theorem 2.2 finishes the proof. □

Theorem 3.8 (Third Version). *Suppose $\mu \in \mathcal{M}(\mathbb{T})$, $\hat{\mu}(n) = 0$ for all $n < 0$. Then μ is absolutely continuous with respect to Lebesgue measure.*

Proof. Since $\hat{\mu}(n) = 0$ for $n \in \mathbb{Z}^-$ one has that $F_r = P_r * \mu$ is analytic on $\mathbb{D}$. By the second version above and the remark preceding it, one concludes that $\mu(d\theta) = f(\theta)\, d\theta$ with $f = \lim_{r\to 1-} F\big(re(\theta)\big) \in L^1(\mathbb{T})$, as claimed. □

The logic of this argument shows that if $\mu \perp m$ (where m is Lebesgue measure), then the harmonic extension u_μ of μ satisfies $u_\mu^* \notin L^1(\mathbb{T})$. It is possible to give a more quantitative version of this fact. Indeed, suppose that μ is a positive measure. Then for some absolute constant C,

$$C^{-1} M\mu < u_\mu^* < CM\mu \tag{3.2}$$

where the upper bound is Lemma 2.7 (applied to the Poisson kernel) and the lower bound follows from the assumption $\mu \geqslant 0$ and the fact that the Poisson kernel dominates the box kernel. Part a) of Problem 3.4 below therefore implies the quantitative non-L^1 statement

$$|\{\theta \in \mathbb{T} \mid u_\mu^*(\theta) \geqslant \lambda\}| \geqslant \frac{C}{\lambda}\|\mu\|$$

for $\mu \perp m$, where μ is a positive measure.

The F. & M. Riesz theorem raises the following question: Given $f \in L^1(\mathbb{T})$, how can one decide if

$$P_r * f + iQ_r * f \in h^1(\mathbb{D}) \ ?$$

We know that necessarily $u_f^* = (P_r * f)^* \in L^1(\mathbb{T})$. A theorem by Burkholder, Gundy, and Silverstein states that this is also sufficient (they proved this for the *non-tangential* maximal function). It is important to note the difference from (3.2), i.e., that this is *not* the same as the Hardy-Littlewood maximal function satisfying $Mf \in L^1(\mathbb{T})$ due to possible cancellation in f. In fact, it is known that

$$Mf \in L^1(\mathbb{T}) \Longleftrightarrow |f| \log(2 + |f|) \in L^1$$

, see Problem 7.6.

We conclude this chapter with another theorem due to the Riesz brothers, which can be seen as a generalization of the uniqueness theorem for analytic functions. In fact, it says that if an analytic function F on the disk does not have too wild growth as one approaches the boundary (expressed through the L^1-boundedness condition), then the boundary values are well-defined and cannot vanish on a set of positive measure (unless F vanishes identically).

THEOREM 3.9 (Second F. & M. Riesz Theorem). *Let F be analytic on $\mathbb{D}$ and L^1-bounded, i.e., $F \in h^1(\mathbb{D})$. Assume $F \not\equiv 0$ and set $f = \lim_{r\to 1-} F_r$. Then $\log|f| \in L^1(\mathbb{T})$. In particular, f does not vanish on a set of positive measure.*

PROOF. The idea is that if $F(0) \neq 0$, then

$$\int_{\mathbb{T}} \log|f(\theta)|\, d\theta \geqslant \log|F(0)| > -\infty$$

Since $\log_+|f| \leqslant |f| \in L^1(\mathbb{T})$ by Theorem 3.8, we should be done. Some care needs to be taken, though, as F attains the boundary values f only in the almost everywhere sense. This issue can easily be handled by means of Fatou's lemma. First, $F^* \in L^1(\mathbb{T})$, so $\log_+|F_r| \leqslant F^*$ implies that $\log_+|f| \in L^1(\mathbb{T})$ by Lebesgue dominated convergence. Second, by subharmonicity,

$$\int_{\mathbb{T}} \log|F_r(\theta)|\, d\theta \geqslant \log|F(0)|$$

so that

$$\begin{aligned}\int_{\mathbb{T}} \log|f(\theta)|\, d\theta &= \int_{\mathbb{T}} \lim_{r\to 1} \log|F_r(\theta)|\, d\theta \\ &\geqslant \limsup_{r\to 1} \int_{\mathbb{T}} \log|F_r(\theta)|\, d\theta \geqslant \log|F(0)|\end{aligned}$$

If $F(0) \neq 0$, then we are done. If $F(0) = 0$, then choose another point $z_0 \in \mathbb{D}$ for which $F(z_0) \neq 0$. Now one either repeats the previous argument with the Poisson kernel instead of the submean value property, or one composes F with an automorphism of the unit disk that moves 0 to z_0. Then the previous argument applies. □

Theorem 3.9 generalizes the following fact from complex analysis, which is proved by Schwarz reflection and Riemann mapping: Let F be analytic in Ω and continuous up to $\Gamma \subset \partial\Omega$ which is an open Lipschitz arc. If $F = 0$ on Γ, then $F \equiv 0$ in Ω.

Notes

For all of this see Hoffman's classic [**25**], Garnett [**20**], and especially Koosis [**34**]. The Burkholder-Gundy-Silverstein theorem, which is proved in [**34**], can be seen as a precursor to the Fefferman-Stein *grand maximal function* which leads to the real-variable theory of Hardy space, see [**46**]. There are some glaring omissions in this chapter such as inner and outer functions. However, we shall not require that aspect of H^p theory in this book.

Problems

Problem 3.1. Establish the maximum principle for subharmonic functions:

a) Show that if u is subharmonic on Ω and satisfies $u(z) \leqslant u(z_0)$ for all $z \in \Omega$ and some fixed $z_0 \in \Omega$, then u is a constant.

b) Show that if Ω is bounded and u is continuous on $\bar{\Omega}$, then $u(z) \leqslant \max_{\partial\Omega} u$ for all $z \in \Omega$ with equality being possibly only if u is a constant.

Problem 3.2. Let u be subharmonic on a domain Ω. Show that there exist a unique measure μ on Ω such that $\mu(K) < \infty$ for every $K \subset\subset \Omega$ (i.e., K is a compact subset of Ω) and so that

$$u(z) = \iint \log |z - \zeta| \, d\mu(\zeta) + h(z)$$

where h is harmonic on Ω. This is called "Riesz's representation of subharmonic functions".

Problem 3.3. With u and μ as in the previous exercise, show that

$$\int_{\mathbb{T}} u(z + re(\theta)) \, d\theta - u(z) = \int_0^r \frac{\mu(D(z,t))}{t} \, dt$$

for all $D(z,r) \subset \Omega$ (this generalizes "Jensen's formula" (3.1) to functions which are not twice continuously differentiable). Recover the classical Jensen's formula from complex analysis by setting $u(z) := \log |f(z)|$ where f is analytic.

Problem 3.4. This problem explores the behavior of the Hardy-Littlewood maximal function of measures which are singular relative to Lebesgue measure.

a) Prove that if $\mu \in \mathcal{M}(\mathbb{T})\setminus\{0\}$ satisfies $\mu \perp m$, then $M\mu \notin L^1$. In fact, show that

$$|\{\theta \in \mathbb{T} \mid M\mu(\theta) > \lambda\}| \geqslant \frac{c}{\lambda}\|\mu\|$$

provided $\lambda > \|\mu\|$ with an absolute constant $c > 0$.

b) Prove that there is a numerical constant C such that if $\mu \in \mathcal{M}(\mathbb{T})$ is a *positive* measure and F the associated harmonic function, then $M\mu \leqslant CF^*$. Conclude that if μ is singular, then $F^* \notin L^1$.

Problem 3.5. Show that the class of analytic function $F(z)$ in the unit disk with positive real part is in one-to-one relation with the class of functions of the form

$$F(\zeta) = \int_{\mathbb{T}} \frac{1 + \zeta e(-\theta)}{1 - \zeta e(\theta)} \, \mu(d\theta) + ic, \qquad |\zeta| < 1$$

where μ is a positive Borel measure on $\mathbb{T}$ and c is a real constant.

CHAPTER 4

The Conjugate Harmonic Function

While the previous two chapters focused on the Poisson kernel P_r, this chapter is devoted to the study of the kernel Q_r conjugate to P_r which we introduced in Chapter 2. This turns out to have far-reaching consequences, for example it will allow us to answer the question of L^p-convergence of Fourier series with $1 < p < \infty$ which was raised in Chapter 1. We begin by recalling the definition of *conjugate function*.

Definition 4.1. Let u be real-valued and harmonic in $\mathbb{D}$. Then we define $\tilde{u}$ to be that unique real-valued and harmonic function in $\mathbb{D}$ for which $u + i\tilde{u}$ is analytic and $\tilde{u}(0) = 0$. If u is complex-valued and harmonic, then we set $\tilde{u} := (\operatorname{Re} u)\tilde{} + i(\operatorname{Im} u)\tilde{}$.

The following lemma presents some simple but useful properties of the harmonic conjugate $\tilde{u}$.

Lemma 4.2. *Let $\tilde{u}$ be the harmonic conjugate as above.*

1) *If u is constant, then $\tilde{u} = 0$.*
2) *If u is analytic in $\mathbb{D}$ and $u(0) = 0$, then $\tilde{u} = -iu$. If u is co-analytic (meaning that $\bar{u}$ is analytic), then $\tilde{u} = iu$.*
3) *Any harmonic function u can be written uniquely as $u = c + f + \bar{g}$ with c =constant, f, g analytic, and $f(0) = g(0) = 0$.*

Proof. 1) and 2) follow immediately from the definition, whereas 3) is given by

$$u = u(0) + \frac{1}{2}(u - u(0) + i\tilde{u}) + \frac{1}{2}(u - u(0) - i\tilde{u})$$

Uniqueness of c, f, g is also clear. □

There is a simple relation between the Fourier coefficients of u_r and that of its harmonic conjugate. This relation will be the key to expressing the partial sum operator for Fourier series in terms of the harmonic conjugate. Henceforth, we shall occasionally use $\mathcal{F}$ to denote the Fourier transform for notational reasons.

Lemma 4.3. *Suppose u is harmonic on $\mathbb{D}$. Then for all $n \in \mathbb{Z}$, $n \neq 0$*

$$\mathcal{F}(\tilde{u}_r)(n) = -i \operatorname{sign}(n)\widehat{u_r}(n) \tag{4.1}$$

Proof. By Lemma 4.2 it suffices to consider u = constant, analytic, co-analytic. We present the case u = analytic, $u(0) = 0$. Then $\tilde{u} = -iu$ so that $\mathcal{F}(\tilde{u}_r)(n) = -i\widehat{u_r}(n)$ for all $u \in \mathbb{Z}$. But $\widehat{u_r}(n) = 0$ for $u \leqslant 0$ and thus (4.1) holds in this case. □

Together with the Plancherel theorem, the previous lemma immediately allows us to conclude the following relation between the L^2-norms.

Proposition 4.4. *Let $u \in h^2(\mathbb{D})$ be real-valued. Then $\tilde{u} \in h^2(\mathbb{D})$. In fact,*

$$\|\tilde{u}_r\|_2^2 = \|u_r\|_2^2 - |u(0)|^2$$

Proof. As already mentioned, this follows from Lemma 4.3 and the Plancherel theorem. Alternatively, by Cauchy's theorem,

$$\int_{\mathbb{T}} (u_r + i\tilde{u}_r)^2(\theta)\, d\theta = (u + i\tilde{u})^2(0) = u^2(0)$$

Since the right-hand side is real-valued, the left-hand side is also necessarily real and thus

$$u^2(0) = \int_{\mathbb{T}} u_r^2(\theta)\, d\theta - \int_{\mathbb{T}} \tilde{u}_r^2(\theta)\, d\theta$$

as claimed. □

The previous L^2-boundedness result allows us to determine the boundary values of the conjugates of h^2 functions.

Corollary 4.5. *If $u \in h^2(\mathbb{D})$, then $\lim_{r\to 1} \tilde{u}(re(\theta))$ exists for a.e. $\theta \in \mathbb{T}$ as an $L^2(\mathbb{T})$ function, and the convergence is also in the sense of L^2.*

Proof. By Proposition 4.4 we know that $\tilde{u} \in h^2(\mathbb{D})$ and the result of Chapter 2 lead to the desired conclusion. □

We now consider the case of $h^1(\mathbb{D})$. From the example of the Poisson kernel and its conjugate we see that there is no hope to obtain a result as strong as in the case of h^2. However, we shall now show that the radial maximal function (as defined in the previous chapter) of the conjugate to any $h^1(\mathbb{D})$-function is bounded in weak-$L^1(\mathbb{T})$.

We emphasize, however, that this has nothing to do with the weak-L^1 boundedness of the Hardy-Littlewood maximal function. In fact, it is not possible to control the convolutions with Q_r by means of that maximal function uniformly in $0 < r < 1$.

Nevertheless, the following theorem due to Besicovitch and Kolmogoroff shows that the weak-L^1 property still holds. The proof presented here is based on the notion of *harmonic measure*.

Theorem 4.6. *Let $u \in h^1(\mathbb{D})$. Then*

$$\left|\{\theta \in \mathbb{T} \mid |\tilde{u}^*(\theta)| > \lambda\}\right| \leqslant \frac{C}{\lambda} \|u\|_1$$

with some absolute constant C.

PROOF. By Theorem 2.2, $u_r = P_r * \mu$. Splitting μ into real and imaginary parts, and then each piece into its positive and negative parts, we reduce ourselves to the case $u \geqslant 0$. Let

$$E_\lambda = \{\theta \in \mathbb{T} \mid \tilde{u}^*(\theta) > \lambda\}$$

and set $F = -\tilde{u} + iu$. Then F is analytic and $F(0) = iu(0)$. Define a function

$$\omega_\lambda(x, y) := \frac{1}{\pi} \int_{(-\infty,-\lambda)\cup(\lambda,\infty)} \frac{y}{(x-t)^2 + y^2}\, dt$$

which is harmonic for $y > 0$ and non negative. The following two properties of ω_λ will be important:

(1) $\omega_\lambda(x, y) \geqslant \frac{1}{2}$ if $|x| > \lambda$
(2) $\omega_\lambda(0, y) \leqslant \frac{2y}{\pi\lambda}$.

For the first property it suffices to note that

$$\int_0^\infty \frac{1}{\pi} \frac{y}{x^2 + y^2}\, dx = \frac{1}{2} \qquad \forall y > 0$$

For the second property compute

$$\omega_\lambda(0, y) = \frac{1}{\pi} \int_{(-\infty,-\lambda)\cup(\lambda,\infty)} \frac{y}{t^2 + y^2}\, dt \leqslant \frac{2}{\pi} \int_{\lambda/y}^\infty \frac{dt}{1 + t^2} \leqslant \frac{2y}{\pi\lambda},$$

as claimed.

Observe now that $\omega_\lambda \circ F$ is harmonic and that $\theta \in E_\lambda$ implies

$$(\omega_\lambda \circ F)(re(\theta)) \geqslant \frac{1}{2}$$

for some $0 < r < 1$. Thus

$$|E_\lambda| \leqslant |\{\theta \in \mathbb{T} \mid (\omega_\lambda \circ F)^*(\theta) \geqslant 1/2\}| \leqslant \frac{3}{1/2} \|\omega_\lambda \circ F\|_1 \tag{4.2}$$

by Proposition 3.7. Since $\omega_\lambda \circ F \geqslant 0$, the mean value property implies that

$$\|\omega_\lambda \circ F\|_1 = (\omega_\lambda \circ F)(0) = \omega_\lambda(iu(0)) \leqslant \frac{2}{\pi} \frac{u(0)}{\lambda} = \frac{2}{\pi} \frac{\|u\|_1}{\lambda}.$$

Combining this with (4.2) yields

$$|E_\lambda| \leqslant \frac{12}{\pi} \frac{\|u\|_1}{\lambda}$$

and we are done. □

EXERCISE 4.1. Show that $\omega_\lambda(x, y)$ equals an angle measured from $z = x + iy$. Which angle? Use this to show that $\omega_\lambda(x, y) \geqslant \frac{1}{2}$ provided (x, y) lies outside the semi-circle with radius λ and center 0. Furthermore, show that ω_λ is the unique harmonic function in the upper half plane which equals 1 on $\Gamma := (-\infty, -\lambda] \cup [\lambda, \infty)$, equals 0 on $(-\lambda, \lambda) = \mathbb{R}\backslash\bar{\Gamma}$, and which is globally bounded. It is called the *harmonic measure* of $(-\infty, -\lambda] \cup [\lambda, \infty)$, and can be defined in the same fashion on general domains Ω and any open $\Gamma \subset \bar{\Omega}$ (in fact, more general Γ). It turns out that the harmonic measure of Γ relative to Ω equals the probability that Brownian motion starting at $z \in \Omega$ hits the boundary for the first time at Γ rather than in $\partial\Omega\backslash\Gamma$. Of particular importance is the conformal invariance of this notion, but we make no use of this probabilistic connection in this book.

The following result introduces the *Hilbert transform* and establishes a weak-L^1 bound for it. Formally speaking, the Hilbert transform $H\mu$ of a measure $\mu \in \mathcal{M}(\mathbb{T})$ is defined by

$$\mu \mapsto u_\mu \mapsto \widetilde{u_\mu} \mapsto \lim_{r\to 1} (\widetilde{u_\mu})_r =: H\mu$$

i.e., the Hilbert transform of a function on $\mathbb{T}$ equals the boundary values of the conjugate function of its harmonic extension. By Corollary 4.5 this is well defined if $\mu(d\theta) = f(\theta)\,d\theta$, $f \in L^2(\mathbb{T})$. We now consider the case $f \in L^1(\mathbb{T})$.

Corollary 4.7. *Given $u \in h^1(\mathbb{D})$ the limit $\lim_{r\to 1} \tilde{u}(re(\theta))$ exists for a.e. θ. With $u = P_r * \mu$, $\mu \in M(\mathbb{T})$, this limit is denoted by $H\mu$. There is the weak-L^1 bound*

$$|\{\theta \in \mathbb{T} \mid |H\mu(\theta)| > \lambda\}| \leqslant \frac{C}{\lambda}\|\mu\|$$

Proof. If $\mu(d\theta) = f(\theta)\,d\theta$ with $f \in L^2(\mathbb{T}$, then $\lim\limits_{r\to 1} \widetilde{u_f}(re(\theta))$ exists for a.e. θ by Corollary 4.5. If $f \in L^1(\mathbb{T})$ and $\varepsilon > 0$, then let $g \in L^2(\mathbb{T})$ such that $\|f - g\|_1 < \varepsilon$. Denote, for any $\delta > 0$,

$$E_\delta := \{\theta \in \mathbb{T} \mid \limsup_{r,s\to 1} |\widetilde{u_f}(re(\theta)) - \widetilde{u_f}(se(\theta))| > \delta\}$$

and

$$F_\delta := \{\theta \in \mathbb{T} \mid \limsup_{r,s\to 1} |\widetilde{u_h}(re(\theta)) - \widetilde{u_h}(se(\theta))| > \delta\}$$

where $h = f - g$. In view of the preceding theorem and the L^2-case,

$$\begin{aligned} |E_\delta| = |F_\delta| &\leqslant |\{\theta \in \mathbb{T} \mid (\widetilde{u_h})^*(\theta) > \delta/2\}| \\ &\leqslant \frac{C}{\delta}\|u_h\|_1 \leqslant \frac{C}{\delta}\|f - g\|_1 \to 0 \end{aligned}$$

as $\varepsilon \to 0$. This finishes the case where μ is absolutely continuous relative to Lebesgue measure.

To treat singular measures, we first consider measures $\mu = \nu$ which satisfy $|\mathrm{supp}(\nu)| = 0$. Here

$$\mathrm{supp}(\nu) := \mathbb{T}\backslash \cup \{I \subset \mathbb{T} \mid \nu(I) = 0\}$$

I being an arc. Observe that for any $\theta \notin \mathrm{supp}(\nu)$ the limit $\lim_{r\to 1} \widetilde{u_r}(\theta)$ exists since the analytic function $u + i\tilde{u}$ can be continued across that interval J on $\mathbb{T}$ for which $\mu(J) = 0$ and which contains θ. Hence $\lim_{r\to 1} \widetilde{u_r}$ exist a.e. by the assumption $|\mathrm{supp}(\nu)| = 0$. If $\mu \in \mathcal{M}(\mathbb{T})$ is an arbitrary singular measure, then use inner regularity to say that for every $\varepsilon > 0$ there is $\nu \in \mathcal{M}(\mathbb{T})$ with $\|\mu - \nu\| < \varepsilon$ and $|\mathrm{supp}(\nu)| = 0$. Indeed, set $\nu(A) := \mu(A \cap K)$ for all Borel sets A where K is compact and $|\mu|(\mathbb{T}\backslash K) < \varepsilon$. The theorem now follows by passing from the statement for ν to that for μ by means of the same argument that was used in the absolutely continuous case above. $\square$

It is now easy to obtain the L^p boundedness of the Hilbert transform on $1 < p < \infty$. This result is due to Marcel Riesz, who gave a different proof, see the following exercise for his original argument.

Theorem 4.8. *If $1 < p < \infty$, then $\|Hf\|_p \leqslant C_p \|f\|_p$. Consequently, if $u \in h^p(\mathbb{D})$ with $1 < p < \infty$, then $\tilde{u} \in h^p(\mathbb{D})$ and $\|\tilde{u}\|_p \leqslant C_p \|u\|_p$*

Proof. By Proposition 4.4, $\|Hu\|_2 \leqslant \|u\|_2$ with equality if and only if

$$\int_{\mathbb{T}} u(\theta)\, d\theta = 0$$

Interpolating this with the weak-L^1 bound from Corollary 4.7 by means of the Marcinkiewicz interpolation theorem finishes the case $1 < p \leqslant 2$. If $2 < p < \infty$, then we use duality. More precisely, if $f, g \in L^2(\mathbb{T})$, then

$$\begin{aligned} \langle f, Hg\rangle &= \sum_{n\in\mathbb{Z}} \hat{f}(n)\overline{\widehat{Hg}(n)} = \sum_{n\in\mathbb{Z}} i\mathrm{sign}(n)\hat{f}(n)\overline{\hat{g}(n)} \\ &= \sum_{n\in\mathbb{Z}} -\widehat{Hf}(n)\overline{\hat{g}(n)} = -\langle Hf, g\rangle \end{aligned}$$

This shows that $H^* = -H$. Hence, if $f \in L^p(\mathbb{T}) \subset L^2(\mathbb{T})$ and $g \in L^2(\mathbb{T}) \subset L^{p'}(\mathbb{T})$, then

$$|\langle Hf, g\rangle| = |\langle f, Hg\rangle| \leqslant \|f\|_p \|Hg\|_{p'} \leqslant C_{p'} \|f\|_p \|g\|_{p'}$$

and thus $\|Hf\|_p \leqslant C_{p'} \|f\|_p$ as claimed. $\square$

Exercise 4.2. Give a complex variable proof of the $L^{2m}(\mathbb{T})$-boundedness of the Hilbert transform by applying the Cauchy integral formula to $(u + i\tilde{u})^{2m}$ when m is a positive integer, cf. the proof of Proposition 4.4. Obtain the previous theorem from this by interpolation and duality.

Consider the analytic mapping $F = u + iv$ that takes $\mathbb{D}$ onto the strip

$$\{z \mid |\mathrm{Re}\, z| < 1\}$$

Then $u \in h^\infty(\mathbb{D})$ but clearly $v \notin h^\infty(\mathbb{D})$ so that Theorem 4.8 fails on $L^\infty(\mathbb{T})$. By duality, it also fails on $L^1(\mathbb{T})$. The correct substitute for L^1 in this context is the space of *real parts of functions in* $\mathbb{H}^1(\mathbb{D})$. This is a deep result that goes much further than the F. & M. Riesz theorem. The statement is that

$$\|Hf\|_1 \leqslant C\|u_f^*\|_1 \tag{4.3}$$

where u_f^* is the non-tangential maximal function of the harmonic extension u_f of f. Recall that by the Burkholder-Gundy-Silverstein theorem the right-hand side in (4.3) is finite if and only if f is the real part of an analytic L^1-bounded function. A more modern approach to (4.3) is given by the real-variable theory of Hardy spaces, which subsume statements such as (4.3) by the boundedness of singular integral operators on such spaces.

Next, we turn to the problem of expressing Hf in terms of a kernel. By Exercise 2.3, it is clear that one would expect that

$$(H\mu)(\theta) = \int_{\mathbb{T}} \cot(\pi(\theta-\varphi))\,\mu(d\varphi) \tag{4.4}$$

for any $\mu \in \mathcal{M}(\mathbb{T})$. This, however, requires justification as the integral on the right-hand side is not necessarily convergent.

PROPOSITION 4.9. *If $\mu \in \mathcal{M}(\mathbb{T})$, then*

$$\lim_{\epsilon\to 0}\int_{|\theta-\varphi|>\epsilon} \cot(\pi(\theta-\varphi))\,\mu(d\varphi) = (H\mu)(\theta) \tag{4.5}$$

for a.e. $\theta \in \mathbb{T}$. In other words, (4.4) *holds in the principal value sense.*

EXERCISE 4.3. Verify that the limit in (4.5) exists for all $d\mu = f\,d\theta + d\nu$ where $f \in C^1(\mathbb{T})$ and $|\mathrm{supp}(\nu)| = 0$. Furthermore, show that these measures are dense in $\mathcal{M}(\mathbb{T})$.

PROOF OF PROPOSITION 4.9. The idea is to represent a general measure as a limit of measures of the kind given by the previous exercise. As we have seen in the proof of the a.e. convergence result Theorem 2.8, the double limit appearing in such an argument requires a bound on an appropriate maximal function. In this case the natural bound is of the form

$$\mathrm{mes}\Big[\theta \in \mathbb{T} \mid \sup_{0<\epsilon<\frac{1}{2}} \Big|\int_{|\theta-\varphi|>\epsilon} \cot(\pi(\theta-\varphi))\,\mu(d\varphi)\Big| > \lambda\Big] \leqslant \frac{C}{\lambda}\|\mu\| \tag{4.6}$$

for all $\lambda > 0$. We leave it to the reader to check that (4.6) implies the theorem. In order to prove (4.6) we invoke our strongest result on the conjugate function, namely Theorem 4.6. More precisely, we claim that

$$\sup_{0<r<1}\Big|(Q_r * \mu)(\theta) - \int_{|\theta-\varphi|>1-r} \cot(\pi(\theta-\varphi))\,\mu(d\varphi)\Big| \leqslant CM\mu(\theta) \tag{4.7}$$

where $M\mu$ is the Hardy-Littlewood maximal function. Since

$$\sup_{0<r<1} |(Q_r * \mu)(\theta)| = \widetilde{u_\mu}^*(\theta)$$

(4.6) follows from (4.7) by means of Theorem 4.6 and Proposition 2.5. To verify (4.7) write the difference inside the absolute value signs as $(K_r * \mu)(\theta)$, where

$$K_r(\theta) = \begin{cases} Q_r(\theta) - \cot(\pi\theta) & \text{if} \quad 1-r < |\theta| < \frac{1}{2} \\ Q_r(\theta) & \text{if} \quad |\theta| \leqslant 1-r \end{cases}$$

By means of calculus one checks that

$$|K_r(\theta)| \leqslant \begin{cases} C\frac{(1-r)^2}{|\theta|^3} & \text{if} \quad |\theta| > 1-r \\ C(1-r)^{-1} & \text{if} \quad |\theta| \leqslant 1-r \end{cases}$$

This proves that $\{K_r\}_{0<r<1}$ form a radially bounded approximate identity and (4.7) therefore follows from Lemma 2.7. ∎

Proposition 4.9 raises the question whether Theorem 4.8 can be proved based on the representation (4.5) alone, i.e., without using complex variables or harmonic functions. Going even further, we may ask if the L^2-boundedness of the Hilbert transform can be proved without using the Fourier transform. We shall answer all of these questions later in the context of Calderon-Zygmund theory where they play an important role.

Exercise 4.4. Show by means of (4.5) that H is not bounded on $L^\infty(\mathbb{T})$. For example, consider $H\chi_{[0,\frac{1}{2})}$. Also, deduce from the fact that $\cot(\pi\theta) \notin L^1(\mathbb{T})$ that H is not bounded on L^1.

The following proposition shows that the Hilbert transform Hf is exponentially integrable for bounded functions f. In Exercise 4.4, you should find that $H\chi_{[0,\frac{1}{2}]}$ has logarithmic behavior at 0 and $\frac{1}{2}$, which shows that the following result is best possible.

Proposition 4.10. *Let f be a real-valued function on $\mathbb{T}$ with $|f| \leqslant 1$. Then for any $0 \leqslant \alpha < \frac{\pi}{2}$*

$$\int_{\mathbb{T}} e^{\alpha|Hf(\theta)|}\, d\theta \leqslant \frac{2}{\cos\alpha}$$

Proof. Let $u = u_f$ be the harmonic extension of f to $\mathbb{D}$ and set $F = \tilde{u} - iu$. Then $|u| \leqslant 1$ by the maximum principle and hence $\cos(\alpha u) \geqslant \cos\alpha$. Therefore,

$$\mathrm{Re}\left(e^{\alpha F}\right) = \mathrm{Re}\left(e^{\alpha\tilde{u}} \cdot e^{-i\alpha u}\right) = \cos(\alpha u)e^{\alpha\tilde{u}} \geqslant \cos(\alpha)e^{\alpha\tilde{u}} \tag{4.8}$$

By the mean value property,

$$\int_{\mathbb{T}} \mathrm{Re}\, e^{\alpha F_r(\theta)}\, d\theta = \mathrm{Re}\ e^{\alpha F(0)} = \mathrm{Re}\ e^{-i\alpha(u(0))} = \cos(\alpha u(0)) \leqslant 1$$

Combining this with (4.8) yields

$$\int_{\mathbb{T}} e^{\alpha\tilde{u}_r(\theta)}\, d\theta \leqslant \frac{1}{\cos\alpha}$$

and by Fatou's lemma therefore

$$\int_{\mathbb{T}} e^{\alpha(Hf)(\theta)}\, d\theta \leqslant \frac{1}{\cos\alpha}$$

Since this inequality also holds for $-f$, the proposition follows. □

In later chapters we will develop the real-variable theory of singular integrals which contains the results on the Hilbert transform obtained above. The basic theorem, due to Calderon and Zygmund, states that singular integrals are bounded on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$ thus generalizing Theorem 4.8.

The analogue of Proposition 4.10 for singular integrals is given by the fact that they are bounded from L^∞ to BMO (the space of bounded mean oscillation) and that BMO functions are exponentially integrable (this is called John-Nirenberg inequality). The dual question of what happens on L^1 leads into the real variable theory of Hardy spaces. The analogue of (4.3) is then that singular integrals are bounded on $H^1(\mathbb{R}^n)$. The dual space of H^1 is BMO, a classical result of Charles Fefferman. While the real-variable theory will not be developed in full in this text, a later chapter will present some elements of it by means of a dyadic model.

We conclude the theory of the conjugate function by returning to the issue of $L^p(\mathbb{T})$ convergence of Fourier series. Recall from Chapter 1 that this convergence fails for $p = 1$ and $p = \infty$ but we will now deduce from Theorem 4.8 that it *does hold* for the full range $1 < p < \infty$.

THEOREM 4.11. *Let S_N denote the partial sum operator of Fourier series. Then for any $1 < p < \infty$ the S_N are uniformly bounded on $L^p(\mathbb{T})$, i.e.,*

$$\sup_N \|S_N\|_{p\to p} < \infty$$

By Proposition 1.9 this implies convergence of $S_N f \to f$ in $L^p(\mathbb{T})$ with $1 < p < \infty$ for any $f \in L^p(\mathbb{T})$.

PROOF. The point is simply that S_N can be written in terms of the Hilbert transform. Indeed, recall that

$$\widehat{Hf}(n) = -i\,\mathrm{sign}(n)\hat{f}(n)$$

so that, with $e_n(x) = e(nx)$,

$$Tf := \frac{1}{2}(1 + iH)f = \sum_n \chi_{(0,\infty)}(n)\hat{f}(n)e_n$$

In other words, on the Fourier side T is multiplication by $\chi_{(0,\infty)}$ whereas S_N is multiplication by $\chi_{[-N,N]}$. It remains to write $\chi_{[-N,N]}$ as the difference of two shifted $\chi_{(0,\infty)}$, i.e.,

$$\chi_{[-N,N]} = \chi_{(-N-1,\infty)} - \chi_{(N,\infty)}$$

or in terms of H and T,

$$S_N f = \overline{e_{N+1}}\, T(e_{N+1} f) - e_N\, T(\overline{e_N}\, f)$$

Hence, for $1 < p < \infty$

$$\|S_N\|_{p\to p} \leqslant 2\|T\|_{p\to p} \leqslant 1 + \|H\|_{p\to p}$$

uniformly in N, as claimed. □

We invite the reader to investigate the analogous result for higher-dimensional Fourier series.

EXERCISE 4.5. Let $\underline{N} = (N_1, \ldots, N_d)$ denote a vector with positive integer entries. Prove that

$$\sup_{\underline{N}} \|S_{\underline{N}}\|_{p\to p} < \infty \qquad \forall\ 1 < p < \infty$$

where $S_{\underline{N}}$ is the partial sum operator for Fourier series on $\mathbb{T}^d$ defined as follows:

$$S_{\underline{N}} f(x) = \sum_{\nu\in\mathbb{Z}^d} \prod_{j=1}^{d} \chi_{[|\nu_j|<N_j]}\hat{f}(\nu)e(\nu\cdot x)$$

Conclude that $S_{\underline{N}} f$ converges in the L^p sense for any $f \in L^p(\mathbb{T}^d)$ and $1 < p < \infty$ provided each component of $\underline{N} \to \infty$.

We remark that the corresponding result in which expanding rectangles are replaced with Euclidean balls in $\mathbb{Z}^d$ is false; in fact, Charlie Fefferman's ball multiplier theorem states that the analogue of the previous exercise only holds for $p = 2$ in that case. This is delicate, and relies on the existence of *Kakeya sets*. We shall return to these matters in a later chapter.

Notes

The conjugate function and the Hilbert transform provide the link between the complex-variable based theory of Hardy spaces and the real-variable Calderón-Zygmund theory which we turn to in a later chapter. Koosis [**34**] proves the Burkholder-Gundy-Silverstein theorem. This reference also presents some elements of the modern real-variable theory of H^1 spaces, such as the atomic decomposition of H^1. The full real-variable H^p theory in higher dimensional is developed in Stein's book [**46**]. We shall discuss some of the basic ideas of that theory by means of a dyadic model in a later chapter.

Problems

Problem 4.1. Find an expression for the harmonic measure of an arc on $\mathbb{T}$ relative to the disk $\mathbb{D}$. Relate it to the harmonic measure of an interval relative to the upper half plane by means of a conformal mapping.

Problem 4.2. This problem provides a weak form of a.e. convergence of Fourier series on $L^1(\mathbb{T})$. In the following chapter we will see via Kolmogoroff's construction that the full form of the a.e. convergence fails for L^1 functions.

a) Show that, for any $\lambda > 0$

$$\sup_N |\{\theta \in \mathbb{T} \mid |(S_N f)(\theta)| > \lambda\}| \leqslant \frac{C}{\lambda} \|f\|_1$$

with some absolute constant C.

b) What would such an inequality mean with the $\sup_N$ inside, i.e.,

$$|\{\theta \in \mathbb{T} \mid \sup_N |(S_N f)(\theta)| > \lambda\}| \leqslant \frac{C}{\lambda} \|f\|_1$$

for all $\lambda > 0$? Can this be true?

c) Using *a*) show that for every $f \in L^1(\mathbb{T})$ there exists a subsequence $\{N_j\} \longrightarrow \infty$ depending on f such that

$$S_{N_j} f \longrightarrow f \ a.e.$$

as $j \to \infty$.

Problem 4.3. Prove the following multiplier estimate. Let $\bar{m} := \{m_n\}_{n \in \mathbb{Z}}$ be a sequence in $\mathbb{C}$ satisfying

$$\sum_{n \in \mathbb{Z}} |m_n - m_{n-1}| \leqslant B, \qquad \lim_{n \to -\infty} m_n = 0$$

Define the multiplier operator $T_{\tilde{m}}$ by the rule

$$T_{\tilde{m}}f(\theta) = \sum_n m_n \hat{f}(n)\, e(n\theta)$$

for trigonometric polynomials f. Prove that for any such f one has

$$\|T_{\tilde{m}}f\|_p \leqslant C(p)B\|f\|_p$$

for any $1 < p < \infty$.

Problem 4.4. Let $\varphi \in C^\infty(\mathbb{T})$ and denote by A_φ the multiplication operator by φ. With H the Hilbert transform on $\mathbb{T}$, show that

$$[A_\varphi, H] = A_\varphi \circ H - H \circ A_\varphi$$

is a *smoothing* operator, i.e., if $\mu \in \mathcal{M}(\mathbb{T})$ is an arbitrary measure, then

$$[A_\varphi, H]\mu \in C^\infty(\mathbb{T})$$

Problem 4.5. The complex variable methods developed in Chapters 2, 3, as well as this one equally well apply to the upper half plane instead of the disk. For example, the Poisson kernel is

$$P_t(x) = \frac{1}{\pi}\frac{t}{x^2+t^2} \qquad x \in \mathbb{R},\ t > 0$$

and its harmonic conjugate is

$$Q_t(x) = \frac{1}{\pi}\frac{x}{x^2+t^2}$$

Observe that $P_t(x) + iQ_t(x) = \frac{i}{\pi}\frac{1}{x+it} = \frac{i}{\pi z}$ with $z = x + it$, which should remind the reader of Cauchy's formula. The Hilbert transform now reads

$$(Hf)(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(y)}{x-y}\, dy$$

in the principal value sense, the precise statement being just as in Proposition 4.9.

In this problem, the reader is asked to develop the theory of the upper half-plane in the context of the previous chapters. The relation between the disk and upper half-plane given by conformal mapping should be considered as well.

Problem 4.6. By considering $F = (u + i\tilde{u})\log(1 + u + i\tilde{u})$ show that if $u \geqslant 0$ and $\tilde{u} \in h^1(\mathbb{D})$, then $\int_{\mathbb{T}} u \log(1+u)\, d\theta < \infty$.

CHAPTER 5

Fourier Series on $L^1(\mathbb{T})$: Pointwise Questions

This chapter is devoted to the question of almost everywhere convergence of Fourier series for $L^1(\mathbb{T})$ functions. From the previous chapters we know that this is not an elementary question, since the partial sums S_N are given by convolution with the Dirichlet kernels D_N which *do not* form an approximate identity. Note also that we encountered an a.e. convergence question which was not associated with an approximate identity, namely for the Hilbert transform, see Proposition 4.9. However, the needed uniform control was provided by the weak-L^1 bound on the maximal function $\tilde{u}^*$, see Theorem 4.6.

By a natural analogy, one sees that a similar bound on the maximal function associated with $S_N f$, i.e., $\sup_N |S_N f|$, would lead to an a.e. convergence result.

What Kolmogoroff realized is essentially that no bound on this maximal function is possible for $f \in L^1(\mathbb{T})$. This has to do with the oscillatory nature of the Dirichlet kernel D_N and the growth of $\|D_N\|_{L^1}$. In fact, by a careful choice of $f \in L^1$ we may arrange the convolution $(D_N * f)(x)$ so that only the positive peaks of D_N contribute at many points x.

What Stein [**44**] realized is that a weak-type bound on $\sup_N |(S_N f)(x)|$ is not only sufficient but also *necessary* for the a.e. convergence. In view of this fact we will conclude indirectly that $S_N f$ fails to converge a.e. for some $f \in L^1(\mathbb{T})$. Kolmogoroff's original argument involved the construction of a function $f \in L^1(\mathbb{T})$ for which a.e. convergence failed. However, what we do here is very close to his original approach even though it is somewhat more streamlined.

Later in this book we will take up to more challenging question of a.e. convergence of $S_N f$ for $f \in L^2(\mathbb{T})$ or $L^p(\mathbb{T})$ for $1 < p < \infty$. This is Carleson's theorem, and the proof introduces some essentially new and far-reaching ideas.

Next to presenting Kolomogoroff's theorem, this chapter also serves to introduce ideas based on the use of *randomness*. This turns out to be a very powerful device for many purposes. We begin the actual argument with such a probabilistic construction, which is a version of the Borel-Cantelli lemma (but the following proof is self-contained).

LEMMA 5.1. *Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of measurable subsets of $\mathbb{T}$ such that*

$$\sum_{n=1}^{\infty} |E_n| = \infty$$

Then there exists a sequence $\{x_n\}_{n=1}^\infty \in \mathbb{T}$ *so that*

$$\sum_{n=1}^{\infty} \chi_{E_n}(x + x_n) = \infty \tag{5.1}$$

for almost every $x \in \mathbb{T}$.

Proof. View $\Omega := \prod_{n=1}^\infty \mathbb{T}$ as a probability space equipped with the infinite product measure. Given $x \in \mathbb{T}$, let $A_x \subset \Omega$ be the event characterized by (5.1). We claim that

$$\mathbb{P}(A_x) = 1 \quad \forall x \in \mathbb{T} \tag{5.2}$$

By Fubini, it then follows that for almost every $\{x_n\}_{n=1}^\infty \in \Omega$, the event (5.1) holds for almost every $x \in \mathbb{T}$. Hence fix an arbitrary $x \in \mathbb{T}$. Then

$$\begin{aligned} A_x &= \Big\{\{x_n\}_{n=1}^\infty \mid x \text{ belongs to infinitely many } E_n(\cdot + x_n)\Big\} \\ &= \Big\{\{x_n\}_{n=1}^\infty \mid x \in \bigcap_{N=1}^\infty \bigcup_{m=N}^\infty (x_m + E_m)\Big\} \\ &= \bigcap_{N=1}^\infty A_N \end{aligned}$$

where

$$\begin{aligned} A_N^c &:= \Big\{\{x_n\}_{n=1}^\infty \mid x \in \bigcap_{m=N}^\infty (x_m + E_m^c)\Big\} \\ &= \Big\{\{x_n\}_{n=1}^\infty \mid x_m \in x - E_m^c \quad \forall m \geqslant N\Big\} \end{aligned}$$

By definition of the product measure on Ω, it therefore follows that

$$\mathbb{P}(A_N^c) = \prod_{m=N}^\infty (1 - |E_m|) = 0$$

by assumption (5.1). Hence (5.2) holds, and the lemma follows. □

Lemma 5.1 will play an important role in the proof of the following theorem, which establishes the necessity of a weak-type maximal function bound for a.e. convergence. In addition to the previous probabilistic lemma, we shall also require the basic *Kolmogoroff* 0, 1 *law* which we now recall: let $\{X_k\}_{k=1}^\infty$ be independent random variables on a probability space $(\Omega, \mathbb{P})$. Assume that $E \subset \Omega$ is measurable with respect to the tail σ-algebra. This means that E is measurable with respect to the σ-algebra generated by $\{X_k\}_{k=K}^\infty$ for every positive integer K. Then either $\mathbb{P}(E) = 0$ or $\mathbb{P}(E) = 1$. Such an E is called a *tail event*. The proof of this law proceeds by showing that E is independent of itself whence

$$0 = \mathbb{P}(E \cap E^c) = \mathbb{P}(E)(1 - \mathbb{P}(E))$$

yielding the desired conclusion.

In our case, the Carleson maximal operator

$$Cf(x) := \sup_N |S_N f(x)| \tag{5.3}$$

is the main object that needs to be bounded. However, we shall formulate the following "abstract" result in much greater generality. In fact, the reader will easily see that it has nothing to do with any special properties of $\mathbb{T}$, either, although we restrict ourselves to that case for simplicity.

Theorem 5.2. *Let T_n be a sequence of linear operators, bounded on $L^1(\mathbb{T})$, and translation invariant. Define*

$$\mathcal{M}f(x) := \sup_{n\geqslant 1} |T_n f(x)|$$

and assume that $\|\mathcal{M}f\|_\infty < \infty$ for every trigonometric polynomial f on $\mathbb{T}$.

If for any $f \in L^1(\mathbb{T})$,

$$|\{x \in \mathbb{T} \mid \mathcal{M}f(x) < \infty\}| > 0$$

then there exists a constant A so that

$$|\{x \in \mathbb{T} \mid \mathcal{M}f(x) > \lambda\}| \leqslant \frac{A}{\lambda}\|f\|_1$$

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

Proof. We will prove this by contradiction. Thus, assume that there exists a sequence $\{f_j\}_{j=1}^\infty \subset L^1(\mathbb{T})$ with $\|f_j\|_1 = 1$ for all $j \geqslant 1$, as well as $\lambda_j > 0$ so that

$$E_j := \{x \in \mathbb{T} \mid \mathcal{M}f_j(x) > \lambda_j\}$$

satisfies

$$|E_j| > \frac{2^j}{\lambda_j}$$

for each $j \geqslant 1$. Be definition of $\mathcal{M}$ we then also have

$$\lim_{m\to\infty} |\{x \in \mathbb{T} \mid \sup_{1\leqslant k\leqslant m} |T_k f_j(x)| > \lambda_j\}| > \frac{2^j}{\lambda_j}$$

for each $j \geqslant 1$. Hence there are $M_j < \infty$ with the property that

$$|\{x \in \mathbb{T} \mid \sup_{1\leqslant k\leqslant M_j} |T_k f_j(x)| > \lambda_j\}| > \frac{2^j}{\lambda_j}$$

for each $j \geqslant 1$. Let σ_N denote the N^{th} Cesàro sum, i.e., $\sigma_N f = K_N * f$, where K_N is the Fejer kernel. Since each T_j is bounded on L^1, we conclude that

$$\lim_{N\to\infty} |\{x \in \mathbb{T} \mid \sup_{1\leqslant k\leqslant M_j} |T_k \sigma_N f_j(x)| > \lambda_j\}| > \frac{2^j}{\lambda_j}$$

Hence, we assume from now on that each f_j is a trigonometric polynomial. Let m_j be a positive integer with the property that

$$m_j \leqslant \frac{\lambda_j}{2^j} < m_j + 1$$

Then

$$\sum_{j=1}^{\infty} m_j |E_j| = \infty$$

by construction. Counting each of the sets E_j with multiplicity m_j, the previous lemma implies that there exists a sequence of points $x_{j,\ell}$, $j \geqslant 1$, $1 \leqslant \ell \leqslant m_j$, so that

$$\sum_{j=1}^{\infty} \sum_{\ell=1}^{m_j} \chi_{E_j}(x - x_{j,\ell}) = \infty \tag{5.4}$$

for almost every $x \in \mathbb{T}$. Let

$$\delta_j := \frac{1}{j^2 m_j}$$

and define

$$f(x) := \sum_{j=1}^{\infty} \sum_{\ell=1}^{m_j} \pm \delta_j f_j(x - x_{j,\ell})$$

where the signs $\pm$ will be chosen randomly. First note that irrespective of the choice of these signs,

$$\|f\|_1 \leqslant \sum_{j=1}^{\infty} m_j \delta_j < \infty$$

The point is now to choose the signs so that

$$\mathcal{M} f(x) = \infty$$

for almost every $x \in \mathbb{T}$. For this purpose, select $x \in \mathbb{T}$ such that $x \in E_j + x_{j,\ell}$ for infinitely many j and $\ell = \ell(j)$ (just pick one such $\ell(j)$ if there are more than one). Denote the set of these j by $\mathcal{J}$. Since the T_n are translation invariant, so is $\mathcal{M}$. Hence

$$\mathcal{M} f_j(x - x_{j,\ell}) > \lambda_j$$

for all $j \in \mathcal{J}$. We conclude that for those j there is a positive integer $n(j,x)$ so that

$$|T_{n(j,x)} f_j(x - x_{j,\ell})| > \lambda_j$$

At the cost of removing another set of measure zero we may assume that

$$T_n f(x) = \sum_{j=1}^{\infty} \sum_{\ell=1}^{m_j} \pm \delta_j T_n f(x - x_{j,\ell})$$

for all positive integers n. In particular, we have that for all $j \in \mathcal{J}$

$$T_{n(j,x)} f(x) = \sum_{k=1}^{\infty} \sum_{\ell=1}^{m_k} \pm \delta_k T_{n(j,x)} f(x - x_{k,\ell})$$

which implies that

$$\mathbb{P}[|T_{n(j,x)} f(x)| > \delta_j \lambda_j] \geqslant \frac{1}{2}$$

where the probability measure is with respect to the choice of signs $\pm$. Since $\delta_j \lambda_j \to \infty$, we obtain that

$$\mathbb{P}[\,\mathcal{M}f(x) = \infty] \geqslant \frac{1}{2}$$

We claim that the event on the left-hand side is a tail event. Indeed, this holds since

$$\mathcal{M}f(x) \leqslant \sum_{j=1}^{\infty} \mathcal{M}f_j(x)$$

and each summand on the right-hand side here is finite (f_j is a trigonometric polynomial and we are assuming that $\mathcal{M}$ is uniformly bounded on trigonometric polynomials). By Kolmogoroff's zero-one law we therefore have

$$\mathbb{P}[\,\mathcal{M}f(x) = \infty] = 1$$

It follows from Fubini's theorem that almost surely (in the choice of $\pm$)

$$\mathcal{M}f(x) = \infty \text{ for a.e. } x$$

This would contradict our main hypothesis, and we are done. □

The previous lemma reduces Kolmogoroff's theorem on the failure of almost everywhere convergence of Fourier series of L^1 functions to disproving a weak-L^1 bound for the Carleson maximal operator. More precisely, we arrive at the following corollary. It will be technically more convenient later on to formulate this result in terms of measures rather than on $L^1(\mathbb{T})$.

Corollary 5.3. *Suppose $\{S_N f\}_{N=1}^{\infty}$ converges almost everywhere for every $f \in L^1(\mathbb{T})$. Then there exists a constant A such that*

$$|\{x \in \mathbb{T} \mid C\mu(x) > \lambda\}| \leqslant \frac{A}{\lambda}\|\mu\| \tag{5.5}$$

for any complex Borel measure μ on $\mathbb{T}$ and $\lambda > 0$, where C is as in (5.3).

Proof. By the previous theorem, our assumption implies that there exists a constant A such that

$$|\{x \in \mathbb{T} \mid Cf(x) > \lambda\}| \leqslant \frac{A}{\lambda}\|f\|_1$$

for all $\lambda > 0$. This is the same as

$$|\{x \in \mathbb{T} \mid \max_{1 \leqslant n \leqslant N} |S_n f(x)| > \lambda\}| \leqslant \frac{A}{\lambda}\|f\|_1$$

for all $N \geqslant 1$ and $\lambda > 0$. If μ is a complex measure, then we set $f = V_m * \mu$, where V_m is de la Vallée-Poussin's kernel, see (1.22). It follows that for all $N \geqslant 1$

$$|\{x \in \mathbb{T} \mid \max_{1 \leqslant n \leqslant N} |S_n[V_m * \mu](x)| > \lambda\}| \leqslant \frac{A}{\lambda}\|\mu\|$$

for all $m \geqslant 1$ and $\lambda > 0$. Note that for $m \geqslant N$ the kernel V_m can be removed from the left-hand side, and we obtain our desired conclusion. □

The idea behind Kolmogoroff's theorem is to find a measure μ which would violate (5.5). This measure will be chosen to create "resonances", i.e., so that the peaks of the Dirichlet kernel all appear with the same sign. More precisely, for every positive integer N we will choose

$$\mu_N := \frac{1}{N}\sum_{j=1}^{N}\delta_{x_{j,N}} \tag{5.6}$$

where the $x_{j,N}$ are close to $\frac{j}{N}$. Then

$$(S_n\mu_N)(x) = \frac{1}{N}\sum_{j=1}^{N}\frac{\sin((2n+1)\pi(x-x_{j,N}))}{\sin(\pi(x-x_{j,N}))}. \tag{5.7}$$

If x is fixed, we will then argue that there exists n so that the summands on the right-hand have the same sign (for this, we will need to make a careful choice of the $x_{j,N}$). Thus, the size of the entire sum will be about

$$\sum_{j=1}^{N}\frac{1}{j} \simeq \log N$$

because of the denominators in (5.7), which clearly contradicts (5.5).

The choice of the points $x_{j,N}$ is based on the following lemma due to Kronecker.

Lemma 5.4. *Assume that $(\theta_1,\ldots,\theta_d)\in\mathbb{T}^d$ is* incommensurate*, i.e., that for any $(n_1,\ldots,n_d)\in\mathbb{Z}^d\setminus\{0\}$ one has that*

$$n_1\theta_1+\ldots+n_d\theta_d\notin\mathbb{Z}.$$

Then the orbit

$$\big\{(n\theta_1,\ldots,n\theta_d) \mod \mathbb{Z}^d \mid n\in\mathbb{Z}\big\}\subset\mathbb{T}^d \tag{5.8}$$

is dense in $\mathbb{T}^d$.

Proof. It will suffice to show that for any smooth function f on $\mathbb{T}^d$

$$\frac{1}{N}\sum_{n=1}^{N}f(n\theta_1,\ldots,n\theta_d)\to\int_{\mathbb{T}^d}f(x)\,dx. \tag{5.9}$$

Indeed, if the orbit (5.8) is not dense, then we could find $f\geqslant 0$ so that the set $\{\mathbb{T}^d|f>0\}$ does not intersect it. Cleary, this would contradict (5.9). To prove (5.9), expand f into a Fourier series, cf. Proposition 1.14. Then

$$\begin{aligned}\frac{1}{N}\sum_{n=1}^{N}f(n\theta_1,\ldots,n\theta_d) &= \frac{1}{N}\sum_{n=1}^{N}\sum_{\nu\in\mathbb{Z}^d}\hat{f}(\nu)e^{2\pi in\theta\cdot\nu}\\ &= \hat{f}(0)+\sum_{\nu\in\mathbb{Z}^d\setminus\{0\}}\hat{f}(\nu)\frac{1}{N}\frac{1-e^{2\pi i(N+1)\theta\cdot\nu}}{1-e^{2\pi i\theta\cdot\nu}}\end{aligned}$$

where the ratio on the right-hand side is well-defined by our assumption. Clearly, $\hat{f}(\nu)$ is rapidly decaying in $|\nu|$ since $f \in C^\infty(\mathbb{T}^d)$. Thus, since

$$\Big|\frac{1}{N}\frac{1-e^{2\pi i(N+1)\theta\cdot\nu}}{1-e^{2\pi i\theta\cdot\nu}}\Big| \leqslant 1$$

for all $N \geqslant 1$, $\nu \neq 0$, it follows that

$$\lim_{N\to\infty}\sum_{\nu\in\mathbb{Z}^d\setminus\{0\}} \hat{f}(\nu)\frac{1}{N}\frac{1-e^{2\pi i(N+1)\theta\cdot\nu}}{1-e^{2\pi i\theta\cdot\nu}} = 0$$

and we are done. □

We can now carry out our construction of the μ_N.

Lemma 5.5. *There exists a sequence μ_N of probability measures on $\mathbb{T}$ with the property that*

$$\limsup_{n\to\infty}\,(\log N)^{-1}|S_n\mu_N(x)| > 0$$

for almost every $x \in \mathbb{T}$.

Proof. For every $N \geqslant 1$ and $1 \leqslant j \leqslant N$ choose $x_{j,N} \in \mathbb{T}$ which satisfy

$$|x_{j,N} - \frac{j}{N}| \leqslant N^{-2}$$

and so that $\{x_{j,N}\}_{j=1}^N \in \mathbb{T}^N$ is an incommensurate vector. This can be done since the commensurate vectors have measure zero. Clearly, the set of $x \in \mathbb{T}$ such that $\{2(x - x_{j,N})\}_{j=1}^N \in \mathbb{T}^N$ is a commensurate vector is at most countable. It follows that for almost every $x \in \mathbb{T}$,

$$\big\{\{2n(x - x_{j,N})\}_{j=1}^N \mod \mathbb{Z}^N \mid n \in \mathbb{Z}\big\}$$

is dense in $\mathbb{T}^N$. Hence, for almost every $x \in \mathbb{T}$,

$$\big\{\{(2n+1)(x - x_{j,N})\}_{j=1}^N \mod \mathbb{Z}^N \mid n \in \mathbb{Z}\big\}$$

is also dense in $\mathbb{T}^N$. It follows that for almost every $x \in \mathbb{T}$ there are infinitely many choices of $n \geqslant 1$ so that

$$\sin((2n+1)\pi(x - x_{j,N})) \geqslant \frac{1}{2}$$

for all $1 \leqslant j \leqslant N$. In particular, for those n the sum in (5.7) satisfies

$$|S_n\mu_N(x)| \geqslant \frac{1}{2N}\sum_{j=1}^N \frac{1}{|\sin(\pi(x - x_{j,N}))|} \geqslant C\,\frac{1}{N}\sum_{j=1}^N \frac{1}{j/N} \geqslant C\,\log N,$$

as desired. □

Finally, combining Lemma 5.5 with Corollary 5.3 yields Kolmogoroff's theorem. The second statement in the following theorem follows from Theorem 5.2.

THEOREM 5.6. *There exists $f \in L^1(\mathbb{T})$ so that $S_n f$ does not converge almost everywhere. In fact, there exists $f \in L^1(\mathbb{T})$ such that*

$$\{x \in \mathbb{T} \mid Cf(x) < \infty\}$$

has measure 0.

It is known that this statement also holds with everywhere divergence. In fact, Konyagin [**35**] showed the following stronger result: Let $\phi : [0, \infty) \to [0, \infty)$ be non-decreasing and satisfy $\phi(u) = o(u\sqrt{\log u}/\sqrt{\log\log u})$ as $u \to \infty$. Then there exists $f \in L^1(\mathbb{T})$ so that

$$\int_{\mathbb{T}} \phi(|f(x)|)\, dx < \infty$$

and $\limsup_{m\to\infty} |S_m f(x)| = \infty$ for every $x \in \mathbb{T}$.

Notes

Kolmogoroff's original reference is [**33**]. Much more on pointwise and a.e. convergence questions can be found in Zygmund [**58**] and Katznelson [**30**].

Problems

PROBLEM 5.1. Consider the following variants of Theorem 5.2:

1) Let $\{\mu_n\}_{n=1}^\infty$ be a sequence of positive measures on $\mathbb{R}^d$ supported in a common compact set. Define
$$\mathcal{M}f(x) := \sup_{n\geqslant 1} |(f * \mu_n)(x)|$$
Let $1 \leqslant p < \infty$ and assume that for each $f \in L^p(\mathbb{R}^d)$
$$\mathcal{M}f(x) < \infty \text{ on a set of positive measure.}$$
Then show that $f \mapsto \mathcal{M}f$ is of weak-type (p, p).
2) Now suppose μ_n are *complex* measures of the form $\mu_n(dx) = K_n(x)\, dx$, but again with common compact support. Show that the conclusion of part 1) holds, but only for $1 \leqslant p \leqslant 2$.

PROBLEM 5.2. If ω is an irrational number, show that

$$\Big\| \frac{1}{N} \sum_{n=1}^N f(\cdot + n\omega) - \int_{\mathbb{T}} f(\theta)\, d\theta \Big\|_{L^2} \to 0$$

for any $f \in L^2(\mathbb{T})$. In particular, if $f \in L^2$ is such that $f(x + \omega) = f(x)$ for a.e. x, then f = const.

PROBLEM 5.3. Let $\{x_n\}_{n=1}^\infty$ be an infinite sequence of real numbers. Show that the following three conditions are equivalent:

a) For any $f \in C(\mathbb{T})$,
$$\frac{1}{N} \sum_{n=1}^N f(x_n) \to \int_{\mathbb{T}} f(x)\, dx$$

b) $\frac{1}{N}\sum_{n=1}^{N} e(kx_n) \to 0$ for all $k \in \mathbb{Z}^+$

c) $\lim_{N\to\infty} \sup_{I\subset\mathbb{T}} \left|\frac{1}{N}\#\{1 \leqslant j \leqslant N \mid x_j \in I \mod 1\} - |I|\right| = 0$

If these conditions hold we say that $\{x_n\}_{n=1}^{\infty}$ is *uniformly distributed* modulo 1 (abbreviated by u.d.). The quantity

$$D\big(\{x_j\}_{j=1}^N\big) := \sup_{I\subset\mathbb{T}} \big|\#\{1 \leqslant j \leqslant N \mid x_j \in I \mod 1\} - N|I|\big|$$

is called the *discrepancy* of the finite sequence $\{x_j\}_{j=1}^N$.

Problem 5.4. Using Problem 1.4 with a suitable choice of T, prove the following: if $\{x_n\}_{n=1}^{\infty}$ is a sequence for which $\{x_{n+k} - x_n\}_{n=1}^{\infty}$ is u.d. modulo 1 for any $k \in \mathbb{Z}^+$, then $\{x_n\}_1^{\infty}$ is also u.d. mod 1. In particular, show that $\{n^d\omega\}_{n=1}^{\infty}$ is u.d. mod 1 for any irrational ω and $d \in \mathbb{Z}^+$.

Problem 5.5. Let $p \geqslant 2$ be a positive integer. Show that for a.e. $x \in \mathbb{T}$ the sequence $\{p^n x\}_{n=1}^{\infty}$ is u.d. modulo 1. Can you characterize those x which have this property? *Hint:* Consider expansions to power p.

Problem 5.6. Let $\{\theta_j\}_{j=1}^N \subset \mathbb{T}$ be N distinct points and consider

$$P(z) = \prod_{j=1}^{N}(z - e(\theta_j))$$

Assume that $\sup_{|z|=1}|P(z)| \leqslant e^{\tau}$. We may assume that $1 \leqslant \tau \leqslant N\log 2$. Prove the following bound on the discrepancy of $\{\theta_j\}_{j=1}^N$ as defined above:

$$\Delta_N := D\big(\{\theta_j\}_{j=1}^N\big) \leqslant C\sqrt{\tau N}$$

Conclude that

$$\big|\{\theta \in \mathbb{T} \mid |P(e(\theta))| < e^{-AN}\}\big| \leqslant Ce^{-\frac{AN}{\Delta_N}} \leqslant Ce^{-\frac{A\sqrt{N}}{\tau}}$$

for any $A \geqslant 0$. The constants C are absolute but may differ. *Hint:* Consider the function $F(\theta) = \mu([\theta_0, \theta)) - N(\theta - \theta_0)$ with $\mu = \sum_{j=1}^N \delta_{\theta_j}$ where $\theta_0 \in (-\frac{1}{2}, \frac{1}{2})$ is chosen such that $\int_{-\frac{1}{2}}^{\frac{1}{2}} F(\theta)\, d\theta = 0$. Relate F to to $u(\theta) := \log|P(e(\theta))|$ via the Hilbert transform. Then write $F = K * \mu$ with a suitable kernel K.

CHAPTER 6

Fourier Analysis on $\mathbb{R}^d$

It is natural to extend the definition of the Fourier transform for the circle given in Chapter 1 to Euclidean spaces. Thus, let $\mu \in \mathcal{M}(\mathbb{R}^d)$ be a complex-valued Borel measure and set

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi}\,\mu(dx) \qquad \forall\, \xi \in \mathbb{R}^d$$

which is called the Fourier transform of μ. Clearly, $\|\hat{\mu}\|_\infty \leqslant \|\mu\|$, the total variation of μ. Moreover, it follows immediately from the dominated convergence theorem that $\hat{\mu}$ is continuous. If $\mu(dx) = f(x)\,dx$, then we also write $\hat{f}(\xi)$. In particular, the Fourier transform is well-defined on $L^1(\mathbb{R}^d)$ and defines a contraction into $C(\mathbb{R}^d)$, the space of continuous functions. Many elementary properties from Chapter 1 carry over to this setting, such as convolutions, Young's inequality, and the action of the Fourier transform on convolutions and translations. To be precise, we collect a few facts in the following lemma.

LEMMA 6.1. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$ and let τ_y denote translation by $y \in \mathbb{R}^d$: $(\tau_y\mu)(E) = \mu(E-y)$. Then*

$$\widehat{\tau_y\mu}(\xi) = e^{-2\pi i\xi\cdot y}\hat{\mu}(\xi) \qquad \forall\, \xi \in \mathbb{R}^d$$

Let $e_\eta(x) := e^{2\pi i x\cdot\eta}$. Then $\widehat{e_\eta\mu}(\xi) = \hat{\mu}(\xi-\eta)$. If $f, g \in L^1(\mathbb{R}^d)$, then the integral

$$\int_{\mathbb{R}^d} f(x-y)g(y)\,dy$$

is absolutely convergent for a.e. $x \in \mathbb{R}^d$ and is denoted by $(f*g)(x)$. One has $f*g \in L^1(\mathbb{R}^d)$, and $\widehat{f*g} = \hat{f}\hat{g}$.

Finally, if $g \in L^p$ for $1 \leqslant p \leqslant \infty$ and $f \in L^1$, then $f*g \in L^p$ and $\|f*g\|_p \leqslant \|f\|_1\|g\|_p$.

We leave the elementary proof to the reader. We record one more very important fact, namely the *change of variables formula* for the Fourier transform: let A be an invertible $d\times d$-matrix with real entries, and denote by $f\circ A$ the function $f(Ax)$. Then for any $f \in L^1(\mathbb{R}^d)$ one has

$$\widehat{f\circ A} = |\det A|^{-1}\hat{f}\circ A^{-t} \tag{6.1}$$

where A^{-t} is the transpose of the inverse of A. A special case of this is the dilation identity. For any $\lambda > 0$ let $f_\lambda(x) = f(\lambda x)$. Then

$$\widehat{f_\lambda}(\xi) = \lambda^{-d}\hat{f}(\xi/\lambda) \tag{6.2}$$

In analogy with the Fourier transform on the circle, we expect that $L^2(\mathbb{R}^d)$ plays a special role. The be more precise, we expect the Plancherel theorem $\|\hat{f}\|_2 = \|f\|_2$ to hold. Notice that as it stands this is only meaningful for $f \in L^1 \cap L^2(\mathbb{R}^d)$, since otherwise the defining integral is not absolutely convergent. In fact, it is convenient to work with a much smaller space called *Schwartz* space. In what follows, we use the standard multi-index notation: for any $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{Z}_0^d$ (where $\mathbb{Z}_0$ are the nonnegative integers)

$$x^\alpha := \prod_{j=1}^d x_j^{\alpha_j}, \qquad x = (x_1, x_2, \ldots, x_d)$$

$$\partial^\alpha := \prod_{j=1}^d \frac{\partial^{\alpha_j}}{\partial x_j^{\alpha_j}}$$

Moreover, $|\alpha| = \sum_{j=1}^d \alpha_j$ denotes the *length* of the multi-index.

Definition 6.2. The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is defined as all functions in $C^\infty(\mathbb{R}^d)$ which decay rapidly together with all derivatives. In other words, $f \in \mathcal{S}(\mathbb{R}^d)$ if and only if $f \in C^\infty(\mathbb{R}^d)$ and

$$x^\alpha \partial^\beta f(x) \in L^\infty(\mathbb{R}^d) \qquad \forall\ \alpha, \beta$$

where α, β are arbitrary multi-indices. We introduce the following notion of convergence in $\mathcal{S}(\mathbb{R}^d)$: a sequence $f_n \in \mathcal{S}(\mathbb{R}^d)$ converges to $g \in \mathcal{S}(\mathbb{R}^d)$ if and only if

$$\|x^\alpha \partial^\beta (f_n - g)\|_\infty \to 0 \qquad n \to \infty$$

for all α, β.

For readers familiar with Fréchet spaces, we remark that the *semi-norms*

$$p_{\alpha\beta}(f) := \|x^\alpha \partial^\beta f\|_\infty \qquad \alpha, \beta \in \mathbb{Z}_0^d$$

turn $\mathcal{S}(\mathbb{R}^d)$ into such a space. Since Fréchet spaces are metrizable, it follows that the notion of convergence introduced in Definition 6.2 completely characterizes the topology in $\mathcal{S}(\mathbb{R}^d)$. However, we make no use of this fact, and only work with the sequential characterization.

One has that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$ for all $1 \leqslant p < \infty$ since this is already the case for smooth functions with compact support. It is evident that the map $f \mapsto x^\alpha \partial^\beta f(x)$ is continuous on $\mathcal{S}(\mathbb{R}^d)$. Somewhat less obvious is that the Fourier transform has the same property.

Proposition 6.3. *The Fourier transform is a continuous operation from the Schwartz space into itself.*

Proof. Fix any $f \in \mathcal{S}(\mathbb{R}^d)$. The proof hinges on the pointwise identities

$$\widehat{\partial^\alpha f}(\xi) = (2\pi i)^{|\alpha|}\ \xi^\alpha \hat{f}(\xi)$$

$$\partial^\beta \hat{f}(\xi) = (-2\pi i)^{|\beta|}\ \widehat{x^\beta f}(\xi)$$

valid for all multi-indices α, β. The first one follows by integration by parts, whereas the second is obtained by differentiating under the integral sign. We leave the formal verification of these identities to the reader (introduce difference quotients, justify the limits by the dominated convergence theorem etc.). This is easy, due to the rapid decay and smoothness of f.

We first note from these identities that $\hat{f} \in C^\infty$ and, moreover, that

$$\xi^\alpha \partial_\xi^\beta \hat{f} \in L^\infty(\mathbb{R}^d) \qquad \forall\ \alpha, \beta \in \mathbb{Z}_0^d$$

which implies by definition that $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$. For the continuity, let f_n be a sequence in $\mathcal{S}(\mathbb{R}^d)$ so that $f_n \to 0$ as $n \to \infty$. Then

$$\|\xi^\alpha \partial_\xi^\beta \hat{f}_n\|_\infty \leqslant C_{\alpha,\beta}\, \|\partial_x^\alpha x^\beta f_n\|_1 \to 0$$

as $n \to \infty$. □

In view of the preceding fact, it is most natural to ask if the Fourier transform takes $\mathcal{S}(\mathbb{R}^d)$ *onto* itself. We will now prove that this is indeed the case, which is a special case of the *Fourier inversion theorem*. The latter is the analogue the Fourier summation problem for Fourier series. The Schwartz space is a most convenient setting in which to formulate the inversion.

Proposition 6.4. *The Fourier transform takes the Schwartz space onto itself. In fact, for any $f \in \mathcal{S}(\mathbb{R}^d)$ one has*

$$f(x) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, \hat{f}(\xi)\, d\xi \qquad \forall\ x \in \mathbb{R}^d \tag{6.3}$$

Proof. We would like to proceed by inserting the expression

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

into (6.3) and show that this recovers $f(x)$. The difficulty here is of course that interchanging the order of integration leads to the formal expression

$$\int_{\mathbb{R}^d} e^{2\pi i(x-y)\cdot\xi}\, d\xi$$

which needs to equal the $\delta_0(x-y)$ measure for the inversion formula (6.3) to hold. Although this is "essentially" the case, the problem here is of course that the previous integral is not well-defined. We shall therefore need to be more careful.

The standard procedure is to introduce a Gaussian weight since the Fourier transform has an explicit form on such weights. In fact, in Exercise 6.1 below we ask the reader to check the following identity

$$\widehat{e^{-\pi|x|^2}}(\xi) = e^{-\pi|\xi|^2} \tag{6.4}$$

Taking this for granted, we conclude from this and (6.2) that

$$\widehat{e^{-\pi\varepsilon^2|x|^2}}(\xi) = \varepsilon^{-d} e^{-\frac{\pi|\xi|^2}{\varepsilon^2}}$$

for any $\varepsilon > 0$.

Now fix $f \in \mathcal{S}(\mathbb{R}^d)$. We commence the proof of (6.3) with the observation that

$$\int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, \hat{f}(\xi)\, d\xi = \lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, e^{-\pi\varepsilon^2|\xi|^2}\, \hat{f}(\xi)\, d\xi$$

for each $x \in \mathbb{R}^d$. This follows from the dominated convergence theorem since $\hat{f} \in L^1$. Next, one has

$$\begin{aligned}
&\int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, e^{-\pi\varepsilon^2|\xi|^2}\, \hat{f}(\xi)\, d\xi \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{2\pi i(x-y)\cdot\xi}\, e^{-\pi\varepsilon^2|\xi|^2}\, d\xi\, f(y)\, dy \\
&= \int_{\mathbb{R}^d} \varepsilon^{-d}\, e^{-\pi\varepsilon^{-2}|x-y|^2}\, f(y)\, dy
\end{aligned}$$

The final observation is that

$$\varepsilon^{-d}\, e^{-\pi\varepsilon^{-2}|x|^2}$$

is an approximate identity in the sense of Chapter 1 with respect to the limit $\varepsilon \to 0$. Hence, since $f \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ one sees that

$$\lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} \varepsilon^{-d}\, e^{-\pi\varepsilon^{-2}|x-y|^2}\, f(y)\, dy = f(x)$$

for all $x \in \mathbb{R}^d$, which concludes the proof. □

Exercise 6.1. Prove the identity (6.4). *Hint:* Use contour integration in the complex plane.

Let us now return to the dilation identity (6.2) with f a smooth bump function. We note that this identity expresses an important normalization principle (at least heuristically), namely that the Fourier transforms maps an L^1-normalized bump function onto an L^∞-normalized one, and vice versa.

The dual of $\mathcal{S}$, denoted by $\mathcal{S}'$ is the space of *tempered distributions*. In other words, any $u \in \mathcal{S}'$ is a continuous functional on $\mathcal{S}$, and we write $\langle u, \phi\rangle$ for u applied to $\phi \in \mathcal{S}$. The space $\mathcal{S}'$ is equipped with the weak-$*$ topology. Thus, $u_n \to u$ in $\mathcal{S}'$ if and only if $\langle u_n, \phi\rangle \to \langle u, \phi\rangle$ as $n \to \infty$ for every $\phi \in \mathcal{S}$. Naturally, $L^p(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$ by the rule

$$\langle f, \phi\rangle = \int_{\mathbb{R}^d} f(x)\phi(x)\, dx, \qquad f \in L^p(\mathbb{R}^d),\ \phi \in \mathcal{S}(\mathbb{R}^d)$$

By a little functional analysis we may see that a linear functional u on $\mathcal{S}$ is continuous if and only if there exists N such that

$$|\langle u, \phi\rangle| \leqslant C \sum_{|\alpha|,|\beta|\leqslant N} \|x^\alpha \partial^\beta \phi\|_\infty \qquad \forall\ \phi \in \mathcal{S}$$

It follows that tempered distributions can be differentiated any number of times, and multiplied by any element of $C^\infty(\mathbb{R}^d)$ of at most polynomial growth. One has

$$\langle \partial^\alpha u, \phi\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha \phi\rangle \qquad \forall\ \phi \in \mathcal{S}(\mathbb{R}^d)$$

Any measure of finite total variation is an element of $\mathcal{S}'$, in particular the Dirac measure δ_0 is.

The Fourier transform of $u \in \mathcal{S}'$ is defined by the relation $\langle \hat{u}, \phi \rangle := \langle u, \hat{\phi} \rangle$ for all $\phi \in \mathcal{S}(\mathbb{R}^d)$. Since the Fourier transform is an isomorphism on $\mathcal{S}(\mathbb{R}^d)$, this definition is meaningful and establishes the distributional Fourier transform as an isomorphism on $\mathcal{S}'$. For example, $\hat{1} = \delta_0$ since

$$\langle 1, \hat{\phi} \rangle = \int_{\mathbb{R}^d} \hat{\phi}(\xi)\, d\xi = \phi(0)$$

In particular, we can meaningfully speak of the Fourier transform of any function in $L^p(\mathbb{R}^d)$ with $1 \leqslant p \leqslant \infty$.

The following exercise presents a very useful fact, namely that an element of $\mathcal{S}'(\mathbb{R}^d)$ which is supported at a point, say $\{0\}$ is necessarily a linear combination of δ_0 and finitely many derivatives thereof. "Support" here refers to the property that $\langle u, \varphi \rangle = 0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ that vanish near 0.

Exercise 6.2. Let $u, v \in \mathcal{S}'$ with $\langle u, \varphi \rangle = \langle v, \varphi \rangle$ for all $\varphi \in \mathcal{S}$ with $\mathrm{supp}(\hat{\varphi}) \subset \mathbb{R}^d \setminus \{0\}$. Then

$$u - v = P$$

for some polynomial P.

The following exercise introduces an important example of a distributional Fourier transform. The functions $|x|^{-\alpha}$ go by the name of *Riesz potentials*.

Exercise 6.3. For any $0 < \alpha < d$ compute the distributional Fourier transform of $|x|^{-\alpha}$ in $\mathbb{R}^d$. *Hint:* Show that the Fourier transform equals a smooth function of $\xi \neq 0$. Use dilation and rotational symmetry to determine this smooth function. Nonzero multiplicative constants do not need to be determined. Finally, exclude any contributions coming from $\xi = 0$, see the previous exercise.

An interesting example of a distributional identity is given by the Poisson summation formula.

Proposition 6.5. *For any $f \in \mathcal{S}(\mathbb{R}^d)$ one has*

$$\sum_{n \in \mathbb{Z}^d} f(n) = \sum_{\nu \in \mathbb{Z}^d} \hat{f}(\nu)$$

Equivalently, with the convergence being in the sense of $\mathcal{S}'$,

$$\sum_{\nu \in \mathbb{Z}^d} e^{2\pi i x \cdot \nu} = \sum_{n \in \mathbb{Z}^d} \delta_n(x)$$

where δ_n denotes the Dirac distribution at n.

PROOF. Given $f \in \mathcal{S}(\mathbb{R}^d)$, define $F(x) := \sum_{n\in\mathbb{Z}^d} f(x-n)$ which lies in $C^\infty(\mathbb{T}^d)$. The Fourier coefficients of F are with $e(\theta) = e^{2\pi i\theta}$ as usual,

$$\begin{aligned}\hat{F}(\nu) &= \sum_{n\in\mathbb{Z}^d} \int_{[0,1]^d} f(x-n)e(-x\cdot\nu)\,dx \\ &= \sum_{n\in\mathbb{Z}^d} \int_{[0,1]^d-n} f(x)e(-x\cdot\nu)\,dx = \int_{\mathbb{R}^d} f(x)e(-x\cdot\nu)\,dx \\ &= \hat{f}(\nu)\end{aligned}$$

From Proposition 1.14 we therefore conclude that

$$F(x) = \sum_{\nu\in\mathbb{Z}^d} \hat{f}(\nu)e(x\cdot\nu)$$

which implies the first identity. The second simply follows by chasing definitions. $\square$

Henceforth, we shall use the notation $\check{f}$ for the integral in (6.3). It is now easy to establish the Plancherel theorem.

COROLLARY 6.6. *For any $f \in \mathcal{S}(\mathbb{R}^d)$ one has $\|\hat{f}\|_2 = \|f\|_2$. In particular, the Fourier transform extends unitarily to $L^2(\mathbb{R}^d)$.*

PROOF. We start from the following identity which is of interest in its own right: for any $f, g \in \mathcal{S}(\mathbb{R}^d)$ one has

$$\int_{\mathbb{R}^d} f(\xi)\hat{g}(\xi)\,d\xi = \int_{\mathbb{R}^d} \hat{f}(x)g(x)\,dx$$

The proof is an immediate consequence of the definition of the Fourier transform and Fubini's theorem. Therefore, one also has

$$\int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,d\xi = \int_{\mathbb{R}^d} f(x)\hat{\bar{\hat{g}}}(x)\,dx$$

However,

$$\hat{\bar{\hat{g}}} = \overline{\check{\hat{g}}} = \bar{g}$$

by the previous proposition, whence we obtain Parseval's indentity

$$\int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,d\xi = \int_{\mathbb{R}^d} f(x)\bar{g}(x)\,dx$$

which is equivalent to the Plancherel theorem and the unitarity of the Fourier transform. $\square$

EXERCISE 6.4. Show that the Schwartz space is an algebra under both multiplication and convolution. In particular, prove that $\widehat{fg} = \hat{f} * \hat{g}$ for any $f, g \in \mathcal{S}(\mathbb{R}^d)$.

We now take another look at the proof of Proposition 6.4. Denote

$$\Gamma_\varepsilon(x) := \varepsilon^{-d}\, e^{-\pi\varepsilon^{-2}|x|^2}$$

Then the proof makes use of the identity

$$\int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, e^{-\pi\varepsilon^2|\xi|^2}\, \hat{f}(\xi)\, d\xi = (f * \Gamma_\varepsilon)(x) \tag{6.5}$$

valid pointwise for all Schwartz functions f and any $\varepsilon > 0$. Let $I_\varepsilon(f)$ be the operator defined by the previous line. Since Γ_ε is an approximate identity as $\varepsilon \to 0$, we conclude the following useful observation which parallels results we encountered for Fourier series on the circle.

Corollary 6.7. *For any $1 \leqslant p < \infty$ and $f \in L^p(\mathbb{R}^d)$ one has $I_\varepsilon(f) \to f$ in $L^p(\mathbb{R}^d)$. If $f \in C(\mathbb{R}^d)$ and vanishes at infinity, then $I_\varepsilon(f) \to f$ uniformly.*

In particular, one has *uniqueness*.

Corollary 6.8. *Suppose $f \in L^1(\mathbb{R}^d)$ satisfies $\hat{f} = 0$ everywhere. Then $f = 0$.*

Exercise 6.5. Let μ be a compactly supported measure in $\mathbb{R}^d$. Show that $\hat{\mu}$ extends to an entire function in $\mathbb{C}^d$. Conclude that $\hat{\mu}$ cannot vanish on an open set in $\mathbb{R}^d$.

Finally, we note the following basic estimate which is called *Hausdorff-Young* inequality.

Lemma 6.9. *Fix $1 \leqslant p \leqslant 2$. Then one has the property that $\hat{f} \in L^{p'}(\mathbb{R}^d)$ for any $f \in L^p(\mathbb{R}^d)$. Moreover, there is the bound $\|\hat{f}\|_{p'} \leqslant \|f\|_p$.*

Proof. Note that $f \in L^p(\mathbb{R}^d)$ can be written as $f = f_1 + f_2$ where

$$f_1 = f\chi_{[|f|<1]} \in L^2(\mathbb{R}^d), \qquad f_2 = f\chi_{[|f|\geqslant 1]} \in L^1(\mathbb{R}^d)$$

Hence $\hat{f}_1 \in L^2$ is well-defined by the L^2-extension of the Fourier transform, and $\hat{f}_2 \in L^\infty$ by integration. The estimate follows by interpolating between $p = 1$ and $p = 2$. □

As in the case of the circle, one can easily express all degrees of smoothness via the Fourier transform. For example, one can do this on the level of the Sobolev spaces H^s and $\dot{H}^s$ which are defined in terms of the norms

$$\|f\|_{H^s} = \|\langle\xi\rangle^s \hat{f}\|_2, \qquad \|f\|_{\dot{H}^s} = \||\xi|^s \hat{f}\|_2 \tag{6.6}$$

for any $s \in \mathbb{R}$ where $\langle\xi\rangle = \sqrt{1 + |\xi|^2}$. These spaces are the completion of $\mathcal{S}(\mathbb{R}^d)$ under the norms in (6.6). For the homogeneous space $\dot{H}^s(\mathbb{R}^d)$ one needs the restriction $s > -\frac{d}{2}$ since otherwise $\mathcal{S}$ is not dense in it.

Perhaps the most basic question relating to the Sobolev spaces concerns their *embedding properties*. An example is given by the following result.

Lemma 6.10. *For any $f \in H^s(\mathbb{R}^d)$ one has*

$$\|f\|_p \leqslant C(s)\|f\|_{H^s(\mathbb{R}^d)} \qquad \forall\ 2 \leqslant p \leqslant \infty$$

provided $s > \frac{d}{2}$. In fact $H^s(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ for any such s and all $2 \leqslant p \leqslant \infty$.

Proof. By Cauchy-Schwarz

$$\|f\|_\infty \leqslant \|\hat{f}\|_1 \leqslant \|\langle\xi\rangle^{-s}\|_2 \,\|\langle\xi\rangle^s \hat{f}(\xi)\|_2$$

for any $s > \frac{d}{2}$. The continuity of f follows by approximation by Schwartz functions. Interpolation with L^2 implies the lemma. □

By duality, one obtains for any $s < -\frac{d}{2}$

$$\|f\|_{H^s(\mathbb{R}^d)} \leqslant C(s,d)\|f\|_{L^1(\mathbb{R}^d)} \tag{6.7}$$

Exercise 6.6. Show that Lemma 6.10 fails at $s = \frac{d}{2}$. *Hint:* Argue by contradiction and use duality. Thus, show that (6.7) cannot hold at $s = -\frac{d}{2}$.

A simple generalization of Lemma 6.10 is the following result, which is an example of what is called a *trace lemma*.

Lemma 6.11. *Let $V \subset \mathbb{R}^d$ be an affine subspace of dimension $d - k$ where $1 \leqslant k \leqslant d$. Then one has*

$$\|f\|_{L^2(V)} \leqslant C(s)\,\|f\|_{H^s(\mathbb{R}^d)}$$

where $s > \frac{k}{2}$.

Proof. This follows by choosing coordinates parallel to $\mathcal{M}$ as well as perpendicular to it. For the parallel ones, one uses Plancherel on $\mathcal{M}$, whereas for the others Cauchy-Schwarz leads to the desired conclusion as in the previous proof. □

An analogous statement also holds when V is not flat but a smooth manifold (or rather, a compact piece thereof).

It is also natural to ask about embeddings $\dot{H}^s(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$ when $s \geqslant 0$. Of course, when $s = 0$ this is tautologically true with $p = 2$, but when $s > 0$ we expect an "improvement" in terms of a better regularity statement expressed by some $p > 2$. Note that the corresponding estimate

$$\|f\|_{L^p(\mathbb{R}^d)} \leqslant C\,\|f\|_{\dot{H}^s(\mathbb{R}^d)} \tag{6.8}$$

is necessarily scaling invariant. Thus, if we replace f by $f(\lambda\cdot)$, then the right-hand side scales with $\lambda^{s-\frac{d}{2}}$, whereas the left-hand side scales like $\lambda^{-\frac{d}{p}}$. Hence we arrive at the relation

$$\frac{1}{2} - \frac{1}{p} = \frac{s}{d} \tag{6.9}$$

This excludes the case $\frac{d}{2}$ by Exercise 6.6, and leaves open the possibility that for any $s \in (0, \frac{d}{2})$ and p as in (6.9) the estimate (6.8) does hold.

A good place to begin in this case is to consider $f \in L^2(\mathbb{R}^d)$ with $\mathrm{supp}(\hat{f}) \subset \{R \leqslant |\xi| \leqslant 2R\}$ and some $R > 0$. In that case, the following Bernstein inequality holds. The basic idea is the same as in Theorem 1.12, viz. functions with bounded Fourier support are arbitrarily smooth and this smoothness can be quantified in a number of ways, for example on the L^p-scale. We remark that (6.10) is scaling invariant.

Lemma 6.12. *Let* $f \in \mathcal{S}(\mathbb{R}^d)$ *satisfy* $\mathrm{supp}(\hat{f}) \subset \{|\xi| \leqslant R\}$. *Then*

$$\|f\|_q \leqslant C\, R^{d(\frac{1}{p}-\frac{1}{q})}\, \|f\|_p \tag{6.10}$$

$1 \leqslant p \leqslant q \leqslant \infty$.

Proof. Since (6.10) is scaling invariant, it suffices to consider $R = 1$. Then one has $\hat{f} = \chi \hat{f}$ where χ is smooth, compactly supported and $\chi = 1$ on the unit ball. It follows that $f = \check{\chi} * f$ and by Young's inequality therefore

$$\|f\|_q \leqslant \|\check{\chi}\|_r \|f\|_p$$

with $1 + \frac{1}{q} = \frac{1}{r} + \frac{1}{p}$, which is the same as $\frac{1}{p'} + \frac{1}{q} = \frac{1}{r}$. Note that $1 \leqslant p \leqslant q \leqslant \infty$ guarantees that $1 \leqslant r \leqslant p'$. $\square$

Exercise 6.7. Formulate and prove the general Young's inequality that was used in the previous proof.

In particular, if $\hat{f}$ is supported on $\{R \leqslant |\xi| \leqslant 2R\}$, then

$$\|f\|_p \leqslant CR^{d(\frac{1}{2}-\frac{1}{p})}\|f\|_2 = C\|f\|_{\dot{H}^s(\mathbb{R}^d)}$$

with s as in (6.9). This proves (6.8) for such functions, and the challenge is now to obtain the general case which is indeed true. The approach is to write $f = \sum_{j\in\mathbb{Z}} P_j f$ where $P_j f$ has Fourier support roughly on dyadic shells of size 2^j. While

$$\sum_{j\in\mathbb{Z}} \|P_j f\|_{\dot{H}^s}^2 \simeq \|f\|_{\dot{H}^s}^2$$

by definition, the problem is to show that

$$\|f\|_p^2 \leqslant C \sum_j \|P_j f\|_p^2 \tag{6.11}$$

This is indeed true for finite $p \geqslant 2$, but requires some machinery. We shall develop the necessary tools in the Littlewood-Paley chapter below. There are alternative routes leading to (6.8), such as fractional integration.

We now take up a topic that has proven to be very important for various applications, especially in partial differential equations: the decay properties of measures with smooth compactly supported densities that live on smooth submanifolds of $\mathbb{R}^d$. Let us start with a concrete example, namely the Cauchy problem for the Schrödinger equation

$$i\partial_t \psi + \Delta\psi = 0, \qquad \psi(0) = \psi_0 \tag{6.12}$$

where ψ_0 is a fixed function, say a Schwartz function, and $\psi(0)$ refers to the function $\psi(t, x)$ evaluated at $t = 0$. Applying the Fourier transform converts (6.12) into

$$i\partial_t \hat{\psi}(\xi) - 4\pi^2|\xi|^2\hat{\psi}(\xi) = 0, \qquad \hat{\psi}(0) = \widehat{\psi_0}$$

where $\hat{\psi}$ is the Fourier transform with respect to the space variable alone. The solution to the previous ordinary differential equation with respect to time is given by

$$\hat{\psi}(t,\xi) = e^{-4it\pi^2|\xi|^2}\widehat{\psi_0}(\xi)$$

whence the actual solution is after Fourier inversion

$$\psi(t,x) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, e^{-4it\pi^2|\xi|^2}\, \widehat{\psi_0}(\xi)\, d\xi \tag{6.13}$$

This formula is of course very relevant as it gives a solution to (6.12). For our purposes, we would like to re-interpret the integral on the right-hand side in the following form. Consider the paraboloid

$$\mathcal{P} := \left\{(\xi, -4\pi^2|\xi|^2) \mid \xi\in\mathbb{R}^d\right\} \subset \mathbb{R}^{d+1}$$

and place the measure μ onto $\mathcal{P}$ by the formula (for given ψ_0 as above)

$$\mu(d\xi, d\tau) = \widehat{\psi_0}(\xi)\, d\xi$$

via the relation

$$\int_{\mathbb{R}^{d+1}} F(\xi,\tau)\mu(d\xi,d\tau) = \int_{\mathbb{R}^d} F(\xi, -4\pi^2|\xi|^2)\, \widehat{\psi_0}(\xi)\, d\xi$$

valid for any continuous and bounded F. Then we see that (6.13) can be interpreted as

$$\psi(t,x) = \int_{\mathbb{R}^{d+1}} e^{2\pi i(x\cdot\xi + t\tau)}\, \mu(d\xi, d\tau)$$

In other words, the solution is given by the inverse Fourier transform of the measure μ which lives on the paraboloid $\mathcal{P}$. For a variety of reasons it has turned out to be of fundamental importance to understand the decay properties of such Fourier transforms. Before investigating this question, we ask the reader to verify a few properties of (6.13).

Exercise 6.8. Let $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$. Prove that (6.13) defines a function $\psi(t,x)$ which is smooth in t and belongs to the Schwartz class for all fixed times. Moreover, show that ψ solves (6.12) and satisfies

$$\|\psi(t)\|_{L^2} = \|\psi_0\|_2 \qquad \forall\ t\in\mathbb{R}$$

Finally, verify that for every $t > 0$

$$\psi(t,x) = c_d\, t^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{i\frac{\pi^2|x-y|^2}{4t}}\, \psi_0(y)\, dy \tag{6.14}$$

with a suitable constant c_d. Conclude that

$$\|\psi(t)\|_\infty \leqslant c_d\, t^{-\frac{d}{2}}\, \|\psi_0\|_1$$

for all $t > 0$.

Now let $\mathcal{M} \subset \mathbb{R}^d$ be a smooth manifold of dimension $k < d$ with Riemannian measure μ. Fix some smooth, compactly supported function ϕ on $\mathcal{M}$. The problem is to describe the asymptotic behavior of the Fourier transform $\widehat{\phi\mu}(\xi)$ for large ξ. For example, consider $\mathcal{M}$ to be a plane of dimension $d-1$. By translation and rotational symmetry we may take $\mathcal{M} = \mathbb{R}^{d-1}$, the subspace given by the first $d-1$ coordinates and ϕ is a Schwartz function of $x' := (x_1, \ldots, x_{d-1})$, whereas μ is simply Lebesgue measure on $\mathbb{R}^{d-1}$. Therefore, with $\xi = (\xi', \xi_d)$,

$$\widehat{\phi\mu}(\xi) = \int_{\mathbb{R}^{d-1}} e^{-2\pi i x'\cdot\xi'} \phi(x')\, dx'$$

which evidently is a Schwartz function in ξ' but does not depend at all on ξ_d whence also does not decay in that direction. What we have shown is that the Fourier transform of a measure that lives on a plane does not decay in the direction normal to the plane. In contrast to this case, we shall demonstrate that *nonvanishing Gaussian curvature* of a hypersurface $\mathcal{M}$ *always guarantees decay* at a universal rate.

To facilitate computations, fix some $\mathcal{M}$ and a measure on it with smooth compactly supported density. Working in local coordinates and fixing a direction in which we wish to describe the decay reduces matters to *oscillatory integrals* of the form

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi)\, d\xi\,, \tag{6.15}$$

with a compactly supported function $a \in C_0^\infty$ and smooth, real-valued phase $\phi \in C^\infty$. The asymptotic behavior of the integral $I(\lambda)$ as $\lambda \to \infty$ is described by the method of *(non)stationary phase*. This refers to the distinction between ϕ having a critical point on $\mathrm{supp}(a)$ or not. We begin with the easier case of non-stationary phase. The reader should note that the standard Fourier transform is of this type.

LEMMA 6.13. *If* $\nabla\phi \neq 0$ *on* $\mathrm{supp}(a)$ *then the integral* (6.15) *decays as*

$$\Big| \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi)\, d\xi \Big| \leqslant C(N, a, \phi)\, \lambda^{-N}, \quad \lambda \to \infty$$

for arbitrary $N \geqslant 1$.

PROOF. Note that the exponential $e^{i\lambda\phi}$ is an eigenfunction for the operator

$$L := \frac{1}{i\lambda} \frac{\nabla\phi}{|\nabla\phi|^2} \nabla \tag{6.16}$$

Therefore we can apply L to the exponential inside of (6.15) as often as we wish. The adjoint of L is

$$L^* = \frac{i}{\lambda} \nabla\Big(\frac{\nabla\phi}{|\nabla\phi|^2}\cdot\Big)$$

Then for any positive integer N, one has

$$\begin{aligned}
\Big| \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi)\, d\xi \Big| &= \Big| \int_{\mathbb{R}^d} L^N(e^{i\lambda\phi(\xi)}) a(\xi)\, d\xi \Big| \\
&= \Big| \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} (L^*)^N (a(\xi))\, d\xi \Big| \\
&\leqslant \int_{\mathbb{R}^d} |(L^*)^N (a(\xi))|\, d\xi \\
&\leqslant C(N, a, \phi) \lambda^{-N}
\end{aligned}$$

as claimed. $\square$

The situation changes when ϕ has a critical point inside $\mathrm{supp}(a)$. In that case one no longer has arbitrary decay and it can be very difficult to determine the exact rate. However, if the critical point is *nondegenerate*, i.e., ϕ has a nondegenerate Hessian at that point, then there is a precise answer as we now show.

Lemma 6.14. *If $\nabla\phi(\xi_0) = 0$ for some $\xi_0 \in \mathrm{supp}(a)$, $\nabla\phi \neq 0$ away from ξ_0 and the Hessian of ϕ at the stationary point ξ_0 is non-degenerate, i.e.,* $\det D^2\phi(\xi_0) \neq 0$, *then for all $\lambda \geqslant 1$,*

$$\Big| \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi)\, d\xi \Big| \leqslant C(d, a, \phi)\, \lambda^{-\frac{d}{2}} \tag{6.17}$$

In fact,

$$\Big| \partial_\lambda^k \Big[e^{-i\lambda\phi(0)} \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi)\, d\xi \Big] \Big| \leqslant C(d, a, \phi, k)\, \lambda^{-\frac{d}{2}-k} \tag{6.18}$$

for any integer $k \geqslant 1$.

Proof. Without loss of generality we assume $\xi_0 = 0$. By Taylor expansion

$$\phi(\xi) = \phi(0) + \frac{1}{2}\langle D^2\phi(0)\xi, \xi\rangle + O(|\xi|^3)$$

Since $D^2\phi(0)$ is non-degenerate we have $|\nabla\phi(\xi)| \gtrsim |\xi|$ on $\mathrm{supp}(a)$. We split the integral into two parts,

$$\int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi)\, d\xi = I + II$$

where the first part localizes the contribution near $\xi = 0$,

$$I := \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi) \chi_0(\lambda^{\frac{1}{2}}\xi)\, d\xi$$

and the second part restores the original integral, viz.

$$II = \int_{\mathbb{R}^d} e^{i\lambda\phi(\xi)} a(\xi) (1 - \chi_0(\lambda^{\frac{1}{2}}\xi))\, d\xi$$

The first term has the desired decay

$$|I| \leqslant C \int_{\mathbb{R}^d} |\chi_0(\lambda^{\frac{1}{2}}\xi)|\, d\xi \leqslant C\, \lambda^{-\frac{d}{2}}$$

To bound the second term, we proceed as in the proof of the nonstationary case. Thus, let L be as in (6.16) and note that

$$\Big| D^\alpha \Big(\frac{\nabla\phi}{|\nabla\phi|^2} \Big)(\xi) \Big| \leqslant C_\alpha |\xi|^{-1-|\alpha|}$$

for any multi-index α. Thus, let $N > d$ be an integer. Then the Leibnitz rule yields

$$
\begin{aligned}
|II| &\leqslant C \int\limits_{[|\xi| > \lambda^{-\frac{1}{2}}]} |(L^*)^N (1 - \chi_0(\lambda^{\frac{1}{2}}\xi)) a(\xi)|\, d\xi \\
&\leqslant C\,\lambda^{-N} \int\limits_{\lambda^{-\frac{1}{2}}}^{\infty} (\lambda^{\frac{N}{2}} r^{-N} + r^{-2N})\, r^{d-1}\, dr \leqslant C\,\lambda^{-\frac{d}{2}}
\end{aligned}
$$

as claimed. The stronger bound (6.18) follows from the observation that

$$|\phi(\xi) - \phi(0)| \leqslant C\,|\xi|^2$$

and the same argument as before (with $N > d + 2k$) concludes the proof. $\square$

It is possible to obtain asymptotic expansions of $I(\lambda)$ in inverse powers of large λ. In fact, the estimate provided by the preceding stationary phase lemma is optimal.

Let us now return to the Fourier transform of $d\mu = \phi\, d\sigma$ where σ is the surface measure of a hypersurface $\mathcal{M}$ in $\mathbb{R}^d$ and ϕ is compactly supported and smooth. Denote the unit normal vector to $\mathcal{M}$ at $x \in \mathcal{M}$ by $N(x)$.

Corollary 6.15. *Let $\xi = \lambda\nu$ where $\nu \in S^{d-1}$ and $\lambda \geqslant 1$. If ν is not parallel to $N(x)$ for any $x \in$ supp(ϕ), then $\hat{\mu} = O(\lambda^{-k})$ as $\lambda \to \infty$ for any positive integer k. Assume that $\mathcal{M}$ has nonvanishing Gaussian curvature on* supp(ϕ). *Then*

$$|\hat{\mu}(\xi)| \leqslant C(\mu, \mathcal{M})\, |\xi|^{-\frac{d-1}{2}}$$

for all $\lambda \geqslant 1$ and all ν. This is optimal if ν does belong to the normal bundle of $\mathcal{M}$ on supp(ϕ).

Proof. One has

$$\hat{\mu}(\xi) = \int_{\mathcal{M}} e^{i\lambda\langle \nu, x\rangle}\, \phi(x)\, \sigma(dx)$$

Note that the function $x \mapsto \langle \nu, x\rangle$ as a function on $\mathcal{M}$ has a critical point at $x_0 \in \mathcal{M}$ if and only if ν is perpendicular to $T_{x_0}\mathcal{M}$, the tangent space at x_0 to $\mathcal{M}$. Lemma 6.13 implies the first statement. For the second statement we note that the critical point at x_0 is nondegenerate if and only if the Gaussian curvature at x_0 does not vanish, and thus the bound follows from stationary phase. $\square$

Lemma 6.14 in fact allows for a more precise description of $\hat{\mu}$. We illustrate this by means of $\widehat{\sigma_{S^{d-1}}}(x)$ for large x, where $\sigma_{S^{d-1}}$ is the surface measure of the unit sphere in $\mathbb{R}^d$. This representation is very useful in certain applications where the oscillatory nature of the Fourier transform of the surface measure is needed, and not just its decay.

Note that the function $\widehat{\sigma_{S^{d-1}}}(x)$ is smooth and bounded, so we are interested in its asymptotic behavior for large $|x|$, both in terms of oscillations and size. We need to understand the oscillations in order to bound derivatives of $\widehat{\sigma_{S^{d-1}}}(x)$. We remark that the following result can also be obtained from the asymptotics of Bessel functions, but we make no use of special functions in this book.

COROLLARY 6.16. *One has the representation*

$$\widehat{\sigma_{S^{d-1}}}(x) = e^{i|x|}\omega_+(|x|) + e^{-i|x|}\omega_-(|x|) \qquad |x| \geqslant 1 \tag{6.19}$$

where $\omega_\pm$ are smooth and satisfy

$$|\partial_r^k \omega_\pm(r)| \leqslant C_k\, r^{-\frac{d-1}{2}-k} \quad \forall\, r \geqslant 1$$

and all $k \geqslant 0$.

PROOF. By definition,

$$\widehat{\sigma_{S^{d-1}}}(x) = \int_{S^{d-1}} e^{ix\cdot\xi}\, \sigma_{S^{d-1}}(d\xi)$$

which is rotationally invariant. Therefore we choose $x = (0, \ldots, 0, |x|)$ and denote the integral on the right-hand side by $I(|x|)$. Thus,

$$I(|x|) = \int_{S^{d-1}} e^{i|x|\xi_d}\, \sigma_{S^{d-1}}(d\xi)$$

Note that ξ_d, viewed as a function on S^{d-1}, has exactly two critical points, namely the north and south poles

$$(0, \ldots, 0, \pm 1)$$

which are, moreover, nondegenerate. We therefore partition the sphere into three parts: a small neighborhood around each of these poles and the remainder. The Fourier integral $I(|x|)$ then splits into three integrals by means of a smooth partition of unity adapted to these three open sets, and we write accordingly

$$I(|x|) = \sum_{\pm} I_\pm(|x|) + I_{\mathrm{eq}}(|x|)$$

The latter integral has non-stationary phase and decays like $|x|^{-N}$ for any N, see Lemma 6.13. On the other hand, the integrals $I_\pm$ exactly fit into the frame work of Lemma 6.14 which yields the desired result. Note that one can absorb the equatorial piece $I_{\mathrm{eq}}(|x|)$ into either $e^{\pm i|x|}\omega_\pm(|x|)$ without changing the stated properties of $\omega_\pm$, and (6.19) follows. $\square$

Notes

Many important topics are omitted here, such as spherical harmomics. A classical as well as comprehensive discussion of the Euclidean Fourier transform is the book by Stein and Weiss [**49**]. Folland [**17**] and especially Rudin [**40**] contain more information on the topic of tempered distributions. Hörmander [**26**] and Wolff [**56**] show that the Fourier transform on L^p with $p > 2$ take values in the distributions outside of the Lebesgue spaces (see Theorem 7.6.6 in [**26**]). For more on the topic of Sobolev embeddings and trace estimates see for example the comprehensive treatments in Gilbarg and Trudinger [**22**], as well as Evans [**13**], which do not invoke the Fourier transform. Stein [**45**] develops

Sobolev spaces in the context of the Fourier transform, but also uses the "fundamental theorem of calculus" approach. A standard reference for asymptotic expansions in stationary phase is the classical reference by Hörmander [**26**]. For Exercise 6.2 see Theorem 6.25 in Rudin [**40**].

Problems

PROBLEM 6.1. Compute the Fourier transform of P.V.$\frac{1}{x}$. In other words, determine

$$\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} e^{-2\pi i x\cdot\xi}\,\frac{dx}{x}$$

for every $\xi \in \mathbb{R}$. Conclude that (up to an irrelevant normalization) the Hilbert transform is an isometry on $L^2(\mathbb{R})$. This is the analogue for the line of Lemma 4.3.

PROBLEM 6.2. Solve the boundary value problem in the upper half plane for the Laplace equation via the Fourier transform. Compare your findings with Problem 4.5. Generalize to higher dimensions.

PROBLEM 6.3. Carry out the analogue Problem 1.11 with $\mathbb{R}$ instead of $\mathbb{T}$ and $\mathbb{R}^d$ instead of $\mathbb{T}^d$ (the case of the Schrödinger equation was already addressed in Exercise 6.8 above). Compute the explicit form of the heat kernel, i.e., the kernel G_t so that $G_t * u_0$ with $t > 0$ is the solution to the Cauchy problem for the heat equation $u_t - \Delta u = 0$, $u(0) = u_0$. Discuss the limit $t \to 0+$.

PROBLEM 6.4. Prove the following generalization of Bernstein's inequality, cf. Lemma 6.12: if $\operatorname{supp}(\hat f) \subset E \subset \mathbb{R}^d$ where E is measurable and f is Schwartz, say, prove that

$$\|f\|_q \leqslant |E|^{\frac{1}{p}-\frac{1}{q}}\|f\|_p \qquad \forall\ 1 \leqslant p \leqslant q \leqslant \infty \tag{6.20}$$

where $|E|$ stands for the Lebesgue measure. *Hint:* First handle the case $q = \infty$, $p = 2$, by Plancherel and Cauchy-Schwarz, and then dualize and interpolate.

PROBLEM 6.5. Show that if $f \in L^2(\mathbb{R})$ is supported in $[-R, R]$, then $\hat f$ extends to $\mathbb{C}$ as an entire function of exponential growth $2\pi R$. The Paley-Wiener theorem states the converse: if F is an entire function of exponential growth $2\pi R$ and such that $F \in L^2(\mathbb{R})$, then $F = \hat f$ where $f \in L^2(\mathbb{R})$ is supported in $[-R, R]$. Prove this via a deformation of contour, and extend to higher dimensions.

PROBLEM 6.6. Consider the kernel $K(x, y) = \frac{1}{x+y}$ on $\mathbb{R}^2_+$ and define the operator

$$(Tf)(x) = \int_0^\infty \frac{f(y)}{x+y}\,dy \qquad \forall\, x > 0$$

Show that T is bounded on $L^p(\mathbb{R}_+)$ for every $1 < p < \infty$, but not at $p = 1$ or $p = \infty$.

PROBLEM 6.7. The Laplace transform is defined as

$$(\mathcal{L}f)(s) := \int_0^\infty e^{-sx} f(x)\,dx, \quad s > 0 \tag{6.21}$$

Show that $\mathcal{L}$ is bounded on $L^2(\mathbb{R}_+)$ but not on any $L^p(\mathbb{R}_+)$ with $p \neq 2$. Verify that $\|\mathcal{L}\|_{2\to 2} = \sqrt{\pi}$. Use this to show that $\|T\|_{2\to 2} = \pi$ where T is as in (6.21). *Hint:* For the L^2 boundedness, write $e^{-sx} = e^{-\frac{sx}{2}}x^{-\frac{1}{4}}e^{-\frac{sx}{2}}x^{\frac{1}{4}}$ and apply Cauchy-Schwarz. For T consider $\mathcal{L}^2$.

PROBLEM 6.8. With $\mathcal{L}$ the Laplace transform on $L^2(\mathbb{R}_+)$ as in the previous problem, find an inversion formula.

PROBLEM 6.9. Establish a representation as in Corollary 6.16 for the ball instead of the sphere. I.e., find the asymptotics as in that corollary for $\widehat{\chi_B}$ where $B \subset \mathbb{R}^d$ is the unit-ball with $d \geqslant 2$. Generalize to the case of ellipsoids, i.e.,

$$\Big\{x = (x_1, \ldots, x_d) \in \mathbb{R}^d \mid \sum_{j=1}^{d} \lambda_j^{-2} x_j^2 \leqslant 1\Big\}$$

where $\lambda_j > 0$ are constants. Contrast this to the Fourier transform of the indicator function of the cube.

PROBLEM 6.10. Let A be a positive $d \times d$-matrix. Compute the Fourier transform of the Gaussian $e^{-\langle Ax, x\rangle}$.

PROBLEM 6.11. Compute the inverse Fourier transform of $(4\pi^2|\xi|^2 + k^2)^{-1}$ in $\mathbb{R}^3$ where $k > 0$ is a constant. Denote this function by $G(x; k)$. Show that the operator

$$(R(k)f)(x) = \int_{\mathbb{R}^d} G(x - y; k) f(y)\, dy, \qquad f \in \mathcal{S}(\mathbb{R}^3)$$

is bounded on $L^2(\mathbb{R}^d)$ and that $R(k)f \in C^\infty(\mathbb{R}^3)$ as well as

$$(-\Delta + k^2)R(k)f = f \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^3)$$

Conclude that $R(k) = (-\Delta + k^2)^{-1}$ for $k > 0$, i.e., $R(k)$ is the resolvent of the Laplacian.

CHAPTER 7

Calderón-Zygmund Theory of Singular Integrals

In this chapter we take up the problem of developing a real-variable theory of the Hilbert transform. In fact, we will subsume the Hilbert transform in a class of operators defined in any dimension. We shall make no reference to harmonic functions as we did in Chapter 4, but rather rely only the properties of the kernel alone. For simplicity, we shall mostly restrict ourselves to the translation invariant setting.

We begin with the following class of kernels which define operators by convolution. As for the Hilbert transform, we include a cancellation condition, see part *iii*) in the following definition. For example, this condition guarantees that the principal value as in (7.1) exists (contrast this to the case $K(x) = |x|^{-d}$). In addition, this condition will be convenient in proving the L^2-boundedness of the associated operators.

This being said, it is often preferable to assume the L^2-boundedness of the Calderón-Zygmund operator instead of the cancellation condition. The logic is then to develop the L^2-boundedness theory separately from the L^p-theory, with the goal of finding necessary and sufficient conditions for L^2-boundedness while the L^p theory requires only condition *ii*) in the following definition. In many ways, *ii*), which is called *Hörmander condition*, is therefore the most important one.

Definition 7.1. Let $K : \mathbb{R}^d \backslash \{0\} \to \mathbb{C}$ satisfy, for some constant B,

i) $|K(x)| \leqslant B|x|^{-d}$ for all $x \in \mathbb{R}^d$
ii) $\int_{|x|>2|y|} |K(x) - K(x-y)|\, dx \leqslant B$ for all $y \neq 0$
iii) $\int_{r<|x|<s} K(x)\, dx = 0$ for all $0 < r < s < \infty$

Then K is called a *Calderón-Zygmund* kernel.

Our first goal is With such a kernel we associate a translation invariant operator by means of the principal value. Thus, the *singular integral operator* (or Calderón-Zygmund operator) with kernel K is defined as

$$Tf(x) := \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} K(x-y) f(y)\, dy \tag{7.1}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$.

Exercise 7.1.

a) Check that the limit exists in (7.1) for all $f \in \mathcal{S}(\mathbb{R}^d)$. How much regularity on f suffices for this property (you may take f to be compactly supported if you wish)?

b) Check that the Hilbert transform on the line with kernel $\frac{1}{x}$ is a singular integral operator.

There is the following simple condition that guarantees *ii*).

Lemma 7.2. *Suppose $|\nabla K(x)| \leqslant B|x|^{-d-1}$ for all $x \neq 0$ and some constant B. Then*

$$\int_{|x|>2|y|} |K(x) - K(x-y)|\, dx \leqslant CB \tag{7.2}$$

with $C = C(d)$.

Proof. Fix $x, y \in \mathbb{R}^d$ with $|x| > 2|y|$. Connect x and $x - y$ by the line segment $x - ty$, $0 \leqslant t \leqslant 1$. This line segment lies entirely inside the ball $B(x, |x|/2)$. Hence

$$\begin{aligned} |K(x) - K(x-y)| &= \Big| - \int_0^1 \nabla K(x-ty) y\, dt \Big| \\ &\leqslant \int_0^1 |\nabla K(x-ty)||y|\, dt \leqslant B 2^{d+1} |x|^{-d-1} |y| \end{aligned}$$

Inserting this bound into the left-hand side of (7.2) yields the desired bound. □

Exercise 7.2. Check that, for any fixed $0 < \alpha \leqslant 1$, and all $x \neq 0$,

$$\sup_{|y| < \frac{|x|}{2}} \frac{|K(x) - K(x-y)|}{|y|^\alpha} \leqslant B|x|^{-d-\alpha}$$

also implies (7.2).

We now prove L^2-boundedness of T, by computing $\hat{K}$ and applying Plancherel. This is an instance where the full force of Definition 7.1 comes into play. We shall use all three conditions, in particular the cancellation condition.

Proposition 7.3. *Let K be as in Definition 7.1. Then $\|T\|_{2\to 2} \leqslant CB$ with $C = C(d)$.*

Proof. Fix $0 < r < s < \infty$ and consider

$$(T_{r,s} f)(x) = \int_{\mathbb{R}^d} K(y) \chi_{[r<|y|<s]}(y) f(x-y)\, dy$$

Let

$$m_{r,s}(\xi) := \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi} \chi_{[r<|x|<s]} K(x)\, dx$$

be the Fourier transform of the restricted kernel. By Plancherel's theorem it suffices to prove that

$$\sup_{0<r<s} \|m_{r,s}\|_\infty \leqslant CB \tag{7.3}$$

Indeed, if (7.3) holds, then

$$\|T_{r,s}\|_{2\to 2} = \|m_{r,s}\|_\infty \leqslant CB$$

uniformly in r, s. Moreover, for any $f \in \mathcal{S}(\mathbb{R}^d)$ one has

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

pointwise in $x \in \mathbb{R}^d$. Fatou's lemma therefore implies that $\|Tf\|_2 \leqslant CB\|f\|_2$ for any $f \in \mathcal{S}(\mathbb{R}^d)$. To verify (7.3) we split the integration in the Fourier transform into the regions $|x| < |\xi|^{-1}$ and $|x| \geqslant |\xi|^{-1}$. In the former case

$$\begin{aligned}\Big|\int_{r<|x|<|\xi|^{-1}} e^{-2\pi i x\cdot\xi} K(x)\,dx\Big| &= \Big|\int_{r<|x|<|\xi|^{-1}} (e^{-2\pi i x\cdot\xi} - 1)K(x)\,dx\Big|\\ &\leqslant \int_{|x|<|\xi|^{-1}} 2\pi|x||\xi||K(x)|\,dx \leqslant\\ &\leqslant 2\pi|\xi| \int_{|x|<|\xi|^{-1}} B|x|^{-d+1}\,dx\\ &\leqslant CB|\xi||\xi|^{-1} \leqslant CB\end{aligned}$$

as desired. Note that we used the cancellation condition *iii*) in the first equality sign. To deal with the case $|x| > |\xi|^{-1}$ one uses the cancellation in $e^{-2\pi i x\cdot\xi}$ which in turn requires smoothness of K, i.e., condition *ii*) (but compare Lemma 7.2). Firstly, observe that

$$\int_{s>|x|>|\xi|^{-1}} K(x)e^{-2\pi i x\cdot\xi}\,dx = -\int_{s>|x|>|\xi|^{-1}} K(x)e^{-2\pi i\left(x+\frac{\xi}{2|\xi|^2}\right)\cdot\xi}\,dx \tag{7.4}$$

$$= -\int_{s>|x-\frac{\xi}{2|\xi|^2}|>|\xi|^{-1}} K\Big(x - \frac{\xi}{2|\xi|^2}\Big)\,e^{-2\pi i x\cdot\xi}\,dx \tag{7.5}$$

Denoting the expression on the left-hand side of (7.4) by F one thus has

$$2F = \int_{s>|x|>|\xi|^{-1}} \Big(K(x) - K(x - \frac{\xi}{2|\xi|^2})\Big)e^{-2\pi i x\cdot\xi}\,dx + O(1) \tag{7.6}$$

The $O(1)$ term here stands for a term bounded by CB. Its origin is of course the difference between the regions of integration in (7.4) and (7.5). We leave it to the reader to check that condition *i*) implies that this error term is really no larger than CB. Estimating the integral in (7.6) by means of *ii*) now yields

$$|2F| \leqslant \int_{|x|>|\xi|^{-1}} \Big|K(x) - K\Big(x - \frac{\xi}{2|\xi|^2}\Big)\Big|\,dx + CB \leqslant CB$$

as claimed. We have shown (7.3) and the proposition follows. □

EXERCISE 7.3. Supply the omitted details concerning the $O(1)$ term in the previous proof.

Our next goal is to show that singular integrals are bounded as operators from $L^1(\mathbb{R}^d)$ into weak-L^1. This requires the following basic decomposition lemma due to Calderón-Zygmund for L^1 functions. The proof is an example of a *stopping time argument*.

LEMMA 7.4. *Let $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$. Then one can write $f = g + b$ where $|g| \leqslant \lambda$ and $b = \sum_Q \chi_Q f$ where the sum runs over a collection $\mathcal{B} = \{Q\}$ of disjoint cubes such that for each Q one has*

$$\lambda < \frac{1}{|Q|}\int_Q |f| \leqslant 2^d \lambda \tag{7.7}$$

Furthermore,

$$\Big|\bigcup_{Q\in\mathcal{B}} Q\Big| < \frac{1}{\lambda}\|f\|_1 \tag{7.8}$$

PROOF. For each $\ell \in \mathbb{Z}$ we define a collection $\mathcal{D}_\ell$ of dyadic cubes by means of

$$\mathcal{D}_\ell = \big\{\Pi_{i=1}^d [2^\ell m_i, 2^\ell(m_i+1)) \mid m_1,\ldots,m_d \in \mathbb{Z}\big\}$$

Notice that if $Q \in \mathcal{D}_\ell$ and $Q' \in \mathcal{D}_{\ell'}$, then either $Q \cap Q' = \emptyset$ or $Q \subset Q'$ or $Q' \subset Q$. Now pick ℓ_0 so large that

$$\frac{1}{|Q|}\int_Q |f|\,dx \leqslant \lambda$$

for every $Q \in \mathcal{D}_{\ell_0}$. For each such cube consider its 2^d "children" of size 2^{ℓ_0-1}. Any such cube Q' will then have the property that either

$$\frac{1}{|Q'|}\int_{Q'} |f(x)|\,dx \leqslant \lambda \ \text{ or } \frac{1}{|Q'|}\int_{Q'} |f(x)|\,dx > \lambda \tag{7.9}$$

In the latter case we stop, and include Q' in the family $\mathcal{B}$ of "bad cubes". Observe that in this case

$$\frac{1}{|Q'|}\int_{Q'} |f(x)|\,dx \leqslant \frac{2^d}{|Q|}\int_Q |f(x)|\,dx \leqslant 2^d\lambda$$

where Q denotes the parent of Q'. Thus (7.7) holds. If, however, the first inequality in (7.9) holds, then subdivide Q' again into its children of half the size. Continuing in this fashion produces a collection of disjoint (dyadic) cubes $\mathcal{B}$ satisfying (7.7). Consequently, (7.8) also holds, since

$$\Big|\bigcup_{\mathcal{B}} Q\Big| \leqslant \sum_{\mathcal{B}} |Q| < \sum_{\mathcal{B}} \frac{1}{\lambda}\int_Q |f(x)|\,dx \leqslant \frac{1}{\lambda}\int_{\mathbb{R}^d} |f(x)|\,dx$$

Now let $x_0 \in \mathbb{R}^d \backslash \bigcup_{\mathcal{B}} \overline{Q}$. Then x_0 is contained in a decreasing sequence $\{Q_j\}$ of dyadic cubes each of which satisfies

$$\frac{1}{|Q_j|}\int_{Q_j} |f(x)|\,dx \leqslant \lambda$$

By Lebesgue's theorem $|f(x_0)| \leqslant \lambda$ for a.e. such x_0. Since moreover $\mathbb{R}^d \setminus \cup_{\mathcal{B}} Q$ and $\mathbb{R}^d \setminus \cup_{\mathcal{B}} \overline{Q}$ differ only by a set of measure zero, we can set

$$g = f - \sum_{Q \in \mathcal{B}} \chi_{\overline{Q}} f$$

so that $|g| \leqslant \lambda$ a.e. as desired. □

We can now state the crucial weak-L^1 bound for singular integrals. We shall do this under more general hypotheses than those of Definition 7.1.

Proposition 7.5. *Suppose T is a linear operator bounded on $L^2(\mathbb{R}^d)$ such that*

$$(Tf)(x) = \int_{\mathbb{R}^d} K(x-y) f(y)\, dy \qquad \forall\, f \in L^2_{\mathrm{comp}}(\mathbb{R}^d)$$

and all $x \notin \operatorname{supp}(f)$. Assume that K satisfies condition iii) of Definition 7.1, but not necessarily any of the other conditions. Then for every $f \in \mathcal{S}(\mathbb{R}^d)$ there is the weak-L^1 bound

$$\big|\{x \in \mathbb{R}^d \mid |Tf(x)| > \lambda\}\big| \leqslant \frac{CB}{\lambda} \|f\|_1 \qquad \forall\, \lambda > 0$$

where $C = C(d)$.

Proof. Dividing by B if necessary, we may assume that $B = 1$. Now fix $f \in \mathcal{S}(\mathbb{R}^d)$ and let $\lambda > 0$ be arbitrary. By Lemma 7.4 one can write $f = g + b$ with this value of λ. We now set

$$f_1 = g + \sum_{Q \in \mathcal{B}} \chi_Q \frac{1}{|Q|} \int_Q f(x)\, dx$$

$$f_2 = b - \sum_{Q \in \mathcal{B}} \chi_Q \frac{1}{|Q|} \int_Q f(x)\, dx = \sum_{Q \in \mathcal{B}} f_Q$$

where we have set

$$f_Q := \chi_Q \Big(f - \frac{1}{|Q|} \int_Q f(x)\, dx \Big)$$

Notice that $f = f_1 + f_2$, $\|f_1\|_\infty \leqslant C\lambda \|f\|_1$, $\|f_2\|_1 \leqslant 2\|f\|_1$, $\|f_1\|_1 \leqslant 2\|f\|_1$, and

$$\int_Q f_Q(x)\, dx = 0$$

for all $Q \in \mathcal{B}$. We now proceed as follows:

$$\begin{aligned}
&\big|\{x \in \mathbb{R}^d \mid |(Tf)(x)| > \lambda\}\big| \\
&\leqslant \Big|\Big\{x \in \mathbb{R}^d \mid |(Tf_1)(x)| > \frac{\lambda}{2}\Big\}\Big| + \Big|\Big\{x \in \mathbb{R}^d \mid |(Tf_2)(x)| > \frac{\lambda}{2}\Big\}\Big| \\
&\leqslant \frac{C}{\lambda^2} \|Tf_1\|_2^2 + \Big|\Big\{x \in \mathbb{R}^d \mid |(Tf_2)(x)| > \frac{\lambda}{2}\Big\}\Big| \qquad (7.10)
\end{aligned}$$

The first term in (7.10) is controlled by Proposition 7.3:

$$\frac{C}{\lambda^2} \|Tf_1\|_2^2 \leqslant \frac{C}{\lambda^2} \|f_1\|_2^2 \leqslant \frac{C}{\lambda^2} \|f_1\|_\infty \|f_1\|_1 \leqslant \frac{C}{\lambda} \|f\|_1$$

To estimate the second term in (7.10) we define, for any $Q \in \mathcal{B}$, the cube Q^* to be the dilate of Q by the fixed factor $2\sqrt{n}$ (i.e., Q^* has the same center as Q but side length equal to $2\sqrt{n}$ times that of Q). Thus

$$\begin{aligned}
&\Big|\big\{x \in \mathbb{R}^d \mid |(Tf_2)(x)| > \frac{\lambda}{2}\big\}\Big| \\
&\leqslant \Big| \cup_{\mathcal{B}} Q^*\Big| + \Big|\big\{x \in \mathbb{R}^d \setminus \cup_{\mathcal{B}} Q^* \mid |(Tf_2)(x)| > \frac{\lambda}{2}\big\}\Big| \\
&\leqslant C\sum_Q |Q| + \frac{2}{\lambda}\int_{\mathbb{R}^d\setminus\cup Q^*} |(Tf_2)(x)|\,dx \\
&\leqslant \frac{C}{\lambda}\|f\|_1 + \frac{2}{\lambda}\sum_{Q\in\mathcal{B}}\int_{\mathbb{R}^d\setminus Q^*} |(Tf_Q)(x)|\,dx
\end{aligned}$$

Perhaps the most crucial point of this proof is the fact that f_Q has mean zero which allows one to exploit the smoothness of the kernel K. More precisely, for any $x \in \mathbb{R}^d\setminus Q^*$,

$$\begin{aligned}
(Tf_Q)(x) &= \int_Q K(x-y) f_Q(y)\,dy \\
&= \int_Q [K(x-y) - K(x-y_Q)] f_Q(y)\,dy
\end{aligned} \tag{7.11}$$

where y_Q denotes the center of Q. Thus

$$\begin{aligned}
\int_{\mathbb{R}^d\setminus Q^*} |(Tf_Q)(x)|\,dx &\leqslant \int_{\mathbb{R}^d\setminus Q^*}\int_Q |K(x-y) - K(x-y_Q)||f_Q(y)|\,dy\,dx \\
&\leqslant \int_Q |f_Q(y)|\,dy \leqslant 2\int_Q |f(y)|\,dy
\end{aligned}$$

To pass to the second inequality sign we used condition *ii*) in Definition 7.1 and our choice of Q^*. Hence the second term on the right bound side of (7.11) is no larger than

$$\frac{C}{\lambda}\sum_Q\int_Q |f(y)|dy \ \leqslant \frac{C}{\lambda}\|f\|_1$$

and we are done. □

The assumption $f \in \mathcal{S}(\mathbb{R}^d)$ in the previous proposition was for convenience only. It ensured that one could define Tf by means of the principal value (7.1). However, observe that Proposition 7.3 allows one to extend T to a bounded operator on L^2. This in turn implies that the weak-L^1 bound in Proposition 7.5 holds for all $f \in L^1 \cap L^2(\mathbb{R}^d)$. Indeed, inspection of the proof reveals that apart from the L^2 boundedness of T, see (7.10), the definition of T in terms of K was only used in (7.11) where x and y are assumed to be sufficiently separated so that the integrals are absolutely convergent.

Theorem 7.6 (Calderón-Zygmund). *Let T be a singular integral operator as in Definition 7.1. Then for every $1 < p < \infty$ one can extend T to a bounded operator on $L^p(\mathbb{R}^d)$ with the bound $\|T\|_{p\to p} \leqslant CB$ with $C = C(p,d)$.*

PROOF. By Proposition 7.3 and 7.5 (and the previous remark) one obtains this statement for the range $1 < p \leqslant 2$ from the Marcinkiewicz interpolation theorem. The range $2 \leqslant p < \infty$ now follows by duality. Indeed, for $f, g \in \mathcal{S}(\mathbb{R}^d)$ one has

$$\langle Tf, g\rangle = \langle f, T^*g\rangle \text{ where } T^*g(x) = \lim_{\varepsilon\to 0}\int_{|x-y|>\varepsilon} K^*(x-y)g(y)\,dy$$

and $K^*(x) := \overline{K(-x)}$. Since K^* clearly verifies conditions *i*)–*iii*) in Definition 77.1, we are done. □

REMARK. It is important to realize that the cancellation condition *iii*) was only used to prove L^2 boundedness, but did not appear in the proof of Proposition 7.5 explicitly. Therefore, T is bounded on $L^p(\mathbb{R}^d)$ provided it is bounded for $p = 2$ and conditions *i*) and *ii*) of Definition 7.1 hold.

We now present some of the most basic examples of singular integrals. Consider the Poisson equation $\triangle u = f$ in $\mathbb{R}^d$ where $f \in \mathcal{S}(\mathbb{R}^d)$, $d \geqslant 2$. In the following exercise we ask the reader to show that in dimensions $d \geqslant 3$

$$u(x) = C_d \int_{\mathbb{R}^d} |x-y|^{2-d} f(y)\,dy \tag{7.12}$$

with some dimensional constant C_d is a solution to this equation. In fact, one can show that (7.12) gives the *unique bounded solution* but we shall not need to know that here. In two dimensions, the formula reads

$$u(x) = C_2 \int_{\mathbb{R}^2} \log(|x-y|) f(y)\,dy \tag{7.13}$$

The kernels $|x|^{2-d}$ and $\log|x|$ are called *Newton potentials*. It turns out that the second derivatives $\frac{\partial^2 u}{\partial x_i \partial x_j}$ with u defined by these convolutions can be expressed as Calderón-Zygmund operators acting on f (the so-called “double Riesz transforms”).

EXERCISE 7.4.

a) With $f \in \mathcal{S}$ and a suitable constant C_d, show that (7.12) satisfies

$$\Delta u = f$$

Hint: pass the Laplacian onto f and introduce a cut-off to the set $[|x-y| > \varepsilon]$ in (7.12). Apply Green's formula, and let $\varepsilon \to 0$ to recover $f(x)$.

b) With u as in (7.12), show that for $f \in \mathcal{S}(\mathbb{R}^d)$ with $d \geqslant 3$

$$\frac{\partial^2 u}{\partial x_i \partial x_j}(x) = C_d\,(d-2)(d-1)\int_{\mathbb{R}^d} K_{ij}(x-y) f(y)\,dy$$

in the principal value sense, where

$$K_{ij}(x) = \begin{cases} \frac{x_i x_j}{|x|^{d+2}} & \text{if} \quad i \neq j \\ \frac{x_i^2 - \frac{1}{d}|x|^2}{|x|^{d+2}} & \text{if} \quad i = j \end{cases}$$

Find the analogue for $d = 2$.

c) Verify that K_{ij} as above are singular integral kernels. Also show that $K_i(x) = \frac{x_i}{|x|^{d+1}}$ is a singular integral kernel.

The operators R_i and R_{ij} defined in terms of the kernels K_i and K_{ij}, respectively, are called the Riesz transforms or double Riesz transforms. By Exercise 7.4

$$R_{ij}(\triangle\varphi) = \frac{\partial^2\varphi}{\partial x_i \partial x_j} \tag{7.14}$$

for any $\varphi \in \mathcal{S}(\mathbb{R}^d)$. The following estimate is one of the most basic L^p estimate for elliptic equations.

COROLLARY 7.7. *Let $u \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$\sup_{1\leqslant i,j\leqslant d} \left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_{L^p(\mathbb{R}^d)} \leqslant C_{p,d} \|\triangle u\|_{L^p(\mathbb{R}^d)} \tag{7.15}$$

for any $1 < p < \infty$. Here $C_{p,d}$ only depends on p and d.

PROOF. This is an immediate consequence of the representation formula (7.14) and the Calderón-Zygmund Theorem 7.6. □

Corollary 7.7 is remarkable since it states that the Hessian is controlled by its trace. This cannot hold in the pointwise sense, which means that the corollary fails at $p = \infty$ and by duality therefore also at $p = 1$.

EXERCISE 7.5. This exercise explains the aforementioned failure at $p = 1$ or $p = \infty$ in more detail. The idea is simply to take u to be the Newton potential. Then $\triangle u = \delta_0$ which, at least heuristically, belongs to $L^1(\mathbb{R}^d)$. However, one checks that $\frac{\partial^2 u}{\partial x_i \partial x_j} \notin L^1(B(0,1))$ for any i, j, see the kernels K_{ij} form Exercise 7.4). In oder to transfer δ_0 to L^1, we convolve the Newton potentials with an approximate identity (this is called *mollifying*); in addition, we truncate to a bounded set.

a) Let $\varphi \geqslant 0$ be in $C_0^\infty(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \varphi\, dx = 1$. Set $\varphi_\varepsilon(x) := \varepsilon^{-d}\varphi(\frac{x}{\varepsilon})$ for any $0 < \varepsilon < 1$. Clearly, $\{\varphi_\varepsilon\}_{0<\varepsilon<1}$ form an approximate identity provided the latter are defined analogously to Definition 1.3 on $\mathbb{R}^d$. Moreover, let $\chi \in C_0^\infty(\mathbb{R}^d)$ be arbitrary with $\chi(0) \neq 0$. Verify that, with $\Gamma_d(x) = |x|^{2-d}$ for $n \geqslant 3$ and $\Gamma_2(x) = \log|x|$, $u_\varepsilon(x) := (\varphi_\varepsilon * \Gamma_d)(x)\chi(x)$ has the following properties:

$$\sup_{\varepsilon>0} \|\triangle u_\varepsilon\|_{L^1} < \infty$$

and

$$\limsup_{\varepsilon\to 0} \|\frac{\partial^2 u_\varepsilon}{\partial x_i \partial x_j}\|_{L^1} = \infty$$

for any $1 \leqslant i, j \leqslant d$. Thus Corollary 7.7 fails on $L^1(\mathbb{R}^d)$.

b) Now show that Corollary 7.7 also fails on L^∞.

The convolution in (7.12) is an example of what one calls *fractional integration*. More generally, for any $0 < \alpha < d$ define

$$(I_\alpha f)(x) := \int_{\mathbb{R}^d} |x-y|^{\alpha-d} f(y)\,dy, \qquad f \in \mathcal{S}(\mathbb{R}^d) \tag{7.16}$$

We cannot use Young's inequality since the kernel is not in any $L^r(\mathbb{R}^d)$ space, but it is clearly in weak-L^r with $r = \frac{d}{d-\alpha}$. The following proposition shows that nevertheless one still has (up to an endpoint) the same conclusion as Young's inequality would imply.

Proposition 7.8. *For any $1 < p < q < \infty$ and any $0 < \alpha < d$ there is the bound*

$$\|I_\alpha f\|_q \leqslant C(p,q,d)\|f\|_p \ \text{ provided } \ \frac{1}{q} = \frac{1}{p} - \frac{\alpha}{d}$$

for any $f \in \mathcal{S}(\mathbb{R}^d)$.

Proof. By Marcinkiewicz's interpolation theorem it suffices for this bound to hold with weak L^q. Thus, we need to show that

$$|\{x \in \mathbb{R}^d \mid |I_\alpha f(x)| > \lambda\}| \leqslant C\lambda^{-q}\|f\|_p^q \tag{7.17}$$

Fix some f with $\|f\|_p = 1$ and an arbitary $\lambda > 0$. With $r > 0$ to be determined, break up the kernel as follows

$$k_\alpha(x) := |x|^{\alpha-d} = k_\alpha(x)\chi_{[|x|<r]} + k_\alpha(x)\chi_{[|x|>r]} =: k_\alpha^{(1)}(x) + k_\alpha^{(2)}(x)$$

Then

$$\|k_\alpha^{(2)} * f\|_\infty \leqslant \|k_\alpha^{(2)}\|_{p'}\|f\|_p = Cr^{\alpha-\frac{d}{p}}$$

Determine r so that the right-hand side equals $\frac{\lambda}{2}$. Then

$$\begin{aligned}|\{x \in \mathbb{R}^d \mid |I_\alpha f(x)| > \lambda\}| &\leqslant |\{x \in \mathbb{R}^d \mid |(k_\alpha^{(1)} * f)(x)| > \lambda/2\}| \\ &\leqslant C\lambda^{-p}\|k_\alpha^{(1)} * f\|_p^p \leqslant C\lambda^{-p}\|k_\alpha^{(1)}\|_1^p \\ &\leqslant C\lambda^{-p}r^{\alpha p} = C\lambda^{-q}\end{aligned}$$

which is (7.17). □

The condition on p and q is of course dictated by scaling. By Exercise 6.3, one has

$$(-\Delta)^{-\frac{s}{2}} f = C(s,d)\, I_s f \qquad \forall\ f \in \mathcal{S}(\mathbb{R}^d) \tag{7.18}$$

Thus, Proposition 7.8 implies the following *Sobolev imbedding estimate*.

Corollary 7.9. *Let $0 \leqslant s < \frac{d}{2}$. Then for all $\frac{1}{2} - \frac{1}{q} = \frac{s}{d}$ one has*

$$\|f\|_{L^q(\mathbb{R}^d)} \leqslant C(s,d)\|f\|_{\dot{H}^s(\mathbb{R}^d)} \qquad \forall\ f \in \mathcal{S}(\mathbb{R}^d) \tag{7.19}$$

Proof. Fix $f \in \mathcal{S}(\mathbb{R}^d)$ and set $g := (-\Delta)^{\frac{s}{2}} f$. Then by (7.18) one has

$$f = C(s,d) I_s g$$

and the estimate (7.19) follows from the previous proposition. □

We now address the question whether a singular integral operator can be defined by means of formula (7.1) even if $f \in L^p(\mathbb{R}^d)$ rather than $f \in \mathcal{S}(\mathbb{R}^d)$, say. This question is the analogue of Proposition 4.9 and should be understood as follows: for $1 < p < \infty$ we defined T "abstractly" as an operator on $L^p(\mathbb{R}^d)$ by extension from $\mathcal{S}(\mathbb{R}^d)$ via the apriori bounds $\|Tf\|_p \leqslant C_{p,d}\|f\|_p$ for all $f \in \mathcal{S}(\mathbb{R}^d)$ (the latter space is dense in $L^p(\mathbb{R}^d)$, cf. Proposition 1.5). We now ask if the principal value (7.1) converges almost everywhere to this extension T for any $f \in L^p(\mathbb{R}^d)$. Based on our previous experience with such questions, we introduce the associated maximal operator

$$T_* f(x) = \sup_{\varepsilon>0} \Big| \int_{|y|>\varepsilon} K(y) f(x-y)\, dy \Big|$$

and seek weak-L^1 bounds for it. As in previous cases where we studied almost everywhere convergence, this will suffice to conclude the a.e. convergence on L^p. In what follows we shall need the Hardy-Littlewood maximal function

$$Mf(x) = \sup_{x\in B} \frac{1}{|B|} \int_B |f(y)|\, dy$$

where the supremum runs over all balls B containing x. M satisfies the same type of bounds as those stated in Proposition 2.5.

We prove the desired bounds on T_* only for a subclass of kernels, namely the homogeneous ones. This class includes the Riesz transforms from above. More precisely, let

$$K(x) := \frac{\Omega(\frac{x}{|x|})}{|x|^d} \quad \text{for} \quad x \neq 0 \tag{7.20}$$

where $\Omega : S^{d-1} \to \mathbb{C}$, $\Omega \in C^1(S^{d-1})$ and $\int_{S^{d-1}} \Omega(x)\sigma(dx) = 0$ where σ is the surface measure on S^{d-1}. Observe the homogeneity $K(tx) = t^{-d}K(x)$ for all $t > 0$.

Exercise 7.6. Check that any such K satisfies the conditions in Definition 7.1.

Proposition 7.10. *Suppose K is of the form* (7.20). *Then T_* satisfies*

$$(T_* f)(x) \leqslant C[M(Tf)(x) + Mf(x)]$$

with some absolute constant C. In particular, T_* is bounded on $L^p(\mathbb{R}^d)$ for $1 < p < \infty$. Furthermore, T_* is also weak-L^1 bounded.

Proof. Let $\tilde{K}(x) := K(x)\chi_{[|x|\geqslant 1]}$ and more generally

$$\tilde{K}_\varepsilon(x) = \varepsilon^{-d}\tilde{K}(\frac{x}{\varepsilon}) = K(x)\chi_{[|x|\geqslant\varepsilon]}$$

Pick a smooth bump function $\varphi \in C_0^\infty(\mathbb{R}^d)$, $\varphi \geqslant 0$, $\int_{\mathbb{R}^d} \varphi\, dx = 1$. Set

$$\Phi := \varphi * K - \tilde{K}$$

and observe that $\varphi * K$ is well-defined in the principal value sense. For any function F on $\mathbb{R}^d$ let $F_\varepsilon(x) := \varepsilon^{-d}F\left(\frac{x}{\varepsilon}\right)$ be its L^1-normalized rescaling. Then $K_\varepsilon = K$, $\tilde{K}_\varepsilon = (\tilde{K})_\varepsilon$, and thus

$$\Phi_\varepsilon = (\varphi * K)_\varepsilon - \tilde{K}_\varepsilon = \varphi_\varepsilon * K_\varepsilon - \tilde{K}_\varepsilon = \varphi_\varepsilon * K - \tilde{K}_\varepsilon$$

Hence, for any $f \in \mathcal{S}(\mathbb{R}^d)$, one has

$$K_\varepsilon * f = \varphi_\varepsilon * (K * f) - \Phi_\varepsilon * f$$

We now invoke the analogue of Lemma 2.7 for radially bounded approximate identities in $\mathbb{R}^d$. This of course requires that $\{\Phi_\varepsilon\}_{\varepsilon>0}$ from such a radially bounded approximate identity, the verification of which we leave to the reader (the case of φ_ε is obvious). Indeed, one checks that

$$|\Phi(x)| \leqslant C \min(1, |x|^{-d-1})$$

which implies the desired property. Therefore,

$$T_* f \leqslant C(M(Tf) + Mf)$$

as claimed. The boundedness of T_* on $L^p(\mathbb{R}^d)$ for $1 < p < \infty$ now follows from that of T and M. The proof of the weak-L^1 boundedness of T_* is a variation of the same property of T, of Proposition 7.5. □

Exercise 7.7. Show that $\{\Phi_\varepsilon\}_{\varepsilon>0}$ as it appears in the previous proof forms a radially bounded approximate identity.

Corollary 7.11. *Let K be a homogeneous singular integral kernel as in* (7.20). *Then for any $f \in L^p(\mathbb{R}^d)$, $1 \leqslant p < \infty$, the limit in* (7.1) *exists almost everywhere.*

Proof. Let

$$\Lambda(f)(x) = |\limsup_{\varepsilon\to 0}(T_\varepsilon f)(x) - \liminf_{\varepsilon\to 0}(T_\varepsilon f)(x)|$$

Observe that $\Lambda(f) \leqslant 2T_* f$. Fix $f \in L^p$ and let $g \in \mathcal{S}(\mathbb{R}^d)$ so that

$$\|f - g\|_p < \delta$$

for a given small $\delta > 0$. Then $\Lambda(f) = \Lambda(f - g)$ and therefore

$$\|\Lambda f\|_p \leqslant C\|\Lambda(f - g)\|_p < C\delta$$

if $1 < p < \infty$. Hence, $\Lambda f = 0$. The case $p = 1$ is similar. □

Operators of the form (7.20) arise very naturally by considering *multiplier operators* of the form $T_m f := (m\hat{f})^\vee$ where m is homogeneous of degree zero and belongs to $C^\infty(\mathbb{R}^d\setminus\{0\})$ (for simplicity we assume infinite regularity). Indeed, with $f \in \mathcal{S}(\mathbb{R}^d)$ one has over the tempered distributions

$$Tf = K * f, \qquad K = \check{m}$$

where $\check{m} \in \mathcal{S}'(\mathbb{R}^d)$ is the distributional inverse Fourier transform. From the dilation law 6.2 one sees that K is homogeneous of degree $-d$ which means that for any $\lambda > 0$

$$\begin{aligned}\langle \widehat{m(\lambda\cdot)}, \varphi\rangle &= \langle m, \lambda^{-d}\hat{\varphi}(\lambda^{-1}\cdot)\rangle = \langle m, \widehat{\varphi(\lambda\cdot)}\rangle \\ &= \langle \hat{m}, \varphi(\lambda\cdot)\rangle = \langle \lambda^{-d}\hat{m}(\lambda^{-1}\cdot), \varphi\rangle\end{aligned}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Observer that constants are homogeneous of degree zero, whence $\hat{1} = \delta_0$ is homogeneous of degree $-d$. This might seem less strange if one remembers that as $\varepsilon \to 0$ one has

$$\frac{1}{|B(0,\varepsilon)|}\chi_{B(0,\varepsilon)} \to \delta_0 \text{ in } \mathcal{S}'(\mathbb{R}^d)$$

since the left-hand scales like a function that is homogeneous of degree $-d$. We now prove the following general statement which is very natural in view of the fact that the representation of $\hat{m}$ needs to be invariant under adding constants to m.

Lemma 7.12. *Let $m \in C^\infty(\mathbb{R}^d\backslash\{0\})$ be homogeneous of degree zero. Let $\langle m\rangle_{S^{d-1}}$ denote the mean of m on S^{d-1}. Then $\hat{m} \in \mathcal{S}'(\mathbb{R}^d)$ satisfies*

$$\hat{m} = \langle m\rangle_{S^{d-1}}\,\delta_0 + \text{P.V.}\,K_\Omega$$

where K_Ω is of the form (7.20) *with $\langle\Omega\rangle_{S^{d-1}} = 0$ and $\Omega \in C^\infty(S^{d-1})$.*

Proof. We may assume that m has mean zero on the sphere, i.e.,

$$\int_{S^{d-1}} m(\omega)\,\sigma(d\omega) = 0$$

Observe that $u_j := \partial_j^d m \in \mathcal{S}'(\mathbb{R}^d)$ is homogeneous of degree $-d$ and has mean zero on S^{d-1}. This implies that P.V. u_j is well-defined in the usual integral sense, and moreover, one has

$$u_j = \text{P.V.}\,u_j + \sum_\alpha c_\alpha\,\partial^\alpha\delta_0$$

by Exercise 6.2 where the sum is finite. In fact, by homogeneity, the sum is necessarily of the form $c_j\,\delta_0$. Passing to the Fourier transforms in $\mathcal{S}'$ one obtains

$$\widehat{u_j} = \widehat{\text{P.V.}\,u_j} + c_j$$

Next, one verifies that $\widehat{\text{P.V.}\,u_j}(x)$ is a smooth function on $\mathbb{R}^d\backslash\{0\}$ which is homogeneous of degree zero. Indeed, with $\chi(\xi)$ a smooth radial, compactly supported function which equals 1 on a neighborhood of the origin one has for any $x \neq 0$

$$\begin{aligned}\widehat{\text{P.V.}\,u_j}(x) &= \int_{\mathbb{R}^d} \partial_j^d m(\xi)\chi(\xi)(e^{-2\pi i x\cdot\xi} - 1)\,d\xi \\ &\quad + \sum_{j=1}^d \frac{x_j}{2\pi i|x|^2}\int_{\mathbb{R}^d} \partial_j\big(\partial_j^d m(\xi)(1-\chi(\xi))\big)e^{-2\pi i x\cdot\xi}\,d\xi\end{aligned}$$

where $p_j = \frac{\partial}{\partial\xi_j}$. Both integrals are absolutely convergent, and the first integral clearly defines a smooth function in x. We may repeat the integration by parts in the second integral any number of times, which implies that the second integral also defines a smooth function. The homogeneity of degree zero follows from the homogeneity of P.V. u_j. We may therefore write

$$\hat{m}(x) = \frac{\Omega\big(\frac{x}{|x|}\big)}{|x|^d} \qquad x \neq 0$$

where $\Omega \in C^\infty(S^{d-1})$. Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ be radial. Then

$$\langle \hat{m}, \varphi \rangle = \langle m, \hat{\varphi} \rangle = 0$$

since m has mean zero on spheres. Thus, Ω also has mean zero on S^{d-1}. By construction, $\hat{m} - \text{P.V.}\, K_\Omega$ is supported at $\{0\}$. By Exercise 6.2 we conclude that this difference is therefore a polynomial in derivatives of δ_0. By homogeneity, it must be a constant. By the vanishing means, this constant is zero and we are done. $\square$

As a consequence, we see that singular integral operators of the type T_m form an algebra, with T_m being invertible if and only if $m \neq 0$ on S^{d-1}. In Chapter 10 we shall see that one can still obtain L^p boundedness of T_m without the strong assumption of m being homogeneous of degree 0. Instead, one requires that $|\xi|^k |\partial^\alpha m(\xi)|$ is bounded for all $|\alpha| = k$ for finitely many k. While these conditions are satisfied if m is homogeneous of degree zero, they are much weaker than that property.

We now wish to *quantify the failure* of the boundedness of singular integral operators on L^∞. It turns out that the correct extension of L^∞ which contains $T(L^\infty)$ for any singular integral operator T is the space of functions of *bounded mean oscillation* on $\mathbb{R}^d$, denoted by $\text{BMO}(\mathbb{R}^d)$. In what follows,

$$\fint_Q f(y)\, dy = |Q|^{-1} \int_Q f(y)\, dy$$

DEFINITION 7.13. Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$. Define

$$f^\sharp(x) := \sup_{Q \ni x} \fint_Q |f(y) - f_Q|\, dy \tag{7.21}$$

where the supremum runs over cubes Q and f_Q is the mean of f over Q. Then $f \in \text{BMO}(\mathbb{R}^d)$ if and only if $f^\sharp \in L^\infty(\mathbb{R}^d)$. Then $\|f\|_{\text{BMO}} := \|f^\sharp\|_\infty$. If Q_0 is a fixed cube, then $\text{BMO}(Q_0)$ is obtained by defining $f^\sharp$ in terms of all subcubes of Q_0.

Of course one may replace cubes here with balls without changing anything. Moreover, $\|\cdot\|_{\text{BMO}}$ is a norm only after factoring out the constants. Clearly, $L^\infty \subset \text{BMO}$ but BMO is strictly larger. In addition, BMO scales like L^∞, i.e., $\|f(\lambda\cdot)\|_{\text{BMO}} = \|f\|_{\text{BMO}}$. Suppose $Q_2 \subset Q_1$ are two cubes with Q_2 being half the size of Q_1. Then

$$|f_{Q_1} - f_{Q_2}| \leqslant \frac{|Q_1|}{|Q_2|} \fint_{Q_1} |f(y) - f_{Q_1}|\, dy \leqslant 2^d \|f\|_{\text{BMO}} \tag{7.22}$$

Iterating this relation leads to the following: suppose $Q_n \subset Q_{n-1} \subset \ldots Q_1 \subset Q_0$ is a sequence of nested cubes so that the diameters decrease by a factor of 2 at each step. Then

$$|f_{Q_n} - f_{Q_0}| \leqslant n\, 2^d \|f\|_{\text{BMO}} \tag{7.23}$$

The important feature here is that while the size of Q_n is by a factor of 2^{-n} smaller than Q_0, the averages increase only linearly in n. As an example, consider the function f defined on the interval $[0,1]$ as follows

$$g = \sum_{n=0}^{\infty} a_n \chi_{[0,2^{-n}]} \tag{7.24}$$

with $1 \leqslant a_n$ for all $n \geqslant 0$.

Exercise 7.8.

- Show that $g \in \mathrm{BMO}([0,1])$ if and only if $a_n = O(1)$. In that case, verify that $g(x) \simeq |\log x|$ for $0 < x < \frac{1}{2}$.
- Show that $\chi_{[|x|<1]} \log |x| \in \mathrm{BMO}([-1,1])$, but

  $$\chi_{[|x|<1]} \operatorname{sign}(x) \log |x| \notin \mathrm{BMO}([-1,1])$$

- Show that for any f as in Definition 7.13

  $$f^{\sharp\sharp}(x) := \sup_{x\in Q} \inf_c \fint_Q |f(y) - c|\, dy$$

  satisfies $f^{\sharp\sharp} \leqslant f^{\sharp} \leqslant 2f^{\sharp\sharp}$.

We shall not distinguish between $f^{\sharp}$ and $f^{\sharp\sharp}$. It is easy to show that singular integrals take $L^\infty \to \mathrm{BMO}$.

Theorem 7.14. *Let T be a singular integral operator as in Definition 7.1. Then*

$$\|Tf\|_{\mathrm{BMO}} \leqslant CB\|f\|_\infty \qquad \forall\ f \in L^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$$

We are assuming here that $f \in L^2(\mathbb{R}^d)$ so that Tf is well-defined.

Proof. We may assume that $B = 1$ in Definition 7.1. Fix some f as above, a ball $B_0 := B(x_0, R) \subset \mathbb{R}^d$ and define

$$c_{B_0} := \int_{[|y-x_0|>2R]} K(x_0 - y) f(y)\, dy$$

Then, with $B_0^* := B(x_0, 2R)$, one estimates

$$\begin{aligned} \int_{B_0} |(Tf)(x) - c_{B_0}|\, dx &\leqslant \int_{B_0} \int_{[|y-x_0|>2R]} |K(x-y) - K(x_0-y)||f(y)|\, dy dx \\ &\quad + \int_{B_0} T\big(\chi_{B_0^*} f\big)(x)\, dx \\ &\leqslant C|B_0|\|f\|_\infty + |B_0|^{\frac{1}{2}} \|T\|_{2\to 2} \|\chi_{B_0^*} f\|_2 \\ &\leqslant C|B_0|\|f\|_\infty \end{aligned}$$

and we are done. □

In fact, it turns out that BMO is the smallest space with contains $T(L^\infty(\mathbb{R}^d))$ for every singular integral operator T as in Definition 7.1, see the notes to this chapter. If BMO is supposed to be a useful substitute to $L^\infty(\mathbb{R}^d)$, then we would like to be able to use it in interpolation theory. This is indeed possible, as the following theorem shows.

THEOREM 7.15. *Let T be a linear operator bonded on $L^{p_0}(\mathbb{R}^d)$ for some $1 \leqslant p_0 < \infty$ and bounded from $L^\infty \to$ BMO. Then T is bounded on $L^p(\mathbb{R}^d)$ for any $p_0 < p < \infty$.*

The proof of this result requires some machinery, more specifically a comparison between the Hardy-Littlewood maximal function and the sharp maximal function as in (7.21). For this purpose it is convenient to use the *dyadic maximal function*, which we denote by M_{dyad} and which is defined as

$$(M_{\mathrm{dyad}} f)(x) := \sup_{x \ni Q} \fint_Q |f(y)|\, dy$$

where the supremum now runs over dyadic cubes which are defined as the collection

$$\mathcal{Q}_{\mathrm{dyad}} := \left\{ 2^k[0,1)^d + 2^k\mathbb{Z}^d \mid k \in \mathbb{Z} \right\}$$

Note that any two dyadic cubes are either disjoint, or one contains the other. Suppose $Q_0 \in \mathcal{Q}_{\mathrm{dyad}}$, and let $f \in L^1(Q_0)$. Then for any

$$\lambda > \fint_{Q_0} |f(y)|\, dy \tag{7.25}$$

we can perform a Calderón-Zygmund decomposition as in Lemma 7.4 to conclude that

$$\{x \in Q_0 \mid (M_{\mathrm{dyad}} f)(x) > \lambda\} = \bigcup_{Q \in \mathcal{B}} Q$$

whence in particular the measure of the left-hand side is $< \lambda^{-1}\|f\|_{L^1(Q_0)}$. In other words, we have reestablished the weak-L^1 bound for the maximal function. The condition (7.25) serves as a "starting condition" for the stopping-time construction in Lemma 7.4. The other values of λ are of no interest, as in that case

$$\lambda^{-1}\|f\|_{L^1(Q_0)} \geqslant |Q_0|$$

so the weak-L^1 bound becomes trivial. The construction of course does not necessarily need to be localized to any Q_0, but can also be carried out on $\mathbb{R}^d$ globally.

We now come to the aforementioned comparison between $M_{\mathrm{dyad}} f$ and $f^\sharp$. While it is clear that $f^\sharp \leqslant 2Mf$, the usual Hardy-Littlewood maximal function, we require a lower bound. In essence, the following theorem says that on L^p with finite p, the sharp function contains no new information.

THEOREM 7.16. *Let $1 \leqslant p_0 \leqslant p < \infty$ and suppose $f \in L^{p_0}(\mathbb{R}^d)$. Then*

$$\int_{\mathbb{R}^d} (M_{\mathrm{dyad}} f)^p(x)\, dx \leqslant C(p,d) \int_{\mathbb{R}^d} \left(f^\sharp(x)\right)^p dx \tag{7.26}$$

Proof. For simplicity, we write M instead of M_{dyad}. The proof is an immediate consequence the following estimate, which is an example of a *good λ inequality*: for all $\gamma > 0$ and $\lambda > 0$ one has

$$\begin{aligned}|\{x \in \mathbb{R}^d \mid Mf(x) > 2\lambda,\ f^\sharp(x) \leqslant \gamma\lambda\}| \\ \leqslant 2^d\gamma|\{x \in \mathbb{R}^d \mid Mf(x) > \lambda\}|\end{aligned} \tag{7.27}$$

To prove it, we write the set on the right-hand side as $\bigcup_{Q\in\mathcal{B}} Q$ with dyadic Q via Lemma 7.4. We use $f \in L^{p_0}$ here since this ensures that we may start the stopping time argument at any level $\lambda > 0$ by taking the size of the first dyadic cube very large. By construction, these Q are the *maximal dyadic cubes* contained in $\{Mf > \lambda\}$. Indeed, if $\tilde{Q} \supset Q$ denotes the unique dyadic cube of twice the side-length of Q (the parent cube) then by definition of $\mathcal{B}$ one has

$$⨍_{\tilde{Q}} |f(y)|\,dy \leqslant \lambda \ \text{ and } \ \lambda < ⨍_Q |f(y)|\,dy \leqslant 2^d\lambda$$

Now fix one such Q and its parent $\tilde{Q}$. Then by the preceding, if $x \in Q$ and

$$(Mf)(x) > 2\lambda \Longrightarrow (M(f - \chi_Q f_{\tilde{Q}}))(x) > \lambda$$

Therefore,

$$\begin{aligned}|\{x \in Q \mid (Mf)(x) > 2\lambda\}| &\leqslant |\{x \in Q \mid (M(f - \chi_Q f_{\tilde{Q}}))(x) > \lambda\}| \\ &\leqslant \lambda^{-1} \int_Q |f - f_{\tilde{Q}}|(y)\,dy \leqslant \gamma|\tilde{Q}|\end{aligned}$$

To pass to the final bound, we used that if there exists some $x_0 \in Q$ so that $f^\sharp(x_0) \leqslant \gamma\lambda$ (which we may assume), then

$$\int_Q |f - f_{\tilde{Q}}|(y)\,dy \leqslant \int_{\tilde{Q}} |f - f_{\tilde{Q}}|(y)\,dy \leqslant \gamma\lambda|\tilde{Q}|$$

Summing over all $Q \in \mathcal{B}$ establishes the claim (7.27). Now

$$\begin{aligned}&\int_0^{\lambda_0} |\{Mf(x) > 2\lambda\}|\, p\lambda^{p-1}\,d\lambda \\ &\leqslant \int_0^{\lambda_0} |\{Mf(x) > 2\lambda,\ f^\sharp(x) \leqslant \gamma\lambda\}|\, p\lambda^{p-1}\,d\lambda + \int_0^{\lambda_0} |\{f^\sharp(x) > \gamma\lambda\}|\, p\lambda^{p-1}\,d\lambda \\ &\leqslant 2^{d+p}\gamma \int_0^{\lambda_0} |\{Mf(x) > 2\lambda\}|\, p\lambda^{p-1}\,d\lambda + \gamma^{-p}\|f^\sharp\|_p^p\end{aligned}$$

Choosing $\gamma = 2^{-d-p-1}$ and letting $\lambda_0 \to \infty$ concludes the proof. □

It is now easy to prove the interpolation result.

Proof of Theorem 7.15. Consider $S : f \mapsto (Tf)^\sharp$. While it is not linear it is sublinear which is sufficient for Marcinkiewicz interpolation. By definition, S is bounded on L^{p_0} and L^∞ whence it is also bounded on $L^p(\mathbb{R}^d)$ for any $p_0 < p < \infty$. Now take $f \in L^p \cap L^{p_0}$. Then $Tf \in L^{p_0}$ and we conclude from Theorem 7.16 and $Tf \leqslant M_{\mathrm{dyad}}(Tf)$ a.e. that T is bounded on L^p. □

We now take a closer look at the property (7.23). In fact, as suggested by Exercise 7.8 we shall now show that any BMO function is exponentially integrable. The precise statement is given by the following *John-Nirenberg* inequality.

THEOREM 7.17. *Let $f \in \mathrm{BMO}(\mathbb{R}^d)$. Then for any cube Q one has*

$$|\{x \in Q \mid |f(x) - f_Q| > \lambda\}| \leqslant C|Q| \exp\Big(- c\frac{\lambda}{\|f\|_{\mathrm{BMO}}}\Big)$$

for any $\lambda > 0$. The constants c, C are absolute.

PROOF. We may take $Q = [0,1)^d$. Assume $\|f\|_{\mathrm{BMO}} \leqslant 1$ with $f_Q = 0$, and let $\lambda > 10$. Perform a Calderón-Zygmund decomposition of f on the cube Q at a level $\mu \geqslant 2$ that we shall fix later. This yields $f = g + b$, $|g| \leqslant 2^d\mu$ and

$$b = \sum_{Q' \in \mathcal{B}} \chi_{Q'}(f - f_{Q'}), \qquad \sum_{Q' \in \mathcal{B}} |Q'| \leqslant \mu^{-1}$$

Define

$$E(\lambda) := \sup_{\|g\|_{\mathrm{BMO}(Q)} \leqslant 1} |\{x \in Q \mid |g(x) - g_Q| > \lambda\}|$$

Then

$$\begin{aligned} E(\lambda) &\leqslant |\{x \in Q \mid |b(x)| > \lambda - 2^d\mu\}| \\ &\leqslant \sum_{Q' \in \mathcal{B}} |\{x \in Q' \mid |f(x) - f_{Q'}| > \lambda - 2^d\mu\}| \\ &\leqslant \sum_{Q' \in \mathcal{B}} E(\lambda - 2^d\mu)|Q'| \leqslant \mu^{-1}E(\lambda - 2^d\mu) \end{aligned} \tag{7.28}$$

Iterating this relation implies

$$E(\lambda) \leqslant \mu^{-n}E(\lambda - 2^d\mu n)$$

Setting for example $\mu = 2$ leads to the desired bound. □

EXERCISE 7.9. Show that one can characterize BMO also by the following condition, where $1 \leqslant r < \infty$ is fixed but arbitrary:

$$\sup_Q \Big(\fint_Q |f(y) - f_Q|^r \, dy \Big)^{\frac{1}{r}} < \infty \tag{7.29}$$

In fact, show that if the left-hand side is $\leqslant 1$, then $\|f\|_{\mathrm{BMO}} \leqslant C(d,r)$. Conversely, by Hölder's $\|f\|_{\mathrm{BMO}} \leqslant 1$ implies that the left-hand side of (7.29) is $\leqslant 1$.

A remarkable characterization of BMO was found by Coifman, Rochberg and Weiss. The considered commutators between Calderón-Zygmund operators and operators given by multiplication by a function, in other words, $[T, b] = Tb - bT$. While it is clear that $[T, b]$ is bounded on L^p for $1 < p < \infty$ if $b \in L^\infty$, these authors discovered that this property remains true for $b \in \mathrm{BMO}(\mathbb{R}^d)$. What is more, if this property holds for a single Riesz-transform (or a single nonzero operator of the form (7.20), then necessarily $b \in \mathrm{BMO}$. We shall now prove one direction of this statement.

THEOREM 7.18. *Let T be a linear operator as in Proposition 7.5 with the added property that for some fixed* $0 < \delta \leqslant 1$ *one has*

$$|K(x-y) - K(x_0 - y)| \leqslant A \frac{|x-x_0|^\delta}{|y-x_0|^{d+\delta}} \qquad \forall\ |y-x_0| > 2|x-x_0| \tag{7.30}$$

as well as $\|T\|_{2\to 2} \leqslant A$. *Then for any* $b \in \mathrm{BMO}(\mathbb{R}^d)$ *one has*

$$\|[T,b]\|_{p\to p} \leqslant C(d,p) A \|b\|_{\mathrm{BMO}}$$

for all $1 < p < \infty$.

PROOF. Since (7.30) implies the Hörmander condition, we see that any operator T as in the theorem is a Calderón-Zygmund operator. We can also assume that $\|b\|_{\mathrm{BMO}} \leqslant 1$ and $A = 1$. Fix $f \in \mathcal{S}(\mathbb{R}^d)$, say. We shall show that for any $1 < r < \infty$

$$([T,b]f)^\sharp \leqslant C(r,d)(M_r(Tf) + M_r f) \tag{7.31}$$

where

$$(M_r f)(x) := \sup_{Q \ni x} \Big(\fint_Q |f(y)|^r \, dy \Big)^{\frac{1}{r}}$$

Letting $1 < r < p < \infty$ and noting that M_r is L^p-bounded, we conclude from (7.31) that $\|([T,b]f)^\sharp\|_p \leqslant C(d,p)\|f\|_p$ and Theorem 7.16 implies the desired estimate (the p_0-condition in that theorem is clear by Theorem 7.17 and $f \in \mathcal{S}(\mathbb{R}^d)$).

To establish (7.31) fix any cube Q, which by scaling and translation invariance may be taken to be the unit cube centered at 0. Denote by $\tilde{Q}$ the cube of side-length $2\sqrt{d}$, and write

$$[T,b]f = T\big((b - b_{\tilde{Q}})f_1\big) + T\big((b-b_{\tilde{Q}})f_2\big) - (b - b_{\tilde{Q}})Tf =: g_1 + g_2 + g_3$$

where $f_1 := \chi_{\tilde{Q}} f$, $f_2 = f - f_1$. First, by Exercise 7.9

$$\fint_Q |g_3(x)| \, dx \leqslant \Big(\fint_Q |b - b_Q|^{r'} \, dx \Big)^{\frac{1}{r'}} \Big(\fint_Q |Tf|^r(x) \, dx \Big)^{\frac{1}{r}} \leqslant C \inf_Q M_r(Tf)$$

Second, choose some $1 < s < r$. Then by boundedness of T on L^s one obtains

$$\begin{aligned} \fint_Q |g_1(x)| \, dx &\leqslant \Big(\fint_Q |g_1(x)|^s \, dx \Big)^{\frac{1}{s}} \leqslant \Big(\fint_{\tilde{Q}} |b - b_Q|^s |f(x)|^s \, dx \Big)^{\frac{1}{s}} \\ &\leqslant C \inf_Q M_r f \end{aligned}$$

where we used (7.22) to compare b_Q with $b_{\tilde{Q}}$. Finally, set

$$c_Q := T\big((b - b_{\tilde{Q}})f_2\big)(0)$$

Then by (7.30) one has the estimate

$$
\begin{aligned}
&⨍_Q |g_2(x) - c_Q|\, dx \\
&\leqslant C \int_Q \int_{\tilde{Q}^c} |K(x-y) - K(0-y)|\, |b(y) - b_{\tilde{Q}}|\, |f(y)|\, dydx \\
&\leqslant C \sum_{\ell \geqslant 0} \int_{2^{\ell+1}\tilde{Q} \setminus 2^{\ell}\tilde{Q}} 2^{-\ell(d+\delta)} |b(y) - b_Q|\, |f(y)|\, dy \\
&\leqslant C \sum_{\ell \geqslant 0} 2^{-\ell\delta} \Big(⨍_{2^{\ell+1}\tilde{Q}} |b(y) - b_Q|^{r'}\, dy \Big)^{\frac{1}{r'}} \Big(⨍_{2^{\ell+1}\tilde{Q}} |f(y)|^{r}\, dy \Big)^{\frac{1}{r}} \\
&\leqslant C \sum_{\ell \geqslant 0} \ell\, 2^{-\ell\delta} \inf_Q M_r f \leqslant C \inf_Q M_r f
\end{aligned}
$$

where we used (7.23) to pass to the final estimate. □

We have merely scratched the surface of BMO theory. Of course one would like to determine a duality relation analogous to the basic $(L^1(\mathbb{R}^d)^* = L^\infty(\mathbb{R}^d)$. This turns out to be $(H^1(\mathbb{R}^d))^* = \mathrm{BMO}(\mathbb{R}^d)$ where $H^1(\mathbb{R}^d)$ is the completion of $\mathcal{S}(\mathbb{R}^d)$ under the norm

$$\|f\|_{H^1} := \|f\|_1 + \sum_{j=1}^{d} \|R_j f\|_1 \tag{7.32}$$

Dually, one can then write

$$\mathrm{BMO} = L^\infty + \sum_{j=1}^{d} R_j L^\infty \tag{7.33}$$

where the action of the Riesz transforms on L^∞ needs to be defined suitably. In particular, this shows that BMO is the smallest space into which singular integral operators are bounded from L^∞. We shall establish some of these facts in a one-dimensional dyadic setting in a later chapter.

We close this introductory chapter by stating a number of natural questions. First, we may ask if a Calderón-Zygmund operator is bounded on Hölder spaces $C^\alpha(\mathbb{R}^d)$ where $0 < \alpha < 1$. As we shall see, this turns out to be the case. The analogue of Corollary 7.7 in this case is known as *Schauder estimates* which play a fundamental role in elliptic equations. There are a number of proofs known of this classical fact, and we shall develop one possible approach in Chapter 10.

Another question that one may ask, and which turned out to have far-reaching consequences, is as follows: is there a proof of the L^2-boundedness of the Hilbert transform which does not rely on the Fourier transform? In other words, we wish to find a proof of the L^2-boundedness which relies exclusively on the properties of the kernel itself. We shall answer this question in the affirmative in the larger context of Calderón-Zygmund operators, and introduce the important method of *almost orthogonality* in the process, see Chapter 8.

And finally, one should of course ask if there is a version of Calderón-Zygmund theory for kernels which are not translation invariant. This question has far-reaching consequences especially with respect to the L^2 boundedness. While it is true that the L^p-theory is quite similar to the translation invariant case (see Problem 7.3), the L^2 theory leads to the $T(1)$ theorem. We shall develop the basic ideas for that theorem by means of a dyadic model on the line.

Notes

For most of the basic material of this chapter, excluding BMO, see Stein's 1970 book [**45**]. For a systematic development of H^1 and BMO see [**46**] which gives a systematic account of many aspects of Calderón-Zygmund theory that appeared from 1970 to roughly 1992. A classic reference for H^1 and the duality with BMO is the paper by Fefferman and Stein [**16**], which contains (7.32). Uchiyama [**55**] gives a constructive proof of the decomposition of a BMO function as a sum of L^∞ and images thereof under the Riesz transforms, see (7.33). Coifman and Meyer [**10**] offer another perspective of Calderón-Zygmund theory with many applications and connections with wavelets. The proof of the Coifman, Rochberg, Weiss commutator estimate is from Janson [**28**] (credited to Strömberg), see also Torchinsky [**54**]. One important aspect of Calderón-Zygmund theory that we do not discuss in this text is that of A_p-weights, which is closely related to BMO. For this material see Stein [**46**], García-Cuerva, Rubio de Francia [**7**], as well as Duoandikoetxea [**12**]. The classical "method or rotations" which can be used to reduce higher-dimensional results to one-dimensional ones, is discussed in [**12**].

Problems

Problem 7.1. Let $\{a_{ij}(x)\}_{i,j=1}^d$ be continuous in some domain $\Omega \subset \mathbb{R}^d$, $d \geqslant 2$, and suppose that these matrices are elliptic in the sense that for all $x \in \Omega$

$$\theta|\xi|^2 \leqslant \sum_{i,j=1}^d a_{ij}(x)\xi^i\xi^j \leqslant \Theta|\xi|^2 \qquad \forall\, \xi \in \mathbb{R}^d$$

for some fixed $0 < \theta < \Theta$. Assume that $u \in C^2(\Omega)$ satisfies

$$\sum_{i,j=1}^d a_{ij}(x)\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}u(x) = f(x)$$

Then show that for any compact $K \subset \Omega$ and any $1 < p < \infty$ one has the bound

$$\|D^2u\|_{L^p(K)} \leqslant C(\|u\|_{L^p(\Omega)} + \|f\|_{L^p(\Omega)})$$

where C only depends on p, K, Ω and the a_{ij}.

Problem 7.2. This problem shows how symmetries limit the boundedness properties of operators which commute with these symmetries. We shall consider translation and dilation symmetries, respectively.

a) Show that for *any* translation invariant non-zero operator $T : L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ one necessarily has $q \geqslant p$.

b) Show that a homogeneous kernel as that in (7.20) can only be bounded from

$$L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$$

if $p = q$.

PROBLEM 7.3. Let T be a bounded linear operator on $L^2(\mathbb{R}^d)$ so that for some function $K(x, y)$ measurable and locally bounded on $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$ with $\Delta = \{x = y\}$ being the diagonal, one has for all $f \in L^2_{\mathrm{comp}}(\mathbb{R}^d)$

$$(Tf)(x) = \int_{\mathbb{R}^d} K(x, y)\, f(y)\, dy, \qquad x \notin \mathrm{supp}(f)$$

Furthermore, assume the Hörmander conditions

$$\begin{aligned} &\int_{[|x-y|>2|y-y'|]} |K(x, y) - K(x, y')|\, dx \leqslant A \qquad \forall\ y \neq y' \\ &\int_{[|x-y|>2|x-x'|]} |K(x, y) - K(x', y)|\, dy \leqslant A \qquad \forall\ x \neq x' \end{aligned} \tag{7.34}$$

Then T is bounded from L^1 to weak-L^1 and strongly on L^p for every $1 < p < \infty$.

PROBLEM 7.4. This is the continuum analogue of Problem 4.3. Let m be a function of bounded variation on $\mathbb{R}$. Show that $T_m f = (m\hat{f})^\vee$ is bounded on $L^p(\mathbb{R})$ for any $1 < p < \infty$. Find a suitable generalization to higher dimensions, starting with $\mathbb{R}^2$.

PROBLEM 7.5. Let $f \in \mathrm{BMO}(\mathbb{R}^d)$. Show that for any $\sigma > d$ and $1 \leqslant r < \infty$ one has

$$\int_{\mathbb{R}^d} \frac{|f(x) - f_{Q_0}|^r}{1 + |x|^\sigma}\, dx \leqslant C(d, \sigma, r)\|f\|_{\mathrm{BMO}}$$

where Q_0 is any cube of side-length 1. What does one obtain by rescaling this inequality?

PROBLEM 7.6. Let $f \in L^1(\mathbb{R}^d)$ have compact support, say $B := B(0, 1)$. Let M be the Hardy-Littlewood maximal function. Show that $Mf \in L^1(B)$ if and only if $f \log(2 + |f|) \in L^1(B)$. Now redo Problem 4.6 based on this fact and the F.&M. Riesz theorem. *Hint:* Reprove the weak-L^1 bound on M by means of the Calderón-Zygmund decomposition. In particular, show that it in order to bound $|\{Mf > \lambda\}|$ it is enough to control the L^1 norm of f on $\{|f| > c\,\lambda\}$ where $c > 0$ is some constant. See Stein [**45**] page 23 for more details.

PROBLEM 7.7. Show that the Hilbert transform H with kernel $\frac{1}{x}$ is bounded on the Hölder spaces $C^\alpha(\mathbb{R})$, $0 < \alpha < 1$. In other words, prove that for any $0 < \alpha < 1$

$$[Hf]_\alpha \leqslant C(\alpha)\, [f]_\alpha$$

for all $f \in C^1_0(\mathbb{R})$, say. Here $[\cdot]_\alpha$ is the C^α-seminorm. Use only properties of the kernel for this problem.

PROBLEM 7.8. Let A, B be reflexive Banach spaces (a reader not familiar with this may take them to be Hilbert spaces). Let $K(x, y)$ be defined, measurable, and locally bounded on $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$ where $\Delta = \{x = y\}$ is the diagonal, and take its values in $\mathcal{L}(A, B)$, the bounded linear operators from A to B. Define an operator T on $C^0_{\mathrm{comp}}(\mathbb{R}^d; A)$, the continuous compactly supported functions taking their values in A, such that

$$(Tf)(x) = \int K(x, y) f(y)\, dy, \qquad x \notin \mathrm{supp}(f)$$

for all $f \in C^0_{\mathrm{comp}}(\mathbb{R}^d; A)$. Assume that for some finite constant C_0

$$\int_{|x-y|>2|x-x'|} \|K(x,y) - K(x',y)\|_{\mathcal{L}(A,B)}\, dy \leqslant C_0 \qquad \forall x \neq x'$$

$$\int_{|x-y|>2|y-y'|} \|K(x,y) - K(x,y')\|_{\mathcal{L}(A,B)}\, dx \leqslant C_0 \qquad \forall y \neq y'$$

Further, assume that $T : L^r(\mathbb{R}^d; A) \to L^r(\mathbb{R}^d; B)$ for some $1 < r < \infty$ with norm bounded by C_0. Show that then $T : L^p(\mathbb{R}^d; A) \to L^p(\mathbb{R}^d; B)$ for all $1 < p < \infty$ with norm bounded by $C_0\, C_1$ where $C_1 = C_1(d, r)$.

Problem 7.9. Investigate the endpoint $p = \frac{d}{\alpha}$ of Proposition 7.8. More precisely, find examples which show that fractional integration in that case is not bounded into L^∞. Verify that your examples get mapped into $\mathrm{BMO}\backslash L^\infty(\mathbb{R}^d)$. For the positive result which proves boundedness $L^p \to \mathrm{BMO}$ see Adams [**1**].

Problem 7.10. Let $f \in L^1(\mathbb{T})$. Given $\lambda > 1$, show that there exists $E \subset \mathbb{T}$ (depending on λ and f) so that $|E| < \lambda^{-1}$ and for all $N \in \mathbb{Z}^+$

$$\frac{1}{N} \int_{\mathbb{T}\backslash E} \sum_{n=1}^{N} |S_n f(\theta)|^2 \, d\theta \leqslant C\lambda \|f\|_1^2$$

C is a constant independent of f, N, λ. *Hint:* First apply the Plancherel theorem in n, and then use a Calderón-Zygmund decomposition.

CHAPTER 8

Almost Orthogonality

The proof of L^2 boundedness of Calderon-Zygmund operators in the previous chapter was based on the Fourier transform. This is quite restrictive as it requires the operator to be translation invariant. We now present a device that allows one to avoid the Fourier transform in the context of L^2-theory in many instances.

Let us start from the basic observation that the operator norm of an infinite diagonal matrix viewed as an operator on $\ell^2(\mathbb{Z})$ is as large as its largest entry. To be specific, let

$$T\big(\{\xi_j\}_{j\in\mathbb{Z}}\big) = \{\lambda_j\xi_j\}_{j\in\mathbb{Z}} \qquad \forall\ \{\xi_j\}_{j\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$$

where $\{\lambda_j\}_{j\in\mathbb{Z}}$ is a fixed sequence of complex numbers. Then

$$\|T\|_{\ell^2\to\ell^2} = \sup_j |\lambda_j|$$

More generally, suppose a Hilbert space $\mathcal{H}$ can be written as an infinite orthogonal sum

$$\mathcal{H} = \bigoplus_j \mathcal{H}_j, \quad \mathcal{H} \ni f = \sum_j f_j, \quad f_j \in \mathcal{H}_j$$

and suppose that T_j are operators on $\mathcal{H}$ with $T_j\mathcal{H}_k = \{0\}$ if $k \neq j$. Furthermore, assume that the range of T_j is a subspace of $\mathcal{H}_j$. Then

$$\begin{aligned}
\|Tf\|_{\mathcal{H}}^2 &= \sum_{j,k}\langle T_jf_j, T_kf_k\rangle = \sum_j \|T_jf_j\|_{\mathcal{H}_j}^2 \\
&\leqslant \sup_j \|T_j\|_{\mathcal{H}_j\to\mathcal{H}_j}^2 \sum_j \|f_j\|_{\mathcal{H}_j}^2 = M^2\|f\|_{\mathcal{H}}^2
\end{aligned}$$

where $M = \sup_j \|T_j\|_{\mathcal{H}_j\to\mathcal{H}_j}$.

Generalizing from the previous example, assume that T is a bounded operator on the Hilbert space $\mathcal{H}$ which admits the representation $T = \sum_j T_j$ such that $\mathrm{Ran}(T_j) \perp \mathrm{Ran}(T_k)$ as well as $\mathrm{Ran}(T_j^*) \perp \mathrm{Ran}(T_k^*)$ for $j \neq k$. In other words, we assume that $T_k^*T_j = 0$ and $T_kT_j^* = 0$ for $j \neq k$. Now recall that $\overline{\mathrm{Ran}(T_j^*)} = \ker(T_j)^\perp$ and denote by P_j the orthogonal projection onto that subspace. Then for any $f \in \mathcal{H}$ one has

$$Tf = \sum_j T_jP_jf = \sum_j T_jf_j, \quad f_j = P_jf$$

which implies that

$$\|Tf\|_{\mathcal{H}}^2 = \sum_j \|T_j f_j\|_{\mathcal{H}}^2 \leqslant \sup_j \|T_j\|^2 \sum_j \|f_j\|_{\mathcal{H}}^2 = M^2 \|f\|_{\mathcal{H}}^2$$

with $M := \sup_j \|T_j\|$.

As a final step, we shall now relax the orthogonality conditions $T_k^* T_j = 0$ and $T_k T_j^* = 0$ for $j \neq k$ from above, which turn out to be too restrictive for most applications. The idea is simply to replace this strong vanishing requirement by a condition which ensures sufficient decay in $|j-k|$. One can think of this as replacing diagonal matrices with matrices whose entries decay in a controllable fashion away from the diagonal. The following lemma, known as *Cotlar-Stein lemma*, carries this idea out.

Lemma 8.1. *Let $\{T_j\}_{j=1}^N$ be finitely many operators on L^2 such that for some function $\gamma : \mathbb{Z} \to \mathbb{R}$, one has*

$$\|T_j^* T_k\| \leqslant \gamma^2(j-k), \quad \|T_j T_k^*\| \leqslant \gamma^2(j-k)$$

for any $1 \leqslant j, k \leqslant N$. Let

$$\sum_{\ell=-\infty}^{\infty} \gamma(\ell) =: A < \infty$$

Then $\| \sum_{j=1}^N T_j \| \leqslant A$.

Proof. For any positive integer n,

$$(T^*T)^n = \sum_{\substack{j_1,\ldots,j_n=1 \\ k_1,\ldots,k_n=1}}^{N} T_{j_1}^* T_{k_1} T_{j_2}^* T_{k_2} \cdots T_{j_n}^* T_{k_n}$$

We now take the operator norm of both sides of this identity, and apply the triangle inequality on the right-hand side. Furthermore, we bound the norm of the products of operators by the products of the norms in two different ways, followed by taking the square root of the product of the resulting estimate.
Therefore, with $\sup_{1 \leqslant j \leqslant N} \|T_j\| =: B$

$$\begin{aligned}
&\|(T^*T)^n\| \\
&\leqslant \sum_{\substack{j_1,\ldots,j_n=1 \\ k_1,\ldots k_n=1}}^{N} \|T_{j_1}\|^{\frac{1}{2}} \|T_{j_1}^* T_{k_1}\|^{\frac{1}{2}} \|T_{k_1} T_{j_2}^*\|^{\frac{1}{2}} \cdots \|T_{k_{n-1}} T_{j_n}^*\|^{\frac{1}{2}} \|T_{j_n}^* T_{k_n}\|^{\frac{1}{2}} \|T_{k_n}\|^{\frac{1}{2}} \\
&\leqslant \sum_{\substack{j_1,\ldots,j_n=1 \\ k_1,\ldots,k_n=1}}^{N} \sqrt{B}\, \gamma(j_1-k_1)\gamma(k_1-j_2)\gamma(j_2-k_2)\cdots\gamma(k_{n-1}-j_n)\gamma(j_n-k_n)\ \sqrt{B} \\
&\leqslant NBA^{2n-1}
\end{aligned}$$

Since T^*T is self-adjoint, the spectral theorem implies that $\|(T^*T)^n\| = \|T^*T\|^n = \|T\|^{2n}$. Hence,

$$\|T\| \leqslant (N\,BA^{-1})^{\frac{1}{2n}} \cdot A$$

Letting $n \to \infty$ yields the desired bound. □

Note that one needs *both* smallness of $T_j^*T_k$ and $T_jT_k^*$ in the Cotlar-Stein lemma, as can be seen from the examples preceding the lemma.

In order to apply Lemma 8.1 one often invokes the following simple device.

Lemma 8.2. *Define an integral operator on a measure space $X \times Y$ with the positive product measure $\mu \otimes \nu$ via*

$$(Tf)(x) = \int_Y K(x,y)f(y)\,\nu(dy)$$

where K is a measurable kernel. One has the following bounds (the first three items are known as Schur's lemma):

- $\|T\|_{1\to 1} \leqslant \sup_{y\in Y} \int_X |K(x,y)|\,\mu(dx) =: A$
- $\|T\|_{\infty\to\infty} \leqslant \sup_{x\in X} \int_Y |K(x,y)|\,\nu(dy) =: B$
- $\|T\|_{p\to p} \leqslant A^{\frac{1}{p}} B^{\frac{1}{p'}}$ *where* $1 \leqslant p \leqslant \infty$.
- $\|T\|_{1\to\infty} \leqslant \|K\|_{L^\infty(X\times Y)}$

Proof. The first two items are immediate from the definitions, and the third then follows by interpolation. Alternatively, one can use Hölder's inequality. The fourth item is again evident from the definitions. □

Henceforth, we shall simply refer to this lemma as "Schur's test". We will now give an alternative proof of L^2 boundedness of singular integrals for kernels as in Definition 7.1 which satisfy the stronger condition

$$|\nabla K(x)| \leqslant B|x|^{-d-1}$$

cf. Lemma 7.2. This requires a certain standard partition of unity over a geometric scale which we now present.

Lemma 8.3. *There exists $\psi \in C^\infty(\mathbb{R}^d)$ with the property that* $\mathrm{supp}(\psi) \subset \mathbb{R}^d\setminus\{0\}$ *is compact and such that*

$$\sum_{j=-\infty}^{\infty} \psi(2^{-j}x) = 1 \qquad \forall\ x \neq 0 \tag{8.1}$$

For any given $x \neq 0$ at most two terms in this sum are nonzero. Moreover, ψ can be chosen to be a radial nonnegative function.

Proof. Let $\chi \in C^\infty(\mathbb{R}^d)$ satisfy $\chi(x) = 1$ for all $|x| \leqslant 1$ and $\chi(x) = 0$ for $|x| \geqslant 2$. Set $\psi(x) := \chi(x) - \chi(2x)$. For any positive N one has

$$\sum_{j=-N}^{N} \psi(2^{-j}x) = \chi(2^{-N}x) - \chi(2^{N+1}x)$$

If $x \neq 0$ is given, then we take N so large that $\chi(2^{-N}x) = 1$ and $\chi(2^{N+1}x) = 0$. This implies (8.1), and the other properties are immediate as well. □

The point of the following corollary is the method of proof rather than the statement (which is weaker than the one in the previous chapter).

COROLLARY 8.4. *Let K be as in Definition 7.1 with the additional assumption that $|\nabla K(x)| \leqslant B|x|^{-d-1}$. Then*

$$\|T\|_{2\to 2} \leqslant CB$$

with $C = C(d)$.

PROOF. We may take $B = 1$. Let ψ be a radial function as in Lemma 8.3 and set $K_j(x) = K(x)\psi(2^{-j}x)$. One now easily verifies that these kernels have the following properties: for all $j \in \mathbb{Z}$ one has

$$\int K_j(x)\,dx = 0,$$
$$\|\nabla K_j\|_\infty \leqslant C\,2^{-j}2^{-jd}$$

In addition, one has the estimates

$$\int |K_j(x)|\,dx < C$$
$$\int |x|\,|K_j(x)|\,dx < C\,2^j$$

with some absolute constant C. Define

$$(T_j f)(x) = \int_{\mathbb{R}^d} K_j(x-y)f(y)\,dy$$

Observe that this integral is absolutely convergent for any $f \in L^1_{\ell oc}(\mathbb{R}^d)$. We shall now check the conditions in Lemma 8.1. Let $\widetilde{K}_j(x) := \overline{K}_j(-x)$. Then it is easy to see that

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

and

$$(T_j T_k^* f)(x) = \int_{\mathbb{R}^d} (K_j * \widetilde{K}_k)(y) f(x-y)\,dy$$

Hence, by Young's inequality,

$$\|T_j^* T_k\|_{2\to 2} \leqslant \|\widetilde{K}_j * K_k\|_1$$

and

$$\|T_j T_k^*\|_{2\to 2} \leqslant \|K_j * \widetilde{K}_k\|_1$$

It suffices to consider the case $j \geqslant k$. Then, using the cancellation condition $\int K_k(y)\,dy = 0$, one obtains

$$\begin{aligned}\left|(\widetilde{K}_j * K_k)(x)\right| &= \left|\int_{\mathbb{R}^d} \overline{K}_j(y-x)K_k(y)\,dy\right| \\ &= \left|\int_{\mathbb{R}^d} \left[\overline{K}_j(y-x) - \overline{K}_j(-x)\right] K_k(y)\,dy\right| \\ &\leqslant \int_{\mathbb{R}^d} \|\nabla K_j\|_\infty |y|\,|K_k(y)|\,dy \\ &\leqslant C\,2^{-j}2^{-jd}2^k\end{aligned}$$

Since

$$\operatorname{supp}\,(\widetilde{K}_j * K_k) \subset \operatorname{supp}(\widetilde{K}_j) + \operatorname{supp}(K_k) \subset\ B(0, C\cdot 2^j)$$

we further conclude that

$$\|\widetilde{K}_j * K_k\|_1 \leqslant C\,2^{k-j} = C\,2^{-|j-k|}$$

Therefore, Lemma 8.1 applies with

$$\gamma^2(\ell) = C\,2^{-|\ell|}$$

and the corollary follows. □

Exercise 8.1.

a) In the previous proof it suffices to consider the case $j > k = 0$. Provide the details of this reduction.

b) Observe that Corollary 8.4 covers the Hilbert transform. In that case, draw the graph of $K(x) = \frac{1}{x}$ and also $K_j(x) = \frac{1}{x}\psi(2^{-j}x)$ and explain the previous argument by means of pictures.

Another simple application of these almost orthogonality ideas is the Calderon-Vaillancourt theorem. This result concerns the L^2-boundedness of *pseudo-differential operators* of the form

$$Tf(x) = \int_{\mathbb{R}^d} e^{ix\cdot\xi}\,a(x,\xi)\,\hat{f}(\xi)\,d\xi, \qquad f \in \mathcal{S}(\mathbb{R}^d) \tag{8.2}$$

where $a \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ is such that

$$\sup_{x,\xi\in\mathbb{R}^d} \left[|\partial_x^\alpha a(x,\xi)| + |\partial_\xi^\alpha a(x,\xi)|\right] \leqslant B \tag{8.3}$$

for all $|\alpha| \leqslant d+1$.

Proposition 8.5. *Under the conditions* (8.3) *the operators in* (8.2) *are bounded on* $L^2(\mathbb{R}^d)$.

Proof. We may assume that $B = 1$. Let χ be smooth and compactly supported such that $\{\chi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ forms a partition of unity in $\mathbb{R}^d$:

$$\sum_{k\in\mathbb{Z}^d} \chi(\xi - k) = 1 \qquad \forall\, \xi \in \mathbb{R}^d$$

To construct such a χ, start from a compactly supported smooth $\eta \geqslant 0$ with the property that

$$\psi(x) := \sum_{k\in\mathbb{Z}^d} \eta(x-k) > 0 \qquad \forall\, x \in \mathbb{R}^d$$

Note that ψ is periodic with respect to $\mathbb{Z}^d$ and therefore uniformly lower bounded. Now define $\chi := \frac{\chi}{\psi}$.

Set

$$\begin{aligned}\chi_{k\ell}(x,\xi) &:= \chi(x-k)\chi(\xi-\ell)\\ a_{k,\ell}(x,\xi) &:= a(x,\xi)\chi(x-k)\chi(\xi-\ell)\end{aligned}$$

and define

$$(T_{k,\ell}f)(x) := \int_{\mathbb{R}^d} e^{ix\cdot\xi}\, a_{k,\ell}(x,\xi)\, f(\xi)\, d\xi$$

for $f \in \mathcal{S}(\mathbb{R}^d)$. Note that we have dropped the Fourier transform of f, which is admissible by Plancherel's theorem. First, Lemma 8.2 implies that

$$\sup_{k,\ell} \|T_{k,\ell}\|_{2\to 2} \leqslant C\, 1$$

Furthermore, we claim that

$$\begin{aligned}\|T^*_{k',\ell'}T_{k,\ell}\|_{2\to2} &\leqslant C\langle k'-k\rangle^{-d-1}\,\langle \ell'-\ell\rangle^{-d-1}\\ \|T_{k',\ell'}T^*_{k,\ell}\|_{2\to2} &\leqslant C\langle k'-k\rangle^{-d-1}\,\langle \ell'-\ell\rangle^{-d-1}\end{aligned} \tag{8.4}$$

for all $k,k',\ell,\ell' \in \mathbb{Z}^d$. Since the bounds on the right-hand side are summable over $k,\ell \in \mathbb{Z}^d$, Lemma 8.1 implies that

$$\Big\| \sum_{|k|<N}\sum_{|\ell|<N} T_{k,\ell}f\Big\|_2 \leqslant C\|f\|_2 \qquad \forall\, f\in\mathcal{S}(\mathbb{R}^d)$$

for any positive integer N with some absolute constant C. Passing to the limit $N\to\infty$ concludes the proof. To prove (8.4) we start from

$$T^*_{k,\ell}g(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\overline{a(x,\xi)}\, g(x)\, dx$$

which shows that $T^*_{k',\ell'}T_{k,\ell} \neq 0$ requires that $|k-k'| \leqslant C$ and similarly, $T_{k',\ell'}T^*_{k,\ell}\neq 0$ requires that $|\ell-\ell'|\leqslant C$. The kernel of $T^*_{k',\ell'}T_{k,\ell}$ is

$$K_{k,\ell;k',\ell'}(\xi,\eta) := \int_{\mathbb{R}^d} e^{-ix\cdot(\xi-\eta)}\overline{a_{k,\ell}(x,\xi)}\, a_{k',\ell'}(x,\eta)\, dx \tag{8.5}$$

To prove the decay in $|\ell-\ell'|$ we may assume that this difference exceeds some constant which is chosen such that $|\xi-\eta|\geqslant 1$ on the support of $\overline{a_{k,\ell}(x,\xi)}\, a_{k',\ell'}(x,\eta)$. We now use the relation

$$i\frac{\xi-\eta}{|\xi-\eta|^2}\nabla_x e^{-ix\cdot(\xi-\eta)} = e^{-ix\cdot(\xi-\eta)}$$

in order to integrate by parts in (8.5) repeatedly, to precise $d+1$ times. This yields

$$|K_{k,\ell;k',\ell'}(\xi,\eta)| \leqslant C|\ell-\ell'|^{-d-1}$$

Since the support of the kernel with respect to either variable is of size $\leqslant C\,1$, we conclude from Schur's test that

$$\|T^*_{k',\ell'}T_{k,\ell}\|_{2\to 2} \leqslant C\,\langle \ell' - \ell\rangle^{-d-1}$$

where may assume that $|k-k'| \leqslant C\,1$. The same type of argument now implies the second relation of (8.4) and we are done. □

We now turn to *Hardy's inequality*, which provides a natural transition between this chapter and next one devoted to the uncertainty principle.

THEOREM 8.6. *For any $0 \leqslant s < \frac{d}{2}$ there is a constant $C(s,d)$ with the property that*

$$\||x|^{-s}f\|_2 \leqslant C(s,d)\,\|f\|_{\dot H^s(\mathbb{R}^d)} \tag{8.6}$$

for all $f \in \dot H^s(\mathbb{R}^d)$.

PROOF. The stated bound is evident for $s=0$, so we may assume that $0<s<\frac{d}{2}$ by interpolation (any reader uncomfortable with this interpolation may simply assume that condition on s without missing anything from the following argument). By density, it suffices to prove this estimate for $f \in \mathcal{S}(\mathbb{R}^d)$. We use the partition of unity from Lemma 8.3 and write $P_kf = (\psi(2^{-k}\xi)\hat f)^\vee$ for any $k \in \mathbb{Z}$. By construction, $f = \sum_k P_kf$ where the series converges in several ways, for example in L^2. Furthermore, we set $\chi_\ell(x) = \psi(2^{-\ell}x)$ with $\ell \in \mathbb{Z}$. Then

$$\begin{aligned}\||x|^{-s}f\|_2 &\leqslant \sum_k \||x|^{-s}P_kf\|_2 \\ &\leqslant C\sum_k\Big(\sum_{\ell+k\leqslant 0} 2^{-2\ell s}\|\chi_\ell P_kf\|_2^2 + 2^{2sk}\|P_kf\|_2^2\Big)^{\frac12}\end{aligned} \tag{8.7}$$

where the final term is the contribution from the region $\{|x| > 2^{-k}\}$. The idea here is that the main contribution should come from $\ell+k=0$ which immediately yields (8.6). The mechanism here derives from the uncertainty principle which we will discuss in the following chapter; an indication of this is given by the transformation law (6.1) and in particular the special case of this, namely the dilation identity (6.2).

To implement this idea, we let $\tilde\psi$ be smooth and compactly supported in $\mathbb{R}^d\backslash\{0\}$ such that $\tilde\psi = 1$ on $\mathrm{supp}(\psi)$. This implies that $\tilde P_kf := (\tilde\psi(2^{-k}\xi)\hat f)^\vee$ for any $k \in \mathbb{Z}$ satisfies $\tilde P_kP_k = P_k\tilde P_k = P_k$ whence

$$\|\chi_\ell P_kf\|_2 = \|\chi_\ell P_k\tilde P_kf\|_2 \leqslant \|\chi_\ell P_k\|_{2\to 2}\|\tilde P_kf\|_2$$

We now claim that

$$\|\chi_\ell P_k\|_{2\to 2}^2 \leqslant C\,2^{d(\ell+k)} \qquad \ell+k\leqslant 0 \tag{8.8}$$

which is simple consequence of Schur's test. We shall provide the details for this estimate at the end of the proof. If this holds, then we can continue from (8.7) as follows:

$$\begin{aligned}\sum_k \Big(\sum_{\ell+k\leqslant 0} 2^{-2\ell s} \|\chi_\ell P_k f\|_2^2 \Big)^{\frac12} &\leqslant C \sum_k \Big(\sum_{\ell+k\leqslant 0} 2^{-2\ell s} 2^{2d(\ell+k)} \|\tilde P_k f\|_2^2 \Big)^{\frac12} \\ &\leqslant C \sum_k \Big(\sum_{\ell+k\leqslant 0} 2^{(d-2s)(\ell+k)} 2^{2sk} \|\tilde P_k f\|_2^2 \Big)^{\frac12} \\ &\leqslant C \sum_k 2^{sk} \|\tilde P_k f\|_2\end{aligned}$$

In summary we arrive at (it is easy to see that we my drop the tilde on P_k)

$$\| |x|^{-s} f\|_2 \leqslant \sum_k 2^{2sk} \|P_k f\|_2$$

But this is strictly *weaker* than what we would like which shows that one cannot give away as much of the (almost) orthogonality of the $P_k f$ as we did.

Hence, we should go about the first step (8.7) a bit differently:

$$\begin{aligned}\| |x|^{-s} f\|_2^2 &\leqslant C \sum_\ell 2^{-2\ell s} \|\chi_\ell f\|_2^2 \\ &\leqslant C \sum_\ell 2^{-2\ell s} \Big(\sum_{k+\ell\leqslant 0} \|\chi_\ell P_k f\|_2 \Big)^2 + \sum_\ell 2^{-2\ell s} \|\chi_\ell P_{>-\ell} f\|_2^2\end{aligned} \tag{8.9}$$

where $P_{-\ell} := \sum_{j>-\ell} P_j$. For the sum involving both k, ℓ we invoke (8.8) to wit

$$\begin{aligned}\sum_\ell 2^{-2\ell s} \Big(\sum_{k+\ell\leqslant 0} \|\chi_\ell P_k f\|_2 \Big)^2 &\leqslant C \sum_\ell \Big(\sum_{k+\ell\leqslant 0} 2^{(\frac d2 - s)(\ell+k)} 2^{sk} \|\tilde P_k f\|_2 \Big)^2 \\ &\leqslant C \sum_\ell 2^{2ks} \|\tilde P_k f\|_2^2 \leqslant C\, \|f\|_{\dot H^s}^2\end{aligned}$$

since $\frac d2 - s > 0$. To pass to the last line one can use Schur's test for sums, for example. The final sum in (8.9) uses that $s > 0$:

$$\begin{aligned}\sum_\ell 2^{-2\ell s} \|\chi_\ell P_{>-\ell} f\|_2^2 &\leqslant \sum_\ell 2^{-2\ell s} \|P_{>-\ell} f\|_2^2 \\ &\leqslant C \sum_\ell 2^{-2(\ell+k)s} \sum_{k+\ell>0} 2^{2sk} \|P_k f\|_2^2 \\ &\leqslant C \sum_k 2^{2sk} \|P_k f\|_2^2 \sum_{\ell+k>0} 2^{-2(\ell+k)s} \leqslant C\, \|f\|_{\dot H^s}^2\end{aligned}$$

which is (8.6).

It remains to prove (8.8). For this write

$$(\chi_\ell P_k f)(x) = \int_{\mathbb{R}^d} \psi(2^{-\ell} x) 2^{kd} \check\psi(2^k(x-y))\, f(y)\, dy$$

and observe that

$$\sup_x \int_{\mathbb{R}^d} |\psi(2^{-\ell}x)2^{kd}\check{\psi}(2^k(x-y))|\,dy \leqslant C\,1$$
$$\sup_y \int_{\mathbb{R}^d} |\psi(2^{-\ell}x)2^{kd}\check{\psi}(2^k(x-y))|\,dx \leqslant C\,2^{(k+\ell)d}$$

while implies the claim by Schur's test. □

Exercise 8.2. Show that one may "drop the tilde" as was done in the previous proof, i.e., prove that

$$\sum_k 2^{2ks}\|\tilde{P}_k f\|_2^2 \simeq \sum_k 2^{2ks}\|P_k f\|_2^2$$

for any s.

Theorem 8.6 is sharp in the following sense: clearly, both sides scale the same so it is impossible to use any other Sobolev space, for example. The condition $s \geqslant 0$ is of course needed as we cannot allow for growing weights. And $s \geqslant \frac{d}{2}$ is not possible since in that case the estimate fails for Schwartz functions which do not vanish at the origin (the norm on the left-hand side is infinite in that case).

Several questions pose themselves naturally now: (i) is there a substitute for $s > \frac{d}{2}$ in the previous theorem? (ii) Is there a version of Theorem 8.6 with L^p instead of L^2?

In case of $s = 1$, we shall answer these questions in the problems.

Notes

For more on the method of almost orthogonality, in particular the use of Cotlar's lemma, see Stein [**46**]. Hardy's inequalities are basic tools that have been generalized in many different directions, see for example Kufner and Opic [**39**]. For fractional s one needs more tools, namely Littlewood-Paley theory.

Problems

Problem 8.1. Let T be of the form (8.2) where $a \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ is such that

$$\sup_{x\in\mathbb{R}^d} \left[|\partial_x^\alpha a(x,\xi)| + |\partial_\xi^\alpha a(x,\xi)|\right] \leqslant B\langle\xi\rangle^s$$

for all $|\alpha| \leqslant d+1$ and all $\xi \in \mathbb{R}^d$. Here $s \geqslant 0$ is arbitrary but fixed. Show that T is bounded form $H^s(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$.

Problem 8.2. Suppose $K(x,y)$ is defined on $(0,\infty)^2$ and is homogeneous of degree -1, i.e., $K(\lambda x, \lambda y) = \lambda^{-1}K(x,y)$ for all $\lambda > 0$. Further, assume that for an arbitrary but fixed $1 \leqslant p \leqslant \infty$ one has

$$\int_0^\infty |K(x,y)|y^{-\frac{1}{p}}\,dy = A$$

Show that $(Tf)(x) := \int_0^\infty K(x,y)\,f(y)\,dy$ satisfies $\|Tf\|_p \leqslant A\|f\|_p$.

PROBLEM 8.3. Prove Hardy's original inequality, valid for every measurable $f \geqslant 0$

$$\int_0^\infty \Big(\int_0^x f(t)\,dt\Big)^p x^{-r-1}\,dx \leqslant \Big(\frac{p}{r}\Big)^p \int_0^\infty (xf(x))^p\, x^{-r-1}\,dx$$
$$\int_0^\infty \Big(\int_x^\infty f(t)\,dt\Big)^p x^{r-1}\,dx \leqslant \Big(\frac{p}{r}\Big)^p \int_0^\infty (xf(x))^p\, x^{r-1}\,dx$$

for every $1 \geqslant p < \infty$, $r > 0$. Equality holds if and only if $f = 0$ a.e. *Hint:* use the previous problem.

PROBLEM 8.4. Show the following Hardy's inequality in $\mathbb{R}^d$

$$\int_{\mathbb{R}^d} |\nabla u(x)|^p\,dx \geqslant \Big(\frac{|d-p|}{p}\Big)^p \int_{\mathbb{R}^d} \frac{|u(x)|^p}{|x|^p}\,dx \tag{8.10}$$

for all $u \in \mathcal{S}(\mathbb{R}^d)$ if $1 \leqslant p < d$ and for all such u which also vanish on a neighborhood of the origin if $\infty > p > d$. The constants here are sharp. *Hint:* use the previous problem.

PROBLEM 8.5. Show that if $1 \leqslant p < d$ and $\frac{1}{p} - \frac{1}{p*} = \frac{1}{d}$, then

$$\int_{\mathbb{R}^d} \frac{|u(x)|^p}{|x|^p}\,dx = C(p,d)\|u\|_{p*}^p$$

provided $u \in \mathcal{S}(\mathbb{R}^d)$ is radial and decreasing. Conclude that for such functions the Sobolev embedding inequality $\|u\|_{p*} \leqslant C(p,d)\|\nabla u\|_p$. By means of the device of *nonincreasing rearrangement* this special case implies the general one. In other words, the Sobolev embedding holds without the radial decreasing assumption. The point here is that the rearrangement does not change $\|u\|_{p*}$ and at most decreases $\|\nabla u\|_p$, see [**36**]. This is useful in order to find the optimal constants, see Frank, Seiringer [**18**].

CHAPTER 9

The Uncertainty Principle

This chapter deal with various manifestations of the heuristic principle that *it is not possible for both f and $\hat{f}$ the be localized on small sets*. We already encountered a rigorous, albeit qualitative and rather weak version of this principle: it is impossible for both $f \in L^2(\mathbb{R}^d)$ and $\hat{f}$ to be compactly supported.

Here we are more interested in quantitative versions, and moreover, the tightness of these quantitative versions. What is meant by that can be seen from the transformation law (6.1): let χ be a smooth bump function, and A be an invertible affine transformation that takes the unit ball onto an ellipsoid

$$\mathcal{E} := \Big\{x \in \mathbb{R}^d \mid \sum_{j=1}^{d} (x_j - y_j)^2 r_j^{-2} \leqslant 1\Big\}$$

where y_j are arbitrary constants and $r_j > 0$. Then $\chi_A := \chi \circ A^{-1}$ can be thought of as a smoothed out indicator function of $\mathcal{E}$. By (6.1), $\widehat{\chi_A}$ then essentially "lives" on the dual ellipsoid

$$\mathcal{E}^* := \Big\{\xi \in \mathbb{R}^d \mid \sum_{j=1}^{d} \xi_j^2 r_j^2 \leqslant 1\Big\}$$

Moreover, while χ_A is L^∞-normalized, $\widehat{\chi_A}$ is L^1-normalized. To "essentially live on" $\mathcal{E}^*$ means that

$$|\widehat{\chi_A}(\xi)| \leqslant C_N\, |\mathcal{E}| \big(1 + |\xi|_{\mathcal{E}^*}^2\big)^{-N}$$

where

$$|\xi|_{\mathcal{E}^*}^2 := \sum_{j=1}^{d} \xi_j^2 r_j^2 \qquad \forall\, \xi$$

is the Euclidean norm relative to $\mathcal{E}^*$. What we can see from this is another, still heuristic but more quantitative form of the uncertainty principle, namely the following: *if $f \in L^2(\mathbb{R}^d)$ is supported in a ball of size R, then it is not possible for $\hat{f}$ to be "concentrated" on a scale much less than R^{-1}.*

We begin the rigorous discussion with another form of Bernstein's estimate. In what follows $B(x_0, r)$ denotes the Euclidean ball in $\mathbb{R}^d$ which is of radius r and centered at x_0.

Lemma 9.1. *Suppose $f \in L^2(\mathbb{R}^d)$ satisfies* $\mathrm{supp}(f) \subset B(0, r)$. *Then*

$$\|\partial^\alpha \hat{f}\|_2 \leqslant (2\pi r)^{|\alpha|} \|\hat{f}\|_2 \tag{9.1}$$

for all multi-indices α.

PROOF. This follows from $\hat{f} \in C^\infty(\mathbb{R}^d)$, the relation

$$\partial^\alpha \hat{f}(\xi) = (-2\pi i)^{|\alpha|} \widehat{x^\alpha f}(\xi)$$

and Plancherel's theorem. □

Heurstically speaking, this can be seen as a version of the non-concentration statement from before. Indeed, if $\hat{f}$ were to be concentrated on a scale of size $\ll r^{-1}$, then we may expect the left-hand side of (9.1) to be much larger than $r^{-|\alpha|}$. Later in this chapter, we shall see a better reason why Bernstein's estimate should be viewed in the context of the uncertainty principle.

Next, we prove Heisenberg's uncertainty principle.

PROPOSITION 9.2. *For any $f \in \mathcal{S}(\mathbb{R})$ one has*

$$\|f\|_2^2 \leqslant 4\pi \|(x - x_0)f\|_2 \|(\xi - \xi_0)\hat{f}\|_2 \tag{9.2}$$

for all $x_0, \xi_0 \in \mathbb{R}$. This inequality is sharp, with extremizers being precisely given by the modulated Gaussians

$$f(x) = C\, e^{2\pi i \xi_0 x}\, e^{-\pi\delta(x-x_0)^2}$$

where C and $\delta > 0$ are arbitrary.

PROOF. Without loss of generality $x_0 = \xi_0 = 0$. Define $D = \frac{1}{2\pi i}\frac{d}{dx}$, and let $(Xf)(x) = xf(x)$. Then the commutator

$$[D, X] = DX - XD = \frac{1}{2\pi i}$$

whence for any $f \in \mathcal{S}(\mathbb{R})$,

$$\begin{aligned} \|f\|_2^2 &= 2\pi i\langle [D, X]f, f\rangle = 2\pi i(\langle Xf, Df\rangle - \langle Df, Xf\rangle) \\ &= 4\pi \mathrm{Im}\,\langle Df, Xf\rangle \leqslant 4\pi\|Df\|_2\|Xf\|_2 \end{aligned} \tag{9.3}$$

as claimed.

For the sharpness, note that if equality is achieved in (9.3) then from the condition on equality in Cauchy-Schwarz one has $Df = \lambda Xf$ for some $\lambda \in \mathbb{C}$. Then $\langle Df, Xf\rangle = \lambda\|Xf\|_2^2$ which means that $\lambda = i\delta$ with some $\delta > 0$. Finally, this implies that

$$f'(x) = -2\pi\delta f(x), \qquad f(x) = Ce^{-\pi\delta x^2}$$

The general case follows by translation in x and ξ (the latter being the same as multiplication by $e^{ix\xi_0}$. □

There is an analogous statement in higher dimension which we leave to the reader. Once again, we can see Proposition 9.2 as a manifestation of the principle that if f is concentrated on scale $R > 0$, then $\hat{f}$ cannot be concentrated on a scale $\ll R^{-1}$.

Hardy's inequality from the previous chapter can be seen as another instance of this principle. Indeed, in $\mathbb{R}^3$ it implies that

$$\| |x - x_0|^{-1} f \|_2 \leqslant C \|\nabla f\|_2 = C \|\xi \hat{f}\|_2 \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^3) \tag{9.4}$$

for any $x_0 \in \mathbb{R}^3$. Thus, if $\hat{f}$ is supported on $B(0,R)$, then

$$\sup_{x_0 \in \mathbb{R}^3} \|f\|_{L^2(B(x_0,\rho))} \leqslant C\rho R \|f\|_2 \qquad 0 < \rho < R^{-1}$$

We leave it to the reader to derive similar statements in all dimensions. In fact, the proof of Hardy's inequality which we gave above clearly shows the dominant role of the dual scales R in x, and R^{-1} in ξ, respectively.

Let us reformulate (9.4) with $x_0 = 0$ as follows:

$$\langle |x|^{-2} f, f \rangle \leqslant C \langle -\Delta f, f \rangle \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^3)$$

which can be rewritten as $-\Delta \geqslant \frac{c}{|x|^2}$ for some $c > 0$ as operators (in other words, $-\Delta - \frac{c}{|x|^2}$ is a nonnegative symmetric operator).

It is interesting to note that we can now give a proof of this special case of Hardy's inequality using commutator arguments, similar to those in the proof of Proposition 9.2. In addition, this will yield the optimal value of the constant.

Proposition 9.3. *One has $-\Delta \geqslant \frac{(d-2)^2}{4|x|^2}$ in $\mathbb{R}^d$ with $d \geqslant 3$, which means that for any $f \in \mathcal{S}(\mathbb{R}^d)$*

$$\langle -\Delta f, f \rangle \geqslant \frac{(d-2)^2}{4} \langle |x|^{-2} f, f \rangle \Longleftrightarrow \| |x|^{-1} f \|_2 \leqslant \frac{(d-2)^2}{4} \|\nabla f\|_2$$

Proof. Let $p = -i\nabla$. The relevant commutator here is, with $r = |x|$,

$$i[r^{-1} p r^{-1}, X] = d r^{-2}$$

This implies that

$$d \|r^{-1} f\|_2^2 = -2 \mathrm{Im}\, \langle r^{-1} p r^{-1} f, X f \rangle \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^d)$$

Since

$$p |x|^{-1} = |x|^{-1} p + i \frac{x}{|x|^3}$$

one obtains

$$(d-2) \|r^{-1} f\|_2^2 = -2 \mathrm{Im}\, \langle p f, X r^{-2} f \rangle \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^d)$$

Cauchy-Schwartz concludes the proof. □

The bound in the previous proposition is optimal, but not attained in the largest admissible class, i.e., $\dot{H}^1(\mathbb{R}^d)$. We shall not dwell on this issue here.

Next, we investigate the following question: *if $E, F \subset \mathbb{R}^d$ are of finite measure, can there be a nonzero $f \in L^2(\mathbb{R}^d)$ with* $\mathrm{supp}(f) \subset E$ *and* $\mathrm{supp}(f) \subset F$ *?*

Note that if there were such an f, then $Tf := \chi_E (\chi_F \hat{f})^\vee$ satisfies $Tf = f$. In particular, this would imply $\|T\|_{2\to 2} \geqslant 1$. We might expect that T has small norm if E and F are small, leading to a contradiction and negative answer to our question. Now note that

$$(Tf)(x) = \int_{\mathbb{R}^d} \chi_E(x)\check{\chi}_F(x-y)f(y)\,dy$$

In other words, T has kernel $K(x,y) = \chi_E(x)\check{\chi}_F(x-y)$ with Hilbert-Schmid norm

$$\int_{\mathbb{R}^{2d}} |K(x,y)|^2\,dxdy = |E|\cdot|F| =: \sigma^2 < \infty$$

We remark that σ is a scaling invariant quantity. Hence, $\|T\| \leqslant \min(\sigma,1)$ in operator norm and T is a compact operator. So if $\sigma < 1$, then we see that there can be no $f \neq 0$ as in the question. Interestingly, the answer is "no" in all cases and one has the following quantitative expression of this property.

Theorem 9.4. *Let E and F be sets of finite measure in $\mathbb{R}^d$. Then*

$$\|f\|_{L^2(\mathbb{R}^d)} \leqslant C(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)}) \tag{9.5}$$

for some constant $C(A,B,d)$.

The reader will easily prove the estimate in case $\sigma < 1$. For the general case, we formulate the following equivalencies with the goal of showing that $\|T\| < 1$ for any E,F as in the theorem. In the lemma supp refers to the essential support.

Lemma 9.5. *Let E,F be measurable subsets of $\mathbb{R}^d$. Then the following are equivalent (C_1 and C_2 denote constants):*

a) $\|f\|_{L^2(\mathbb{R}^d)} \leqslant C_1(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)})$ *for all* $f \in L^2(\mathbb{R}^d)$
b) there exists $\varepsilon > 0$ *so that* $\|f\|^2_{L^2(E)} + \|\hat{f}\|^2_{L^2(F)} \leqslant (2-\varepsilon)\|f\|_2^2$ *for all* $f \in L^2(\mathbb{R}^d)$
c) if $\operatorname{supp}(\hat{f}) \subset F$*, then* $\|f\|_2 \leqslant C_2\|f\|_{L^2(E^c)}$
d) if $\operatorname{supp}(f) \subset E$*, then* $\|\hat{f}\|_2 \leqslant C_2\|\hat{f}\|_{L^2(F^c)}$
e) there exists $0 < \rho < 1$ *so that* $\|\chi_E(\chi_F\hat{f})^\vee\|_2 \leqslant \rho\|f\|_2$ *for all* $f \in L^2(\mathbb{R}^d)$.

Proof. $a) \Longrightarrow b)$: one has

$$\begin{aligned}\|f\|^2_{L^2(E)} + \|\hat{f}\|^2_{L^2(F)} &= 2\|f\|_2^2 - \|f\|^2_{L^2(E^c)} - \|\hat{f}\|^2_{L^2(F^c)}\\ &\leqslant (2-(2C_1^2)^{-1})\|f\|_2^2\end{aligned}$$

$b) \Longrightarrow c)$: for f as in $c)$, $\|f\|^2_{L^2(E)} \leqslant (1-\varepsilon)\|f\|_2^2$ or $\|f\|_2^2 \leqslant \varepsilon^{-1}\|f\|^2_{L^2(E)}$. $c) \Longrightarrow a)$: write $f = P_Ff + P_{F^c}f$ where $P_Ff = (\chi_F\hat{f})^\vee$ and set $Q_{E^c}f = \chi_{E^c}f$. Then

$$\begin{aligned}\|f\|_2 &\leqslant \|P_Ff\|_2 + \|P_{F^c}f\|_2 \leqslant C_2\|Q_{E^c}P_Ff\|_2 + \|P_{F^c}f\|_2\\ &\leqslant C_2\|Q_{E^c}f\|_2 + C_2\|Q_{E^c}P_{F^c}f\|_2 + \|P_{F^c}f\|_2\\ &\leqslant C_2\|Q_{E^c}f\|_2 + (C_2+1)\|P_{F^c}f\|_2\end{aligned}$$

Interchanging f with $\hat{f}$ one sees that property $d)$ to the first three.

$c) \Longrightarrow e)$: with the projection P, Q as above one has

$$\begin{aligned}\|P_F f\|_2^2 &= \|Q_E P_F f\|_2^2 + \|Q_{E^c} P_F f\|_2^2 \\ &\geqslant \|Q_E P_F f\|_2^2 + C_2^{-2} \|P_F\|_2^2\end{aligned}$$

which implies the desired bound with $\rho = (1 - C_2^{-2})^{\frac{1}{2}}$.
$e) \Longrightarrow c)$: clear. □

Proof of Theorem 9.4. We already observed that $T = Q_E P_F$ is compact with Hilbert-Schmid norm $\|T\|_{\mathrm{HS}} \leqslant \sigma$. In particular,

$$\dim\{f \in L^2(\mathbb{R}^d) \mid Tf = \lambda f\} \leqslant \lambda^{-2}\sigma^2 \tag{9.6}$$

To prove this bound, let $\{f_j\}_{j=1}^m$ be an orthonormal sequence in the eigenspace on the left. Then, with K being the kernel of K, one has

$$m\lambda^2 = \sum_j \Big| \int_{\mathbb{R}^{2d}} K(x,y) f_j(x) \overline{f_j}(y)\, dxdy \Big|^2 \leqslant \|K\|_{L^2(\mathbb{R}^{2d}}^2 \tag{9.7}$$

by Bessel's inequality. Furthermore, as product of projections, $\|T\| \leqslant 1$ in operator norm. If $\|T\| = 1$, then it follows from compactness of T that there exists $f \in L^2(\mathbb{R}^d)$ with $f \neq 0$ and $\|Tf\|_2 = \|f\|_2$.

Thus, there exists f with $\mathrm{supp}(f) \subset E$ and $\mathrm{supp}(\hat{f}) \subset F$ (as essential supports). We shall now obtain a contradiction to (9.6) for $\lambda = 1$. Inductively, define

$$\begin{aligned} S_0 &:= \mathrm{supp}(f), \quad S_1 := S_0 \cup (S_0 - x_0) \\ S_{k+1} &:= S_k \cup (S_0 - x_k) \quad k \geqslant 1 \end{aligned}$$

where the translations x_k are chosen such that

$$|S_k| < |S_{k+1}| < |S_k| + 2^{-k}$$

If follows that the collection $f_k := f(\cdot + x_k)$ with $k \geqslant 0$ is linearly independent with

$$\mathrm{supp}(f_k) \subset S_\infty := \bigcup_{\ell=0}^{\infty} S_\ell \qquad \forall\, k \geqslant 0$$

where $|S_\infty| < \infty$, while maintaining $\mathrm{supp}(\hat{f}_k) \subset F$. This gives the desired contradiction. □

Next, we formulate some results that provide further evidence of the non-concentration property of functions with Fourier support in $B(0,1)$.

Proposition 9.6. *Let $\alpha > 0$ and suppose that $E \subset \mathbb{R}^d$ satisfies*

$$|E \cap B| < \alpha |B| \ \textit{ for all balls } B \textit{ or radius } 1$$

Suppose $f \in L^2(\mathbb{R}^d)$ satisfies $\mathrm{supp}(\hat{f}) \subset B(0,1)$. *Then*

$$\|f\|_{L^2(E)} \leqslant \delta(\alpha) \|f\|_2 \tag{9.8}$$

where $\delta(\alpha) \to 0$ as $\alpha \to 0$.

PROOF. Let ψ be a smooth compactly supported function with $\psi = 1$ on $B(0,1)$. Define $Tg := \chi_E(\psi\hat{g})^\vee$. The kernel of this operator is

$$K(x,y) = \chi_E(x)\check{\psi}(x-y)$$

and satisfies the bound, for any $N \geqslant 1$,

$$|K(x,y)| \leqslant C_N \chi_E(x)(1+|x-y|)^{-N}$$

We apply Schur's lemma to this kernel. To this end let $\varepsilon > 0$ be arbitrary but fixed. First, $\|K(x,\cdot)\|_{L^1_y} \leqslant C\,1$. Second, take $N > d$ and choose R so large that

$$\sup_y \int_{|x-y|>R} C_N\,(1+|x-y|)^{-N}\,dx \leqslant \varepsilon$$

Then we see by covering $B(x,R)$ by finitely many balls of radius 1 that

$$\int_{|x-y|<R} C_N \chi_E(x)(1+|x-y|)^{-N}\,dx \leqslant \varepsilon$$

provided $\alpha > 0$ is small enough. Therefore, by Schur's lemma 8.2 we conclude that $\|T\| \leqslant C\sqrt{\varepsilon}$. Hence, if f is as in the hypothesis, then $\chi_E f = Tf$ whence

$$\|f\|_{L^2(E)} \leqslant C\sqrt{\varepsilon}\|f\|_2$$

as claimed. □

We now establish a somewhat more refined version of the previous result, which goes by the name of Logvinenko-Sereda theorem.

THEOREM 9.7. *Suppose a measurable set $E \subset \mathbb{R}^d$ satisfies the following "thickness" condition: there exists $\gamma \in (0,1)$ such that*

$$|E \cap B| > \gamma|B| \text{ for all balls } B \text{ of radius } R^{-1}$$

where $R > 0$ is arbitrary but fixed. Assume that $\mathrm{supp}(\hat{f}) \subset B(0,R)$. *Then*

$$\|f\|_{L^2(\mathbb{R}^d)} \leqslant C\,\|f\|_{L^2(E)} \tag{9.9}$$

where the constant C only depends on d and γ.

By the same argument as in Proposition 9.6 (or directly from that proposition) one proves this result for γ near 1.

EXERCISE 9.1. Prove Theorem 9.7 if γ is close to 1.

We shall base the proof of Theorem 9.7 on arguments involving analytic functions of several complex variables. This is natural in view of the Paley-Wiener theorem, cf. 6.5 which we shall state below. However, we shall not obtain explicit constants $C(d,\gamma)$ from our argument.

In order to keep this discussion self-contained, we briefly review the few facts about analytic functions of several complex variables that will be needed. Let $\Omega \subset \mathbb{C}^d$ be open and connected (a "domain"), and suppose $G \in C(\Omega)$ is complex-valued. We say that G is analytic in Ω if and only if

$$z_j \mapsto G(z_1, z_2, \ldots, z_{j-1}, z_j, z_{j+1}, \ldots, z_d)$$

is analytic on the corresponding sections of Ω as all other z_j are kept fixed. We denote this class by $\mathcal{A}(\Omega)$. The Cauchy integral formula applies via integration along contours in the individual variables. Examples of such functions are power series

$$f(z) = \sum_{\alpha} a_\alpha \, (z - z_0)^\alpha, \qquad z \in \Delta := \{z \mid \max_{1 \leqslant j \leqslant d} |z_j - z_j^0| < M\}$$

provided $|a_\alpha| \leqslant M^{-|\alpha|}$ for all multi-indices α and a constant $M > 0$. Indeed, it is easy to see that the power series converges uniformly in Δ and the sum $f(z)$ is analytic in Δ. In particular, if

$$|a_\alpha| \leqslant M \frac{B^{|\alpha|}}{\alpha!} \qquad \forall \, \alpha$$

then the power series represents an entire function in $\mathbb{C}^d$.

Conversely, if $f \in \mathcal{A}(\Omega)$, then

$$f(z) = \sum_{\alpha} \frac{\partial_z^\alpha f(z_0)}{\alpha!} \, (z - z_0)^\alpha \qquad |z - z_0| < r$$

for some $r > 0$. The *uniqueness theorem* therefore also holds in this setting: if $\partial_z^\alpha f(z_0) = 0$ for all α, then $f \equiv 0$ in Ω. Moreover, if $\Omega \cap \mathbb{R}^d$ contains a set of positive measure on which f vanishes, then $f \equiv 0$ in Ω. Due to the Cauchy estimates, one has Montel's theorem on *normal families:* if $\mathcal{F} \subset \mathcal{A}(\Omega)$ is a locally bounded family, then $\mathcal{F}$ is precompact, i.e., every sequence in $\mathcal{F}$ has a subsequence which converges locally uniformly on Ω to an element of $\mathcal{A}(\Omega)$.

Finally, we recall the Paley-Wiener theorem: suppose $f \in L^2(\mathbb{R}^d)$ satisfies $\mathrm{supp}(\hat{f}) \subset B(0, R)$. Then f extends to an entire function in $\mathbb{C}^d$ with exponential growth:

$$|f(z)| \leqslant C \, R^{\frac{d}{2}} \, e^{2\pi |R| |z|} \|f\|_{L^2(\mathbb{R}^d)}$$

which is proved by placing absolute values inside the integral defining $\hat{f}(z)$ and applying Cauchy-Schwarz.

Proof of Theorem 9.7. We may assume that $R = 1$. Let $A \geqslant 2$ be a constant that we shall fix later. Partition $\mathbb{R}^d$ into congruent cubes of sidelength 2, which we shall denote by Q. We say that one such Q is *good* if

$$\|\partial^\alpha f\|_{L^2(Q)} \leqslant A^{|\alpha|} \|f\|_{L^2(Q)} \qquad \forall \, \alpha$$

We now claim that

$$\|f\|_{L^2(\bigcup_{\mathrm{bad}} Q)} \leqslant C A^{-1} \|f\|_2 \tag{9.10}$$

To prove it, consider the sum

$$\sum_{\alpha \neq 0} \sum_{Q \text{ bad}} A^{-2|\alpha|} \|D^\alpha f\|_{L^2(Q)}^2 \tag{9.11}$$

Since Q is bad, there is a lower bound of $\|f\|^2_{L^2(\bigcup_{\text{bad}} Q)}$. On the other hand, by Lemma 9.1, one has an upper bound of the form

$$\begin{aligned}&\sum_{\alpha\neq 0}\sum_{Q\text{ bad}} A^{-2|\alpha|}(2\pi)^{|\alpha|}\|f\|_2^2\\ &\leqslant \sum_{k=1}^{\infty}(k+1)^d(2\pi A^{-2})^k\|f\|_2^2 \leqslant CA^{-2}\|f\|_2^2\end{aligned}$$

which proves the claim.

Next, we claim that if Q is good, then there exists $x_0 \in Q$ such that

$$|\partial^\alpha f(x_0)| \leqslant C_\alpha A^{2|\alpha|+1}\|f\|_{L^2(Q)}, \qquad \forall\ \alpha \tag{9.12}$$

Indeed, if not, then

$$\begin{aligned}A^2|Q|\|f\|^2_{L^2(Q)} &\leqslant \int_Q \sum_\alpha A^{-3|\alpha|}A^{4|\alpha|+2}\|f\|^2_{L^2(Q)}\,dx\\ &\leqslant \int_Q \sum_\alpha A^{-3|\alpha|}|\partial^\alpha f(x)|^2\,dx\\ &\leqslant \sum_\alpha A^{-3|\alpha|}A^{2|\alpha|}\|f\|^2_{L^2(Q)} \leqslant \sum_{k=0}^{\infty}(k+1)^d 2^{-k}\|f\|^2_{L^2(Q)}\end{aligned}$$

which is a contradiction for $A \geqslant 2$ large.

And finally, we claim that there exists $\eta > 0$ which depends only on d, γ, A so that

$$\|f\|_{L^2(E\cap Q)} \geqslant \eta\|f\|_{L^2(Q)} \text{ for all good } Q \tag{9.13}$$

If not, there exist $\{f_n\}_{n=1}^\infty$ in $L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat{f}_n) \subset B(0,1)$ and measurable sets $E_n \subset \mathbb{R}^d$ with

$$|E_n \cap B(x,1)| \geqslant \gamma|B(x,1)| \qquad \forall\ x \in \mathbb{R}^d$$

and good cubes Q_n such that

$$\|f_n\|_{L^2(Q_n)} = 1, \qquad \|f_n\|_{L^2(E_n\cap Q_n)} \to 0$$

as $n \to \infty$. After translation, $Q_n = [-1,1]^d =: Q_0$. Denote the entire extension of f_n given by the Paley-Wiener theorem by F_n. We expand F_n around a good x_n as in (9.12) to conclude that for any $|z| \leqslant R$ one has

$$\begin{aligned}|F_n(z)| &\leqslant \sum_\alpha \frac{A^{2|\alpha|+1}}{\alpha!}|(z-x_n)^\alpha|\\ &\leqslant \Big(\sum_{\ell=0}^{\infty}\frac{A^{2\ell+1}}{\ell!}(2+|z|)^\ell\Big)^d \leqslant C(A,d,R)\end{aligned}$$

So $\{F_n\}_{n=1}^\infty$ is a normal family in $\mathbb{C}^d$ and we may extract a locally uniform limit $F_n \to F_\infty$ (up to passing to a subsequence) whence $\|F_\infty\|_{L^2(Q_0)} = 1$. On the other hand,

$$\begin{aligned}
&\Big|\big\{x \in Q_0 \cap E_n \mid |f_n(x)| \geqslant \lambda_n\big\}\Big| \\
&\leqslant \lambda_n^{-2}\|f_n\|^2_{L^2(E_n)} = \lambda_n^{-2}\eta_n^2 \\
&\leqslant \frac{\gamma}{2}|B(0,1)|
\end{aligned}$$

if $\eta_n = \lambda_n\sqrt{\frac{\gamma}{2}|B(0,1)|}$. In particular, $\lambda_n \to 0$. By the thickness condition,

$$X_n := \{x \in Q_0 \cap E_n \mid |f_n(x)| \leqslant \lambda_n\}$$

satisfies $|X_n| > \frac{\gamma}{2}$. But then

$$X_\infty := \limsup_{n\to\infty} X_n$$

satisfies $|X_\infty| \geqslant \frac{\gamma}{2}$. By construction, $\lambda_n \to 0$ whence $F_\infty = 0$ on X_∞. By the uniqueness theorem for analytic functions of several variables, $F_\infty \equiv 0$ on Q_0.

This contradiction establishes our final claim. We can now finish the proof of the theorem. Indeed, one has from (9.13) and (9.10)

$$\begin{aligned}
\|f\|^2_{L^2(E\cap\bigcup_{Q\text{ good}} Q)} &= \sum_{Q\text{ good}} \|f\|^2_{L^2(E\cap Q)} \\
&\geqslant \sum_{Q\text{ good}} \eta^2\|f\|^2_{L^2(Q)} = \eta^2\|f\|^2_{L^2(\bigcup_{\text{good}} Q)} \\
&\geqslant \eta^2\big(\|f\|^2_{L^2} - C^2A^{-2}\|f\|_2^2\big) \\
&\geqslant \frac{1}{2}\eta^2\|f\|_2^2
\end{aligned}$$

if A was chosen large enough. This implies the desired estimate. □

Exercise 9.2. Prove that Theorem 9.7 with $R = 1$ can also be proved under the following weaker hypothesis: there exists $\gamma > 0$ and $r > 0$ such that

$$|E \cap B(x,r)| > \gamma \qquad \forall\ x \in \mathbb{R}^d$$

The constant in (9.9) then of course also depends on r.

We now formulate a second version of the Logvinenko-Sereda theorem. We begin with a definition.

Definition 9.8. A set $F \subset \mathbb{R}^d$ is called $B(0,1)$-negligible if and only if there exists $\varepsilon > 0$ so that for all $f \in L^2(\mathbb{R}^d)$ one has either

$$\begin{aligned}
\|\hat{f}\|_{L^2(\mathbb{R}^d\setminus B(0,1))} &\geqslant \varepsilon\|f\|_2 \ \text{ or} \\
\|f\|_{L^2(\mathbb{R}^d\setminus F)} &\geqslant \varepsilon\|f\|_2
\end{aligned} \tag{9.14}$$

By means of Theorem 9.7 we can characterize negligible sets.

THEOREM 9.9. *A set F is $B(0,1)$-negligible if and only if there exist $r > 0$ and $\beta \in (0,1)$ such that for all $x \in \mathbb{R}^d$*

$$|F \cap B(x,r)| < \beta |B(x,r)| \tag{9.15}$$

PROOF. First, we note that the condition (9.15) is equivalent to the following: there exist $r > 0$ and $\gamma > 0$ so that

$$|F^c \cap B(x,r)| > \gamma \qquad \forall\ x \in \mathbb{R}^d \tag{9.16}$$

If (9.16) fails, then there exists $x_j \in R^d$ and $r_j \to \infty$ with

$$|F^c \cap B(x_j, r_j)| \to 0 \qquad j \to \infty$$

Now take φ Schwartz with $\operatorname{supp}(\hat{\varphi}) \subset B(0,1)$ and set $\widehat{\varphi_j}(\xi) := e^{2\pi i x_j \cdot \xi}\hat{\varphi}$. Then $\widehat{\varphi_j}$ is still supported in $B(0,1)$ and with $B_j := B(x_j, r_j)$, we have

$$\begin{aligned}
\|\varphi_j\|_{L^2(\mathbb{R}^d \setminus F)} &\leqslant \|\varphi_j\|_{L^2(B_j \setminus F)} + \|\varphi_j\|_{L^2(\mathbb{R}^d \setminus B_j)} \\
&\leqslant \|\varphi\|_\infty \sqrt{|B_j \setminus F|} + \Big(\int_{[|x-x_j|>r_j]} C|x-x_j|^{-d-1}\,dx \Big)^{\frac{1}{2}}
\end{aligned} \tag{9.17}$$

which vanishes in the limit $j \to \infty$. But this contradicts (9.14) which finishes the necessity.

For the sufficiency, let $E := F^c$. Applying Theorem 9.7, see also Exercise 9.2, we infer that for any $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat{f}) \subset B(0,1)$ the estimate

$$\|f\|_2 \leqslant C\, \|f\|_{L^2(E)}$$

with $C = C(\gamma, r, d)$. Now split $f = f_1 + f_2$ with $f_1 = (\chi_{B(0,1)}\hat{f})^\vee$. If $\|f_2\|_2 \leqslant \varepsilon \|f\|_2$, then

$$\begin{aligned}
\|f\|_{L^2(E)} &\geqslant \|f_1\|_{L^2(E)} - \|f_2\|_2 \\
&\geqslant C^{-1}\|f\|_2 - \varepsilon\|f\|_2 > \varepsilon\|f\|_2
\end{aligned}$$

provided $\varepsilon > 0$ is sufficiently small and we are done. □

We shall now apply these uncertainty principle ideas to the local solvability of partial differential equations with constant coefficients. To be specific, suppose we are given a polynomial

$$p(\xi) = \sum_\alpha a_\alpha\, \xi^\alpha$$

Then with $D := \frac{1}{2\pi i}\partial_x$ one has

$$p(D) := \sum_\alpha a_\alpha\, (2\pi i)^{-\alpha} D^\alpha$$

This definition is chosen so that for all $f \in \mathcal{S}(\mathbb{R}^d)$

$$\widehat{p(D)f}(\xi) = p(\xi)\hat{f}(\xi) \qquad \forall\ \xi \in \mathbb{R}^d$$

The following theorem is known as Malgrange-Ehrenpreis theorem.

THEOREM 9.10. *Let Ω be a bounded domain in $\mathbb{R}^d$ and $p \neq 0$ be a polynomial. Then for all $g \in L^2(\Omega)$ there exists $f \in L^2(\Omega)$ such that $p(D)f = g$ in the sense of distributions.*

The "distributional sense" means that for all $\varphi \in C^\infty(\Omega)$ with compact support $\mathrm{supp}(\varphi) \subset \Omega$ one has

$$(g, \varphi) = (f, \overline{p}(D)\varphi)$$

where $\overline{p}(\xi) := \sum_\alpha \overline{a_\alpha}\, \xi^\alpha$. The pairing $(\cdot, \cdot)$ is the L^2 pairing.

We emphasize that we are *not* making any ellipticity assumptions in the theorem. Even so, one can show that f can be taken to be $C^\infty(\Omega)$, see the Problem 9.6 below. This is very different from the assertion that *every solution* f is C^∞ which is known as *hypoellipticity*. This is not only a large area of research onto itself, but also false in the generality of the Malgrange-Ehrenpreis theorem.

The proof follows the common principle in functional analysis that *uniqueness (for the adjoint operator) implies existence (for the operator).* Recall that this is the essence of Fredholm theory which establishes this fact for compact perturbations of the identity operator. Of course, we cannot expect uniqueness in our case without any boundary conditions, so we introduce a strong vanishing condition at the boundary. Henceforth we take Ω to be a ball, which we may do since Ω bounded. By dilation and translation, the ball may furthermore be taken to be $B(0, 1)$.

PROPOSITION 9.11. *One has the bound*

$$\|\overline{p}(D)\varphi\|_2 \geqslant C^{-1}\|\varphi\|_2$$

for all $\varphi \in C^\infty(B(0, 1))$ which are compactly supported in $B(0, 1)$. The constant C depends on p and the dimension.

PROOF OF THEOREM 9.10. We show how to deduce the theorem from the proposition. Let

$$X := \left\{\overline{p}(D)\varphi \mid \varphi \in C^\infty_{\mathrm{comp}}(B(0, 1))\right\}$$

viewed as a linear subspace of $L^2(B(0, 1))$. Define a linear functional ℓ on X by

$$\ell(\overline{p}(D)\varphi) = (g, \varphi)$$

where the right-hand side is the L^2-pairing. Proposition 9.11 implies that this is well-defined and in fact one has a bound

$$|\ell(\overline{p}(D)\varphi)| \leqslant \|g\|_2\|\varphi\|_2 \leqslant C\|g\|_2\|\overline{p}(D)\varphi\|_2$$

By the Hahn-Banach theorem we may extend ℓ as a bounded linear functional to $L^2(B(0, 1))$ with the same norm. So by the Riesz representation of such functionals there exists $f \in L^2(B(0, 1))$ with the property that

$$\ell(\overline{p}(D)\varphi) = (f, \overline{p}(D)\varphi) = (g, \varphi)$$

for all φ as above which proves the theorem. □

The proof of Proposition 9.11 will be based on Theorem 9.7, which requires some preparatory lemmas on polynomials.

LEMMA 9.12. *Let p be a nonzero polynomial in $\mathbb{R}^d$. There exists $\beta \in (0,1)$ which only depends on the degree of p and the dimension so that*

$$|\{B \mid |p(x)| \leqslant \frac{1}{2} \max_B |p|\}| < \beta |B|$$

for all balls B of radius 1.

PROOF. Denote the degree of p by N. The proof exploits the fact that any two norms on a finite-dimensional space are comparable (in our case the space of polynomials of degree at most N). Fix a unit ball B. Then there exists a constant $C(N)$ such that

$$\max_{x \in B} (|p(x)| + |\nabla p(x)|) \leqslant C(N) \max_{x \in B} |p(x)|$$

whence

$$\max_{x \in B} |\nabla p(x)| \leqslant C(N) \max_{x \in B} |p(x)|$$

By the mean value theorem,

$$|p(x) - p(y)| \leqslant \max_B |\nabla p|\, |x - y| \leqslant C(N) |x - y| \max_B |p|$$

from which we infer that

$$|p(y)| \geqslant |p(x_{\max})| - C(N) \max_B |p|\, |x_{\max} - y| \geqslant \frac{1}{2} |p(x_{\max})|$$

for any $y \in B(x_{\max}, (2C(N))^{-1}) \cap B$. In particular,

$$|\{y \in B \mid 2|p(y)| \geqslant \max_B |p|\}| \geqslant c(N,d) > 0$$

Finally, we may translate B without affecting any of the constants which concludes the proof. □

The following lemma is formulated so as to fit Theorem 9.9.

LEMMA 9.13. *Let p be any nonzero polynomial in $\mathbb{R}^d$. Then there exists $\varepsilon_0 > 0$ depending on p such that*

$$|\{x \in B \mid |p(x)| \leqslant \varepsilon_0\}| \leqslant \beta |B|$$

for all unit balls B. Here $\beta \in (0,1)$ is from Lemma 9.12.

PROOF. By the previous lemma it suffices to show that

$$\inf_B \max_{x \in B} |p(x)| \geqslant \varepsilon_0$$

where the infimum is taken over all unit balls. First, with N the degree of p,

$$\max_{B(0,1)} |p(x)| \geqslant C(N)^{-1} \sum_\alpha |a_\alpha|$$

where a_α are the coefficients of p. Furthermore, one has

$$\max_B |p(x)| \geqslant C(N,p)^{-1} \sum_{|\alpha|=N} |a_\alpha| =: \varepsilon_0$$

uniformly in B. This follows form the previous inequality and the property that the highest order coefficients are invariant under translation. This concludes the proof. $\square$

PROOF OF PROPOSITION 9.11. Lemma 9.13 guarantees that

$$F := \{x \in \mathbb{R}^d \mid |p(x)| \leqslant \varepsilon_0\}$$

is $B(0,1)$-negligible. Thus, by Theorem 9.9,

$$\begin{aligned}\|\bar{p}(D)\varphi\|_2 = \|\bar{p}(\xi)\hat{\varphi}\|_2 &\geqslant \|\bar{p}\hat{\varphi}\|_{L^2(F^c)} \\ &\geqslant \varepsilon_0\|\hat{\varphi}\|_{L^2(F^c)} \geqslant \varepsilon_1\|\hat{\varphi}\|_{L^2(\mathbb{R}^d)}\end{aligned}$$

with some constant ε_1 that depends on p and d. $\square$

There is no analogue of Theorem 9.10 for partial differential operators with nonconstant smooth coefficients. This is known as *Lewy's* example, see for example [**17**].

Notes

A standard reference for the material in this chapter is Havin and Jöricke's book [**24**]. Another basic resource in this area is Charles Fefferman's survey [**15**]. An explicit constant in Theorem 9.4 of the form $e^{C|E||F|}$ in dimension 1 was obtained by Nazarov [**38**], and in higher dimensions by Jaming [**27**]. For the solvability of linear constant coefficient PDEs see Jerison [**29**].

Problems

PROBLEM 9.1. Does there exist a nonzero function $f \in L^2(\mathbb{R}^d)$ so that both $f = 0$ and $\hat{f} = 0$ on nonempty open sets?

PROBLEM 9.2. Does there exist a nonzero $f \in L^2(\mathbb{R})$ with $f = 0$ on $[-1,1]$ and $\hat{f} = 0$ on a half-line? What about $f = 0$ on a set of positive measure and $\hat{f} = 0$ on a half-line?

PROBLEM 9.3. Suppose $E, F \subset \mathbb{R}^d$ have finite measure. Show that for any $g_1, g_2 \in L^2(\mathbb{R}^d)$ there exists $f \in L^2(\mathbb{R}^d)$ with

$$f = g_1 \text{ on } E, \qquad \hat{f} = g_2 \text{ on } F$$

This can be seen as the statement that $f|_E$ and $\hat{f}|_F$ are independent.

PROBLEM 9.4. Prove that for $E, F \subset \mathbb{R}^d$ of finite measure one has

$$\dim\{f \in L^2(\mathbb{R}^d) \mid f = 0 \text{ on } E, \quad \hat{f} = 0 \text{ on } F\} = \infty$$

PROBLEM 9.5. There are several ways of proving Proposition 9.6. As an alternative to the Schur lemma proof we gave, one can use Sobolev embedding on cubes together with Bernstein's theorem. This problem suggests yet another route.

- Prove Poincaré's inequality: for any $f \in H^1(\mathbb{R}^d)$ one has for every $R > 0$

$$\int_{B(0,R)} |f(x) - \langle f\rangle_R|^2\,dx \leqslant CR^2 \int_{B(0,R)} |\nabla f(x)|^2\,dx$$

  with an absolute constant. Here $\langle f\rangle_R$ is the average of f on $B(0,R)$.

- Combine this inequality with Bernstein's estimate as in Lemma 9.1 to give an independent proof of Proposition 9.6.

PROBLEM 9.6. Show that f in Theorem 9.10 may be taken to be $C^\infty(\Omega)$. *Hint:* Use Sobolev spaces in the existence part in addition to L^2.

PROBLEM 9.7. Show Poincaré's inequality for general powers $1 \leqslant p < \infty$, i.e., verify that for any $f \in \mathcal{S}(\mathbb{R}^d)$ one has

$$\int_{B(0,R)} |f(x) - \langle f \rangle_R|^p \, dx \leqslant C(p,d)\, R^p \int_{B(0,R)} |\nabla f(x)|^p \, dx$$

Deduce from this that $\dot{W}^{1,d}(\mathbb{R}^d) \hookrightarrow \mathrm{BMO}(\mathbb{R}^d)$, where $\|f\|_{\dot{W}^{1,d}(\mathbb{R}^d)} = \|\nabla f\|_{L^d(\mathbb{R}^d)}$.

CHAPTER 10

Littlewood-Paley theory

One of the central concerns of harmonic analysis is the study of multiplier operators, i.e., operators T which are of the form

$$(T_m f)(x) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi} m(\xi)\, \hat{f}(\xi)\, d\xi \tag{10.1}$$

where $m : \mathbb{R}^d \to \mathbb{C}$ is bounded. By Plancherel's theorem $\|T_m\|_{2\to 2} \leqslant \|m\|_\infty$. In fact, it is easy to see that one has equality here. Over the distributions we may write $T_m f = K * f$ for any $f \in \mathcal{S}(\mathbb{R}^d)$ with $K := \check{m} \in \mathcal{S}'$. Then T_m is bounded on L^1 if and only if the associated kernel is in L^1, in which case

$$\|T_m\|_{1\to 1} = \|K\|_1$$

There are many cases, though, where $K \notin L^1(\mathbb{R}^d)$ but T is bounded on L^p for some (or all) $1 < p < \infty$. The Hilbert transform is one such example. We shall now discuss one of the basic results in the field, which describes a large class of multipliers m that give rise to L^p bounded operators for $1 < p < \infty$.

In Chapter 7 we previously studied the case where m is homogeneous of degree 0 and we showed that the T_m is a Calderón-Zygmund operator with homogeneous kernel, and thus L^p-bounded for $1 < p < \infty$. In the following theorem, which goes by the name of Mikhlin multiplier theorem, we shall allow for more general functions m which satisfy finitely many derivative conditions *of the zero-degree type.*.

Theorem 10.1. *Let $m : \mathbb{R}^d\backslash\{0\} \to \mathbb{C}$ satisfy, for any multi-index γ of length $|\gamma| \leqslant d+2$*

$$|\partial^\gamma m(\xi)| \leqslant B|\xi|^{-|\gamma|}$$

for all $\xi \neq 0$. Then for any $1 < p < \infty$ there is a constant $C = C(d,p)$ so that

$$\|(m\hat{f})^\vee\|_p \leqslant CB\|f\|_p$$

for all $f \in \mathcal{S}$.

Proof. Let ψ give rise to a dyadic partition of unity as in Lemma 8.3. Define for any $j \in \mathbb{Z}$

$$m_j(\xi) = \psi(2^{-j}\xi) m(\xi)$$

and set $K_j = \check{m}_j$. Now fix some large positive integer N and set

$$K(x) = \sum_{j=-N}^{N} K_j(x)$$

We claim that under our smoothness assumption on m one has the pointwise estimates

$$|K(x)| \leqslant CB|x|^{-d}, \qquad |\nabla K(x)| \leqslant CB|x|^{-d-1} \tag{10.2}$$

where $C = C(d)$. One then applies the Calderon-Zygmund theorem and lets $N \to \infty$.

We will verify the first inequality in (10.2). The second one is similar, and will be only sketched. By assumption, $\|D^\gamma m_j\|_\infty \leqslant CB2^{-j|\gamma|}$ and thus

$$\|D^\gamma m_j\|_1 \leqslant CB2^{-j|\gamma|}2^{jd}$$

for any multi-index $|\gamma| \leqslant d+2$. Similarly,

$$\|D^\gamma(\xi_i m_j)\|_1 \leqslant CB2^{-j(|\gamma|-1)}2^{jd}$$

for the same γ. Hence,

$$\|x^\gamma\, \check{m}_j(x)\|_\infty \leqslant CB2^{j(d-|\gamma|)}$$

and

$$\|x^\gamma D\check{m}_j(x)\|_\infty \leqslant CB2^{j(d+1-|\gamma|)}\ .$$

Since $|x|^k \leqslant C(k,d)\sum_{|\gamma|=k}|x^\gamma|$ one concludes that

$$|\check{m}_j(x)| \leqslant CB2^{j(d-k)}|x|^{-k} \tag{10.3}$$

and

$$|D\check{m}_j(x)| \leqslant CB2^{j(d+1-k)}|x|^{-k} \tag{10.4}$$

for any $0 \leqslant k \leqslant d+2$ and all $j \in \mathbb{Z}$, $x \in \mathbb{R}^d \setminus \{0\}$. We shall use this with $k = 0$ and $k = d+2$. Indeed,

$$\begin{aligned}
|K(x)| &\leqslant \sum_j |\check{m}_j(x)| \leqslant \sum_{2^j \leqslant |x|^{-1}} |\check{m}_j(x)| + \sum_{2^j > |x|^{-1}} |\check{m}_j(x)| \\
&\leqslant CB \sum_{2^j \leqslant |x|^{-1}} 2^{jd} + CB \sum_{2^j > |x|^{-1}} 2^{jd}(2^j|x|)^{-(d+2)} \\
&\leqslant CB|x|^{-d} + CB|x|^2|x|^{-d-2} = CB|x|^{-d}\ ,
\end{aligned}$$

as claimed. To obtain the second inequality in (10.2) one uses (10.4) instead of (10.3). Otherwise the argument is unchanged. Thus we have verified that K satisfies the conditions *i*) and *iii*) of Definition 7.1, see Lemma 7.2. Furthermore, $\|m\|_\infty \leqslant B$ so that $\|(m\hat{f})^\vee\|_2 \leqslant B\|f\|_2$. By Theorem 7.6, and the remark following it, one concludes that

$$\|(m\hat{f})^\vee\|_p \leqslant C(p,d)\|f\|_p$$

for all $f \in \mathcal{S}$ and $1 < p < \infty$, as claimed. □

The number of derivatives entering into the hypothesis can be decreased, see the problem section. Theorem 10.1 allows one to give proofs of Corollaries 7.7 without going through Exercise 7.4. Indeed, one has

$$\widehat{\frac{\partial^2 u}{\partial x_i \partial x_j}}(\xi) = \frac{\xi_i \xi_j}{|\xi|^2}\widehat{\triangle u}(\xi)$$

Since $m(\xi) = \frac{\xi_i\xi_j}{|\xi|^2}$ satisfies the conditions of Theorem 10.1 (this is obvious, as m is homogeneous of degree 0 and smooth away from $\xi = 0$), we are done. Observe that this example also shows that Theorem 10.1 fails if $p = 1$ or $p = \infty$.

Theorem 10.1 is formulated as an apriori inequality for $f \in \mathcal{S}(\mathbb{R}^d)$. In that case $m\hat{f} \in \mathcal{S}'$ so that $(m\hat{f})^\vee \in \mathcal{S}'$. The question arises whether or not $(m\hat{f})^\vee$ is meaningful for $f \in L^p(\mathbb{R}^d)$. If $1 < p \leqslant 2$, then $\hat{f} \in L^2 + L^\infty(\mathbb{R}^d)$ (in fact $\hat{f} \in L^{p'}$ by Hausdorff-Young), so that $m\hat{f} \in \mathcal{S}'$ and therefore $(m\hat{f})^\vee \in \mathcal{S}'$ is well-defined. If $p > 2$, however, it is known that there are $f \in L^p(\mathbb{R}^d)$ such that $\hat{f} \in \mathcal{S}'$ has positive order (see the Notes to Chapter 6). For such f it is in general not possible to define $(m\hat{f})^\vee$ in $\mathcal{S}'$.

We now present an important application of Mikhlin's theorem to Littlewood-Paley theory. With ψ as in Lemma 8.3 we define $P_jf = (\psi_j\hat{f})^\vee$ where $\psi_j(\xi) = \psi(2^{-j}\xi)$ (it is customary to assume $\psi \geqslant 0$). Then by Plancherel,

$$C^{-1}\|f\|_2^2 \leqslant \sum_{j\in\mathbb{Z}} \|P_jf\|_2^2 \leqslant \|f\|_2^2 \tag{10.5}$$

for any $f \in L^2(\mathbb{R}^d)$. Observe that the middle expression is equal to $\|Sf\|_2^2$ with

$$Sf = \Big(\sum_j |P_jf|^2\Big)^{\frac{1}{2}}$$

This is called the Littlewood-Paley *square-function*. It is a basic result of harmonic analysis that (10.5) generalizes to

$$C^{-1}\|f\|_p \leqslant \|Sf\|_p \leqslant C\|f\|_p \tag{10.6}$$

for any $f \in L^p(\mathbb{R}^d)$ provided $1 < p < \infty$ and with $C = C(p, d)$. It is of course important to understand why such a result might hold, and so we offer some heuristic explanations.

Consider a lacunary trigonometric polynomial of the form

$$T(\theta) = \sum_{n=1}^{M} c_n\, e(2^n\pi\theta)$$

where $\theta \in \mathbb{T}$. Let $p \geqslant 1$ be an integer and estimate by Plancherel's theorem

$$\begin{aligned}\int_0^1 \Big|\sum_{n=1}^{M} c_n\, e(2^n\theta)\Big|^{2p} d\theta &= \int_0^1 \Big|\sum_{n_1,\dots,n_p=1}^{M} c_{n_1}\cdots c_{n_p}\, e(2^n\theta)\Big|^2 d\theta \\ &\leqslant C(p) \sum_{n_1,\dots,n_p=1}^{M} |c_{n_1}\cdots c_{n_p}|^2 = \Big(\sum_n |c_n|^2\Big)^p\end{aligned}$$

We used here that for every positive integer N the number of ways in which it can be written in the form

$$N = 2^{n_1} + \cdots + 2^{n_p}, \quad n_j \geqslant 0$$

is bounded by a constant that depends only on p. In other words, given a sequence $\{c_k\} \in \ell^2(\mathbb{Z})$ so that $c_k = 0$ unless $|k|$ is of the form 2^n with $n \geqslant 0$, we have the property

$$\Big(\sum_k |c_k|^2\Big)^{\frac{1}{2}} \leqslant \Big\| \sum_k c_k\, e(k\pi\theta) \Big\|_p \leqslant \Big(\sum_k |c_k|^2\Big)^{\frac{1}{2}} \tag{10.7}$$

for every $p \geqslant 2$ (and by Hölder's inequality one can lower this to $p \geqslant 1$, see the proof of Lemma 10.3 below). Clearly, (10.7) is remarkable as it says that any $L^2(\mathbb{T})$ function f so that $\hat{f}(k) = 0$ unless $|k|$ is a power of 2 has finite L^p-norm for every $p < \infty$ (but of course not for $p = \infty$). We remark in passing that this applies to any lacunary or gap Fourier series, but this is harder to see (the standard argument involves Riesz products).

The rough idea behind (10.6) can be construed as follows: if we consider an arbitrary trigonometric polynomial $\sum_n a_n e(n\theta)$, then (10.7) is of course false. However, we may try to salvage something of (10.7) by setting

$$c_k(\theta) := \sum_{2^{k-1} < |n| \leqslant 2^k} a_n e(n\theta), \qquad \forall\, k \geqslant 1$$

We might then expect that the square function

$$S f(\theta) := \Big(\sum_k |c_k(\theta)|^2\Big)^{\frac{1}{2}}$$

satisfies $\|Sf\|_p \simeq \|f\|_p$.

It is insightful to view lacunary characters such as $\{e(2^n\theta)\}_{n\geqslant 1}$ from a probabilistic angle. Indeed, we can view these functions as being close to independent random variables on the interval $[0,1]$. For example, on any interval $[k2^{-n}, (k+1)2^{-n}) \subset [0,1)$ with integer k, on which $\sin(2^n\pi\theta)$ has a fixed sign, the function $\sin(2^{n+1}\pi\theta)$ is equally likely to be positive or negative. The reader should also compare the graphs of $\operatorname{Re} e(2^n\theta)$ with those of the Rademacher functions r_n defined below.

Taking another leap of faith, we are thus lead to viewing lacunary series from the point of view of the *central limit theorem* in probability theory, which at least heuristically implies that a lacunary Fourier series f as above with $\|f\|_2 = 1$ satisfies

$$\big|\{\theta \in \mathbb{T} \mid |f(\theta)| > \lambda\}\big| \leqslant C e^{-\frac{\lambda^2}{2}}$$

with a suitable constant $C > 0$. In particular, this ensures that f satisfies an estimate of the type (10.7).

It should therefore come as no surprise to the reader that some probabilistic elements enter into the proof of (10.6). In fact, we shall now derive (10.6) by means of a standard randomization technique. Let $\{r_j\}$ be a sequence of independent random variables with $\mathbb{P}[r_j = 1] = \mathbb{P}[r_j = -1] = \frac{1}{2}$, for all j (in other words, the r_j are a coin tossing sequence). Readers familiar with the central limit theorem will not be surprised by the following sub-Gaussian bound.

Lemma 10.2. *For any positive integer N and $\{a_j\}_{j=1}^N \subset \mathbb{C}$ one has*

$$\mathbb{P}\Big[\big|\sum_{j=1}^N r_j a_j\big| > \lambda\big(\sum_{j=1}^N |a_j|^2\big)^{\frac12}\Big] \leqslant 4e^{-\frac{\lambda^2}{2}} \tag{10.8}$$

for all $\lambda > 0$.

Proof. Assume first that $a_j \in \mathbb{R}$. Then

$$\mathbb{E}\, e^{tS_N} = \Pi_{j=i}^N \mathbb{E}\left(e^{t\,r_j a_j}\right) = \Pi_{j=1}^N \cosh(ta_j)$$

where we have set $S_N = \sum_{j=1}^N r_j a_j$. Now invoke the calculus fact

$$\cosh x \leqslant e^{x^2/2} \quad \forall\, x \in \mathbb{R}$$

to conclude that

$$\mathbb{E}\, e^{tS_N} \leqslant \Pi_{j=1}^N e^{t^2 a_j^2/2} = \exp\left(t^2 \sum_{j=1}^N a_j^2/2\right)$$

Hence, with $\sigma^2 = \sum_{j=1}^N a_j^2$,

$$\mathbb{P}\big[S_N > \lambda\sigma\big] \leqslant e^{t^2\sigma^2/2} e^{-\lambda t\sigma} \leqslant e^{-\lambda^2/2}$$

where the final inequality follows by minimizing in t, i.e., by setting $t = \frac{\lambda}{\sigma}$. Similarly,

$$\mathbb{P}[S_N < -\lambda\sigma] \leqslant e^{-\lambda^2/2}$$

so that

$$\mathbb{P}[|S_N| > \lambda\sigma] \leqslant 2e^{-\lambda^2/2}$$

The case of $a_j \in \mathbb{C}$ follows by means of a decomposition into real and imaginary parts. □

This sub-Gaussian estimate on the so-called "tails" of the distribution of the random sum $S_N := \sum_{j=1}^N r_j a_j$ imply that all expressions $(\mathbb{E}\,|S_N|^p)^{\frac1p}$ are comparable. This fact is known as Khinchin's inequality.

Lemma 10.3. *For any $1 \leqslant p < \infty$ there exists a constant $C = C(p)$ so that*

$$C^{-1}\Big(\sum_{j=1}^N |a_j|^2\Big)^{\frac{p}{2}} \leqslant \mathbb{E}\Big|\sum_{j=1}^N r_j a_j\Big|^p \leqslant C\Big(\sum_{j=1}^N |a_j|^2\Big)^{\frac{p}{2}} \tag{10.9}$$

for any choice of positive integer N and $\{a_j\}_{j=1}^N \in \mathbb{C}$.

Proof. We start with the upper bound in (10.9). It suffices to consider the case $\sum_{j=1}^N |a_j|^2 = 1$. Setting $\sum_{j=1}^N a_j r_j = S_N$ one has

$$\mathbb{E}\,|S_N|^p = \int_0^\infty \mathbb{P}[|S_N| > \lambda] p\lambda^{p-1}\, d\lambda \leqslant \int_0^\infty 4e^{-\lambda^2/2} p\lambda^{p-1}\, d\lambda =: C(p) < \infty$$

For the lower bound it suffices to assume that $1 \leqslant p \leqslant 2$, in fact, $p = 1$. By Hölder's inequality,

$$\begin{aligned}\mathbb{E}\,|S_N|^2 &= \mathbb{E}\,|S_n|^{2/3}|S_N|^{4/3} \leqslant (\mathbb{E}\,|S_N|)^{2/3}(\mathbb{E}\,|S_n|^4)^{1/3}\\ &\leqslant C(\mathbb{E}\,|S_N|)^{2/3}(\mathbb{E}\,|S_N|^2)^{4/3}\end{aligned}$$

where the final inequality follows from the case $2 \leqslant p < \infty$ just considered. This implies that

$$\mathbb{E}\,|S_N|^2 \leqslant C(\mathbb{E}\,|S_N|)^2$$

and we are done. □

Khinchin's inequality is often formulated for the *Rademacher* functions, which are a concrete realization of the sequence $\{r_j\}$ on the interval $[0,1]$ given by $r_j(x) = \operatorname{sign}(\sin(\pi 2^j x))$ for $j \geqslant 0$. The explicit form of the Rademacher functions allows for a different proof of Khinchin's inequality, which are based on expanding and integrating out $\mathbb{E}\,|S_N|^p$ when p is an even integer. We invite the reader to determine the growth (at least an upper bound on the growth) of the constant $C(p)$ in Khinchin's inequality as $p \to \infty$.

Exercise 10.1.

- Draw the graphs of the first few Rademacher functions and explain whey they are a realization of independent identically distributed random variables as those appearing above (coin tossing sequence).
- By explicit expansion and integration for even integer p, prove that

$$\int_0^1 \Big|\sum_{j=1}^N a_j\, r_j(t)\Big|^p\, dt \leqslant C(p)\Big(\sum_{j=1}^N |a_j|^2\Big)^{\frac{p}{2}}$$

Then recover the general case of $1 \leqslant p < \infty$.

We are now ready to prove the Littlewood-Paley theorem (10.6).

Theorem 10.4 (Littlewood-Paley). *For any $1 < p < \infty$ there is a constant $C = C(p,d)$ such that*

$$C^{-1}\|f\|_p \leqslant \|Sf\|_p \leqslant C\|f\|_p$$

for any $f \in \mathcal{S}$.

Proof. Let $\{r_j\}$ be as above. The proof rests on the fact that

$$m(\xi) := \sum_{j=-N}^{N} r_j \psi_j(\xi)$$

satisfies the conditions of the Mikhlin multiplier theorem uniformly in N and uniformly in the realization of the random variables $\{r_j\}$. Indeed, for any γ,

$$|D^\gamma m(\xi)| \leqslant \sum_{j=-N}^{N} |D^\gamma \psi_j(\xi)| \leqslant C \sum_{j=-N}^{N} |\xi|^{-\gamma}|(D^\gamma\psi)(2^{-j}\xi)| \leqslant C|\xi|^{-\gamma}$$

To pass to the final inequality one uses that only an absolutely bounded number of terms is non-zero in the sum preceding it for any $\xi \neq 0$. Hence, in view of Lemma 10.3,

$$\int_{\mathbb{R}^d} \Big|(Sf)(x)\Big|^p\,dx \leqslant C \limsup_{N\to\infty} \mathbb{E} \int_{\mathbb{R}^d} \Big|\sum_{j=-N}^{N} r_j (P_j f)(x)\Big|^p\,dx \leqslant C\|f\|_p^p$$

as desired.

To prove the lower bound we use duality: choose a function $\tilde\psi$ so that $\tilde\psi = 1$ on $\operatorname{supp}(\psi)$ and $\tilde\psi$ is compactly supported with $\operatorname{supp}\tilde\psi \subset \mathbb{R}^d\backslash\{0\}$. Defining $\tilde P_j$ like P_j with $\tilde\psi$ instead of ψ yields $\{\tilde P_j\}$ satisfying $\tilde P_j P_j = P_j$. Therefore, for any $f, g \in \mathcal{S}$, and any $1 < p < \infty$,

$$\begin{aligned} |\langle f, g\rangle| = |\sum_j \langle P_j f, \tilde P_j g\rangle| &\leqslant \int_{\mathbb{R}^d} \Big(\sum_j |(P_j f)(x)|^2\Big)^{\frac12} \Big(\sum_k |(\tilde P_k g)(x)|^2\Big)^{\frac12}\,dx \\ &\leqslant \|Sf\|_p \|\tilde S g\|_{p'} \leqslant C\|Sf\|_p \|g\|_{p'} \end{aligned}$$

For the final bound we use that the argument for the upper bound equally well applies to $\tilde S$ instead of S. Thus, $\|f\|_p \leqslant C\|Sf\|_p$, as claimed. □

It is desirable to formulate Theorem 10.4 on $L^p(\mathbb{R}^d)$ rather than on $\mathcal{S}(\mathbb{R}^d)$. This is done in the following corollary. Observe that Sf is defined pointwise if $f \in \mathcal{S}'(\mathbb{R}^d)$ since $P_j f$ is a smooth function.

Corollary 10.5.

a) *Let $1 < p < \infty$. Then for any $f \in L^p(\mathbb{R}^d)$ one has $Sf \in L^p$ and*

$$C_{p,d}^{-1}\|f\|_p \leqslant \|Sf\|_p \leqslant C_{p,d}\|f\|_p$$

b) *Suppose that $f \in \mathcal{S}'$ and that $Sf \in L^p(\mathbb{R}^d)$ with some $1 < p < \infty$. Then $f = g + P$ where P is a polynomial and $g \in L^p(\mathbb{R}^d)$. Moreover, $Sf = Sg$ and*

$$C_{p,d}^{-1}\|Sf\|_p \leqslant \|g\|_p \leqslant C_{p,d}\|Sf\|_p$$

Proof. To prove part $a)$, let $f_k \in \mathcal{S}$ so that $\|f_k - f\|_p \to 0$ as $k \to \infty$. We claim that

$$\lim_{k\to\infty} \|Sf_k - Sf\|_p = 0 \tag{10.10}$$

If (10.10) holds, then passing to the limit $k \to \infty$ in

$$C_{p,d}^{-1}\|f_k\|_p \leqslant \|S(f_k)\|_p \leqslant C_{p,d}\|f_k\|_p$$

implies part $a)$. To prove (10.10) one applies Fatou repeatedly. Fix $x \in \mathbb{R}^d$. Then

$$\begin{aligned} |Sf_k(x) - Sf(x)| &= |\|\{P_j f_k(x)\}_{k=-\infty}^{\infty}\|_{\ell^2} - \|\{P_j f(x)\}_{j=-\infty}^{\infty}\|_{\ell^2}| \\ &\leqslant \|\{P_j f_k(x) - P_j f(x)\}_{j=-\infty}^{\infty}\|_{\ell^2} \\ &= S(f_k - f)(x) \leqslant \liminf_{m\to\infty} S(f_k - f_m)(x) \end{aligned}$$

Therefore,

$$\|S f_k - S f\|_p \leqslant \liminf_{m\to\infty} \|S(f_k - f_m)\|_p \leqslant C_{p,d} \liminf_{m\to\infty} \|f_k - f_m\|_p$$

The claim (10.10) now follows by letting $k \to \infty$.

For the second part we argue as in the proof of the lower bound in Theorem 10.4: Let $f \in \mathcal{S}'$ and $h \in \mathcal{S}$ with $\mathrm{supp}(\hat{h}) \subset \mathbb{R}^d \backslash \{0\}$. Then

$$|\langle f, \bar{h}\rangle| \leqslant C_p \|S f\|_p \|\tilde{S} h\|_{p'} \leqslant C_p \|S f\|_p \|h\|_{p'}$$

where $\tilde{S}$ is the modified square function from the proof of Theorem 10.4. By the Hahn-Banach theorem there exists $g \in L^p$ so that $\langle f, \bar{h}\rangle = \langle g, h\rangle$ for all h as above satisfying $\|g\|_p \leqslant C_p \|S f\|_p$. By Exercise (6.2) one has $f = g + P$. Moreover, $S f = S g$ and

$$\|S f\|_p = \|S g\|_p \leqslant C \|g\|_p$$

by part *i*). □

The following exercise shows that the Littlewood-Paley theorem fails at the endpoints $p = 1$ and $p = \infty$.

Exercise 10.2.

a) Show that Theorem 10.4 fails at $p = 1$. The intuition is of course to take $f = \delta_0$. For that case verify that that

$$(S f)(x) \simeq |x|^{-d}$$

so that $S f \notin L^1(\mathbb{R}^d)$. Next, transfer this logic to $L^1(\mathbb{R}^d)$ by means of approximate identities.

b) Show that Theorem 10.4 fails for $p = \infty$.

It is natural to ask at this point whether the Littlewood-Paley theorem holds for square functions defined in terms of *sharp cut-offs* rather than smooth ones as above. More precisely, suppose we set

$$(S_{\mathrm{new}} f)^2 = \sum_{j\in\mathbb{Z}} |(\chi_{[2^{j-1}\leqslant|\xi|<2^j]} \hat{f}(\xi))^\vee|^2$$

Is it true that

$$C_p^{-1} \|f\|_p \leqslant \|S_{\mathrm{new}} f\|_p \leqslant C_p \|f\|_p \qquad (10.11)$$

for $1 < p < \infty$? Due the boundedness of the Hilbert transform the answer is “yes” in dimension 1, but it is rather deep fact that it is “no” in dimensions $d \geqslant 2$, see the notes and problems to this chapter.

As a first application of (10.6) we now finish the proof of (6.8).

Proposition 10.6. *For any* $0 \leqslant s < d$ *and with* $\frac{1}{2} - \frac{1}{p} = \frac{s}{d}$ *one has*

$$\|f\|_{L^p(\mathbb{R}^d)} \leqslant C(s,d)\, \|f\|_{\dot{H}^s(\mathbb{R}^d)}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$.

Proof. In Chapter 6 we already settled the case where $\hat{f}$ is localized to a dyadic shell $\{|\xi| \simeq 2^j\}$. Thus, it remains to prove (6.11). In fact, we shall now prove that for any finite $p \geqslant 2$

$$\|f\|_p \leqslant C \Big(\sum_{j\in\mathbb{Z}} \|P_j f\|_p^2 \Big)^{\frac{1}{2}}$$

with a constant that only depends on p and d. By (10.6)

$$\begin{aligned}\|f\|_p &\leqslant C(p,d)\, \|Sf\|_p = C(p,d) \| \, \|\{P_j f\}_{j\in\mathbb{Z}}\|_{\ell^2(\mathbb{Z})} \, \|_p \\ &\leqslant C(p,d) \big\| \{\|P_j f\|_p\}_{j\in\mathbb{Z}} \big\|_{\ell^2(\mathbb{Z})}\end{aligned}$$

where we used the triangle inequality in the final step (see the following exercise).

□

Exercise 10.3.

- Show that for any $\infty \geqslant p \geqslant q \geqslant 1$ and on any measure space (X,μ),
$$\big\| \, \|\{f_j\}_{j\in\mathbb{Z}}\|_{\ell^q(\mathbb{Z})} \, \big\|_{L^p(\mu)} \leqslant \big\| \, \{\|f_j\|_{L^p(\mu)}\}_{j\in\mathbb{Z}} \, \big\|_{\ell^q(\mathbb{Z})}$$
Also show that for $\infty \geqslant q \geqslant p \geqslant 1$ one has the reverse inequality.
- Show that (6.11) fails for every $1 \leqslant p < 2$.

As evidenced by the proof of Proposition 10.6, Exercise 10.3 is a useful (and sharp) tool when combined with (10.6). In particular, it is used in comparing Sobolev with Besov spaces, see the problem section below.

As final general comment about Littlewood-Paley theory, we ask the reader to verify that there can be no version of 10.6 in which the square function is replaced by some ℓ^q-based norm.

Exercise 10.4. Suppose that for some $1 \leqslant q \leqslant \infty$ and $1 \leqslant p \leqslant \infty$ one has

$$\big\| \, \|\{P_j f\}_{j\in\mathbb{Z}}\|_{\ell^q(\mathbb{Z})} \, \big\|_{L^p(\mathbb{R}^d)} \simeq \|f\|_p \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^d)$$

with constants that do not depend on f. Show that $q = 2$.

We now present a connection between BMO and Littlewood-Paley theory. This will give us the opportunity to introduce some notions that historically played an important role in the development of the theory, such as *Carleson measure* and the *Calderón reproducing formula*, which can be seen as a continuum version of the discrete Littlewood-Paley decomposition.

Definition 10.7. Let μ be a (complex) measure on the upper half-space $\mathbb{R}^{d+1}_+$. Denote a general dyadic cube in $\mathbb{R}^d$ by I and the cube in $\mathbb{R}^{d+1}_+$ with base I and of height $\ell(I)$ where ℓ is the side-length of I, by $Q(I)$. In other words, $Q(I) := \{(x,t) \mid x \in I,\ 0 < t < \ell(I)\}$. One says that μ is a Carleson measure if

$$\|\mu\|_C := \sup_I \frac{|\mu|(Q(I))}{|I|} < \infty$$

The following result may seem unmotivated for now, but it is a standard device by which a BMO function can be broken up into basic constituents.

Proposition 10.8. *Let $f \in \mathrm{BMO}(\mathbb{R}^d)$ with $\|f\|_{\mathrm{BMO}} \leqslant 1$ have compact support and mean zero. There exist functions $\{b_I\}_I$ indexed by the dyadic cubes in $\mathbb{R}^d$ and scalars λ_I such that*

- $f = \sum_I \lambda_I b_I$
- $\mathrm{supp}(b_I) \subset 3I$
- $\int_{\mathbb{R}^d} b_I(x)\, dx = 0$
- $\|\nabla b_I\|_\infty \leqslant \ell(I)^{-1}$
- $\sum_{I \subset J} |\lambda_I|^2 |I| \leqslant |J|$ *for all dyadic cubes J*

The final property means that $\sum_I |\lambda_I|^2 |I| \delta_{(x_I, \ell(I))}$ is a Carleson measure where x_I is the center of I.

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ be real-valued, radial, and such that $\mathrm{supp}(\varphi) \subset B(0,1)$ and with

$$\int_{\mathbb{R}^d} \hat{\varphi}(t\xi)^2 \frac{dt}{t} = 1 \qquad \forall\, \xi \neq 0 \tag{10.12}$$

Note that the left-hand side is a radial function and by scaling does not even depend on $|\xi|$. In other words, it does not depend on the choice of $\xi \neq 0$ at all. The only requirement is that $\hat{\varphi}(0) = 0$ or equivalently, $\int \varphi = 0$. Moreover, any $\varphi \not\equiv 0$ which satisfies this condition can be normalized so as to fulfil (10.12).

Clearly, (10.12) is a continuum analogue of the discrete partition of unity in Lemma 8.3. The Calderón reproducing formula is now the following statement, which is implied by (10.12) and the vanishing $\int f = 0$:

$$\begin{aligned} f(x) &= \int_0^\infty (\varphi_t * \varphi_t * f)(x) \,\frac{dt}{t} \\ &= \sum_I \int_{T(I)} \varphi_t(x-y)(\varphi_t * f)(y)\, \frac{dt}{t} dy = \sum_I \tilde{b}_I(x) \end{aligned} \tag{10.13}$$

where $T(I) := \{(x,t) \mid x \in I,\ t \in (\ell(I)/2, \ell(I))\}$ and $\varphi_t(x) = t^{-d}\varphi(x/t)$. By construction, $\int \tilde{b}_I = 0$ and $\mathrm{supp}(\tilde{b}_I) \subset 3I$. Moreover, if we define with a suitable constant C

$$\lambda_I := C\Big(|I|^{-1} \int_{T(I)} |\varphi_t * f(y)|^2 t^{-1}\, dtdy\Big)^{\frac{1}{2}}$$

then one checks by means of Cauchy-Schwarz that

$$|\nabla \tilde{b}_I(x)| = \Big| \int_{T(I)} \nabla\varphi_t(x-y)(\varphi_t * f)(y)\, \frac{dt}{t} dy \Big| \leqslant \lambda_I\, \ell(I)^{-1}$$

Define $b_I := \lambda_I^{-1} \tilde{b}_I$. We have verified all properties up to the Carleson measure. To obtain this final claim, we first show that for any $f \in \mathrm{BMO}(\mathbb{R}^d)$ and any (dyadic) cube $I \subset \mathbb{R}^d$

$$\int_{Q(I)} |(\varphi_t * f)(y)|^2\, t^{-1}\, dtdy \leqslant C\|f\|_{\mathrm{BMO}}^2 |I| \tag{10.14}$$

By the scaling and translation symmetries, we may assume that I is the cube $[-1,1]^d$. Define $I^* := [-2,2]^d$ and split

$$f = f_{I^*} + (f - f_{I^*})\chi_{I^*} + (f - f_{I^*})\chi_{\mathbb{R}^d \setminus I^*} =: f_1 + f_2 + f_3$$

The constant f_1 does not contribute to (10.14) and f_2 is estimated by

$$\int_{\mathbb{R}^{d+1}_+} |(\varphi_t * f_2)(y)|^2 \, t^{-1} \, dtdy = \|f_2\|_2^2 \leqslant C\|f\|_{\mathrm{BMO}}^2$$

Finally,

$$\begin{aligned}
&\int_{Q(I)} |(\varphi_t * f_3)(x)|^2 \, t^{-1} \, dtdx \\
&\leqslant \int_0^2 \frac{dt}{t} \int_{(-1,1)^d} dx \, \Big| \int_{\mathbb{R}^d \setminus (-2,2)^d} (f(y) - f_{I^*}) \, t^{-d} \varphi\Big(\frac{x-y}{t}\Big) \, dy \Big|^2 \\
&\leqslant C \int_0^2 \frac{dt}{t} \int_{(-1,1)^d} dx \int_{\mathbb{R}^d \setminus (-2,2)^d} |f(y) - f_{I^*}|^2 \, t^{-d} |\varphi\Big(\frac{x-y}{t}\Big)| \, dy \\
&\leqslant C \int_{\mathbb{R}^d} \frac{|f(y) - f_{I^*}|^2}{1 + |y|^{d+1}} \, dy \leqslant C\|f\|_{\mathrm{BMO}}^2
\end{aligned}$$

as desired. For the final step, see Problem 7.5. Hence (10.14) holds and therefore

$$\begin{aligned}
\sum_{J \subset I} |\lambda_J|^2 |J| &= \sum_{J \subset I} \int_{T(J)} |(\varphi_t * f)(x)|^2 \, t^{-1} \, dtdx \\
&= C \int_{Q(I)} |(\varphi_t * f)(x)|^2 \, t^{-1} \, dtdx \leqslant C\|f\|_{\mathrm{BMO}}^2 |I|
\end{aligned}$$

and we are done. □

Exercise 10.5. Give a detailed proof of (10.13).

To conclude this chapter, we shall show that singular integrals as in Definition 7.1 are bounded on $C^\alpha(\mathbb{R}^d)$, $0 < \alpha < 1$. Problem 7.7 of the previous chapter deals with the C^α-boundedness of the Hilbert transform, which can be proved on the level of the kernel alone. Here we shall follow a different route based on the projections P_j. We begin with a characterization of $C^\alpha(\mathbb{R}^d)$ for $0 < \alpha < 1$ in terms of the projections P_j. The reader should realize that the proof of the following lemma is reminiscent of the proof of Bernstein's Theorem 1.12.

Lemma 10.9. *Let $|f| \leqslant 1$. Then $f \in C^\alpha(\mathbb{R}^d)$ for $0 < \alpha < 1$ if and only if*

$$\sup_{j \in \mathbb{Z}} 2^{j\alpha} \|P_j f\|_\infty \leqslant A \tag{10.15}$$

Moreover, the smallest A for which (10.15) *holds is comparable to* $[f]_\alpha$.

Proof. Set $\check{\psi}_j(x) = 2^{jn}\check{\psi}(2^j x) =: \varphi_j(x)$. Hence $\|\varphi_j\|_1 = \|\check{\psi}\|_1$ for all $j \in \mathbb{Z}$ and

$$\int_{\mathbb{R}^d} \varphi_j(x) x^\gamma \, dx = 0$$

for all multi-indices γ. First assume $f \in C^\alpha$. Then

$$\begin{aligned}|P_j f(x)| &\leqslant \int_{\mathbb{R}^d} |f(x-y) - f(x)|\,|\varphi_j(y)|\,dy \\ &\leqslant \int_{\mathbb{R}^d} [f]_\alpha |y|^\alpha |\varphi_j(y)|\,dy = 2^{-j\alpha}[f]_\alpha \int_{\mathbb{R}^d} |y|^\alpha |\check{\psi}(y)|\,dy\end{aligned}$$

Hence $A \leqslant C[f]_\alpha$, as claimed. Conversely, define

$$g_\ell(x) = \sum_{-\ell \leqslant j \leqslant \ell} (P_j f)(x)$$

for any integer $\ell \geqslant 0$. We need to show that, for all $y \in \mathbb{R}^d$,

$$\sup_\ell |g_\ell(x-y) - g_\ell(x)| \leqslant CA|y|^\alpha$$

with some constant $C = C(d)$. Fix $y \neq 0$ and estimate

$$\Big| \sum_{|y|^{-1} < 2^j \leqslant 2^\ell} (P_j f)(x) \Big| \leqslant \sum_{2^j > |y|^{-1}} A 2^{-j\alpha} \leqslant CA|y|^\alpha \tag{10.16}$$

Secondly, observe that

$$\begin{aligned}|P_j f(x-y) - P_j f(x)| &\leqslant \|\nabla P_j f\|_\infty |y| \leqslant C 2^j \|P_j f\|_\infty |y| \\ &\leqslant C 2^{j(1-\alpha)} A|y|\end{aligned} \tag{10.17}$$

where we invoked Bernstein's inequality to pass to the second inequality sign. Combining (10.16) and (10.17) yields

$$\begin{aligned}&|g_\ell(x-y) - g_\ell(x)| \\ &\leqslant \sum_{2^{-\ell} \leqslant 2^j \leqslant |y|^{-1}} |P_j f(x-y) - P_j f(x)| + \sum_{|y|^{-1} < 2^j \leqslant 2^\ell} 2\|P_j f\|_\infty \\ &\leqslant \sum_{2^j \leqslant |y|^{-1}} CA 2^{j(1-\alpha)}|y| + \sum_{|y|^{-1} < 2^j} 2A 2^{-j\alpha} \leqslant CA|y|^\alpha\end{aligned} \tag{10.18}$$

uniformly in $\ell \geqslant 1$.

So $\{g_\ell - g_\ell(0)\}_{\ell=1}^\infty$ are uniformly bounded on $C^\alpha(K)$ for any compact K. By the Arzela-Ascoli theorem one concludes that

$$g_\ell - g_\ell(0) \to g$$

uniformly on any compact set and therefore $[g]_\alpha \leqslant CA$ by (10.18) (up to passing to a subsequence $\{\ell_i\}$). It remains to show that f has the same property. This will follow from the following claim: $f = g + \text{const}$. To verify this property, note that $g_\ell - g_\ell(0) \to g$ in $\mathcal{S}'$. Thus, also $\hat{g}_\ell - \delta_0 g_\ell(0) \to \hat{g}$ in $\mathcal{S}'$ which is the same as

$$\sum_{-\ell \leqslant j \leqslant \ell} \psi(2^{-j}\xi)\hat{f}(\xi) - \delta_0 g_\ell(0) \to \hat{g} \text{ in } \mathcal{S}'$$

So if $h \in \mathcal{S}$ with $\mathrm{supp}(h) \subset \mathbb{R}^d \setminus \{0\}$, then $\langle \hat{f} - \hat{g}, h \rangle = 0$, i.e., we have $\mathrm{supp}(\hat{f} - \hat{g}) = \{0\}$. By Exercise 6.2 therefore

$$(f - g)(x) = \sum_{|\gamma| \leqslant M} C_\gamma x^\gamma \tag{10.19}$$

On the other hand, since $g(0) = 0$,

$$|(f - g)(x)| \leqslant \|f\|_\infty + |g(x) - g(0)| \leqslant 1 + CA|x|^\alpha$$

Since $\alpha < 1$, comparing this bound with (10.19) shows that the polynomial in (10.19) is of degree zero. Thus, $f - g = \mathrm{const}$, as claimed. □

Next, we need the following result which gives an example of a situation where T maps into $L^1(\mathbb{R}^d)$. We postpone the proof.

Proposition 10.10. *Let K be a singular integral kernel as in Definition 7.1. For any $\eta \in \mathcal{S}(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \eta(x)\, dx = 0$ one has*

$$\|T\eta\|_1 \leqslant C(\eta) B$$

with a constant depending on η.

We can now formulate and prove the C^α bound for singular integrals.

Theorem 10.11. *Let K be as in Definition 7.1 and $0 < \alpha < 1$. Then for any $f \in L^2 \cap C^\alpha(\mathbb{R}^d)$ one has $Tf \in C^\alpha(\mathbb{R}^d)$ and $\left[Tf\right]_\alpha \leqslant C_\alpha B[f]_\alpha$ with $C_\alpha = C(\alpha, d)$.*

Proof. We use Lemma 10.9. In order to do so, notice first that

$$\|Tf\|_\infty \leqslant CB([f]_\alpha + \|f\|_2)$$

which we leave to the reader to check. Therefore, it suffices to show that

$$\sup_j 2^{j\alpha} \|P_j T f\|_\infty \leqslant CB[f]_\alpha \tag{10.20}$$

Let $\tilde{P}_j$ be defined as

$$\tilde{P}_j u = \left(\tilde{\psi}(2^{-j}\cdot)\hat{u}(\xi)\right)^\vee$$

where $\tilde{\psi} \in C_0^\infty(\mathbb{R}^d \setminus \{0\})$ and $\tilde{\psi}\psi = \psi$. Thus, $\tilde{P}_j P_j = P_j$. Hence,

$$\begin{aligned}
\|P_j T f\|_\infty &= \|\tilde{P}_j P_j T f\|_\infty = \|\tilde{P}_j T P_j f\|_\infty \\
&\leqslant \|\tilde{P}_j T\|_{\infty\to\infty} \|P_j f\|_\infty \\
&\leqslant \|\tilde{P}_j T\|_{\infty\to\infty} C[f]_\alpha 2^{-j\alpha}
\end{aligned}$$

It remains to show that $\sup_j \|\tilde{P}_j T\|_{\infty\to\infty} \leqslant CB$, see (10.20). Clearly, the kernel of $\tilde{P}_j T$ is $2^{jd}\, \check{\tilde{\psi}}(2^j\cdot) * K$ so that

$$\|\tilde{P}_j T\|_{\infty\to\infty} \leqslant \|2^{jd}\, \check{\tilde{\psi}}(2^j\cdot) * K\|_1 = \|\check{\tilde{\psi}} * 2^{-jd} K(2^{-j}\cdot)\|_1 \leqslant CB \tag{10.21}$$

by Proposition 10.10. Indeed, set $\eta = \check{\tilde{\psi}}$ in that proposition so that

$$\int_{\mathbb{R}^d} \eta(x)\, dx = \tilde{\psi}(0) = 0$$

as required. Furthermore, we apply Proposition 10.10 with the rescaled kernel $2^{-jd}K(2^{-j}\cdot)$. As this kernel satisfies the conditions in Definition 7.1 *uniformly* in j, one obtains (10.21) and the theorem is proved. □

It remains to show Proposition 10.10. We will build up from the case of an "atom" as defined by the following lemma.

LEMMA 10.12. *Let* $f \in L^\infty(\mathbb{R}^d)$ *with* $\int f(x)\,dx = 0$, $\operatorname{supp}(f) \subset B(0,R)$ *and* $\|f\|_\infty \leqslant R^{-d}$. *Then*

$$\|Tf\|_1 \leqslant CB$$

PROOF. By Cauchy-Schwarz,

$$\begin{aligned}\int_{|x|\leqslant 2R} \big|(Tf)(x)\big|\,dx &\leqslant \|Tf\|_2 CR^{\frac{d}{2}} \leqslant CB\|f\|_2 R^{\frac{d}{2}} \\ &\leqslant CBR^{-d}R^{\frac{d}{2}}R^{\frac{d}{2}} = CB\end{aligned}$$

Furthermore,

$$\int_{|x|>2R} |(Tf)(x)|\,dx \leqslant \int_{\mathbb{R}^d}\int_{|x|>2|y|} |K(x-y)-K(x)|\,dx|f(y)|\,dy \leqslant B\|f\|_1 \leqslant CB$$

as desired. □

In passing, we point out a simple relation between these atoms and BMO. In fact, let $\varphi \in \mathrm{BMO}$ and $a = f$ as defined in Lemma 10.12. Then

$$\begin{aligned}|\langle a,\varphi\rangle| &= \Big|\int_B a(x)(\varphi(x)-\varphi_B)\,dx\Big| \leqslant \|a\|_\infty \int_B |\varphi(x)-\varphi_B|\,dx \\ &\leqslant \fint_B |\varphi(x)-\varphi_B|\,dx \leqslant \|\varphi\|_{\mathrm{BMO}}\end{aligned}$$

Moreover, it is easy to see from this that $\sup_a |\langle a,\varphi\rangle| = \|\varphi\|_{\mathrm{BMO}}$. It is therefore natural to *define the atomic space* $H^1_{\mathrm{atom}}(\mathbb{R}^d)$ as all sums $f = \sum_j c_j a_j$ with $\sum_j |c_j| < \infty$ and to define the norm $\|f\|_{H^1}$ as the smallest possible value of the sum $\sum_j |c_j|$ for a given f.

It is a remarkable fact that this atomic space coincides with H^1 as defined in (7.32). We will establish this in a later chapter in a one-dimensional dyadic setting. In the following lemma we show that one can write a general Schwartz function with mean zero as a superposition of infinitely many atoms.

LEMMA 10.13. *Let* $\eta \in \mathcal{S}(\mathbb{R}^n)$, $\int_{\mathbb{R}^d} \eta(x)\;dx = 0$. *Then one can write*

$$\eta = \sum_{\ell=1}^\infty c_\ell a_\ell$$

with $\int_{\mathbb{R}^n} a_\ell(x)\,dx = 0$, $\|a_\ell\|_\infty \leqslant \ell^{-d}$, $\operatorname{supp}(a_\ell) \subset B(0,\ell)$ *for all* $\ell \geqslant 1$ *and*

$$\sum_{\ell=1}^\infty |c_\ell| \leqslant C(\eta)$$

with a constant $C(\eta)$ depending on η.

Proof. In this proof, we let

$$\langle g\rangle_S := \frac{1}{|S|}\int_S g(x)\,dx$$

for any $g \in L^1(\mathbb{R}^d)$ and $S \subset \mathbb{R}^d$ with $0 < |S| < \infty$. Moreover, $B_\ell := B(0,\ell)$ for $\ell \geqslant 1$, and $\chi_\ell = \chi_{B_\ell}$ (indicator of B_ℓ). Define

$$f_1 := (\eta - \langle\eta\rangle_{B_1})\chi_1\,, \quad \eta_1 := \eta - f_1$$

and set inductively

$$\begin{cases} f_{\ell+1} &:= (\eta_\ell - \langle\eta_\ell\rangle_{B_{\ell+1}})\,\chi_{\ell+1} \quad \text{and} \\ \eta_{\ell+1} &:= \eta_\ell - f_{\ell+1} \end{cases} \tag{10.22}$$

for $\ell \geqslant 1$ (one can take this also with $\ell = 0$ and $\eta_0 := \eta$). Observe that

$$\eta = \eta_1 + f_1 = f_1 + f_2 + \eta_2 = \ldots = \sum_{\ell=1}^{M} f_\ell + \eta_{M+1} \tag{10.23}$$

We need to show that we can pass to the limit $M \to \infty$ and that

$$a_\ell := f_\ell \cdot \frac{\ell^{-d}}{\|f_\ell\|_\infty} \quad \text{for} \quad \ell \geqslant 1 \tag{10.24}$$

have the desired properties. By construction, for all $\ell \geqslant 1$,

$$\int_{\mathbb{R}^d} a_\ell(x)\,dx = 0\,, \quad \operatorname{supp}(a_\ell) \subset B_\ell$$

and $\|a_\ell\|_\infty \leqslant \ell^{-d}$. It remains to show that

$$c_\ell := \ell^d \|f_\ell\|_\infty \tag{10.25}$$

satisfies $\sum_{\ell=1}^\infty c_\ell < \infty$ and that $\|\eta_{M+1}\|_\infty \to 0$, see (10.23). Clearly,

$$\eta_1 = \begin{cases} \langle\eta\rangle_{B_1} & \text{on} \quad B_1 \\ \eta & \text{on} \quad \mathbb{R}^d \backslash B_1 \end{cases}$$

Hence,

$$\eta_2 = \begin{cases} \langle\eta_1\rangle_{B_2} & \text{on} \quad B_2 \\ \eta_1 = \eta & \text{on} \quad \mathbb{R}^d \backslash B_2 \end{cases}$$

By induction, for $\ell \geqslant 1$, one checks that

$$\eta_{\ell+1} = \begin{cases} \langle\eta_\ell\rangle_{B_{\ell+1}} & \text{on} \quad B_{\ell+1} \\ \eta & \text{on} \quad \mathbb{R}^d \backslash B_{\ell+1}. \end{cases} \tag{10.26}$$

Moreover, induction shows that

$$\int_{\mathbb{R}^d} \eta_\ell(x)\,dx = 0 \tag{10.27}$$

for all $\ell \geqslant 0$. Indeed, this is assumed for $\eta_0 = \eta$. Since $\int f_\ell(x)\,dx = 0$ for $\ell \geqslant 1$ by construction, one now proceeds inductively via the formula

$$\eta_{\ell+1} = \eta_\ell - f_{\ell+1}$$

Property (10.27) implies that

$$\begin{aligned} \left|\langle \eta_\ell \rangle_{B_{\ell+1}}\right| &\leqslant \frac{1}{|B_{\ell+1}|} \int_{\mathbb{R}^d \setminus B_{\ell+1}} |\eta_\ell(x)| \; dx \\ &\leqslant \frac{1}{|B_{\ell+1}|} \int_{\mathbb{R}^d \setminus B_{\ell+1}} |\eta(x)| \, dx \leqslant C\ell^{-20d} \end{aligned} \tag{10.28}$$

since $\eta_\ell = \eta$ on $\mathbb{R}^d \setminus B_\ell$, see (10.26), and since η has rapid decay. (10.28) implies that

$$\|\eta_{\ell+1}\|_\infty \leqslant C\ell^{-20d},$$

see (10.26), and also

$$\|f_{\ell+1}\|_\infty \leqslant C\ell^{-20d}$$

see (10.22). We now conclude from (10.23)–(10.25) that

$$\eta = \sum_{\ell=1}^{\infty} f_\ell = \sum_{\ell=1}^{\infty} c_\ell a_\ell$$

with, see (10.25) and (10.24),

$$\sum_{\ell=1}^{\infty} |c_\ell| = \sum_{\ell=1}^{\infty} \ell^d \|f_\ell\|_\infty \leqslant \sum_{\ell=1}^{\infty} C\ell^{-19d} < \infty$$

and the lemma follows. □

Proof of Proposition 8. Let η be as in the statement of the proposition. By Lemma 10.13,

$$\eta = \sum_{\ell=1}^{\infty} c_\ell a_\ell$$

with c_ℓ and a_ℓ as stated there. By Lemma 10.12,

$$\|T(c_\ell a_\ell)\|_1 \leqslant |c_\ell| \|Ta_\ell\|_1 \leqslant CB|c_\ell|$$

for $\ell \geqslant 1$. Hence,

$$\sum_{\ell=1}^{\infty} \|T(c_\ell a_\ell)\|_1 \leqslant C_\eta B$$

and this easily implies that $\|T\eta\|_1 \leqslant C_\eta B$, is claimed. □

A typical application of Theorem 10.11 is to the so-called "Schauder estimate".

Corollary 10.14. *Let $f \in C_0^{2,\alpha}(\mathbb{R}^d)$. Then*

$$\sup_{1 \leqslant i,j \leqslant d} \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]_\alpha \leqslant C(\alpha, d)[\triangle f]_\alpha \tag{10.29}$$

for any $0 < \alpha < 1$.

Proof. As in the L^p case, see Corollary 7.7, this follows from the fact that

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = R_{ij}(\triangle f) \text{ for } 1 \leqslant i, j \leqslant d$$

where R_{ij} are the double Riesz transforms. Now apply Theorem 10.11 to the singular integral operators R_{ij}. □

For an extension of this corollary to variable coefficients, see Problem 10.6 below.

Exercise 10.6. Show that Theorem 10.11 fails at $\alpha = 0$ and $\alpha = 1$.

As a final application of Lemma 10.9 we remark that it easily implies Sobolev's imbedding $\dot{W}^{1,p}(\mathbb{R}^d) \hookrightarrow C^\alpha(\mathbb{R}^d)$ in the range $d < p < \infty$ with $\alpha = 1 - \frac{d}{p}$. See Problem 10.12.

Notes

The Littlewood-Paley theorem can be proved in a number of different ways, see for example Stein [**47**], which also contains a martingale difference version of this theorem. For many applications of Littlewood-Paley theory (such as to Besov spaces) and connections with wavelet theory, see Frazier, Jawerth and Weiss [**19**].

Concerning (10.11), it is a relatively easy consequence of the L^p-boundedness of the Hilbert transform that the answer is "yes" in one dimension. On the other hand, it is a very remarkable result of Charles Fefferman that the answer is "no" for dimensions $d \geqslant 2$. In fact, Fefferman [**14**] showed that the ball multiplier χ_B where B is any ball in $\mathbb{R}^d$ is bounded on $L^p(\mathbb{R}^d)$ only for $p = 2$. This latter result is based on the existence of Kakeya sets and will shall return to it in a later chapter. The positive answer for $d = 1$ is presented as Problem 10.4 below.

The "atoms" of Lemma 10.12 are exactly what one calls an H^1-*atom* where $H^1(\mathbb{R}^d)$ is the real-variable Hardy space. Lemma 10.13 is an instance of an *atomic decomposition* which exists for any $f \in H^1$, see Stein [**46**] and Koosis [**34**].

Proposition 10.8 is from [**55**].

Problems

Problem 10.1. The conditions in Theorem 10.1 can be relaxed. Indeed, it suffices to assume the following:

$$\sup_{R>0} R^{2|\alpha|} \fint_{R<|\xi|<2R} |\partial_\xi^\alpha m(\xi)|^2 \, d\xi \leqslant A^2 < \infty$$

for all $|\alpha| \leqslant k$ where k is the smallest integer $> \frac{d}{2}$.

Hint: Verify the Hörmander condition *ii*) of Definition 7.1 directly rather than going through Lemma 7.2.

Problem 10.2. Let T by a bounded linear operator on $L^p(X,\mu)$ where (X,μ) is some measure space and $1 \leqslant p < \infty$. Show the following vector-valued L^p estimate:

$$\big\| \{Tf_j\}_{\ell^2} \big\|_p \leqslant C(p) \|T\|_{p\to p} \big\| \{f_j\}_{\ell^2} \big\|_p$$

for any sequence $\{f_j\}$ of measurable functions for which the right-hand side is finite. Such vector-valued extensions are very useful. A challenging question is how to extend this property to sub-linear operators such as the Hardy-Littlewood maximal function, see [**46**], page 51.

Problem 10.3. Let $1 \leqslant p < 2$. By means of Khintchin's inequality show that for every $\varepsilon > 0$ there exists $f \in \mathcal{S}(\mathbb{R}^d)$ such that $\|\hat{f}\|_{p'} \leqslant \varepsilon \|f\|_p$, cf. Corollary 6.8.

Problem 10.4. In this problem the dimension is $d = 1$.

a) Deduce (10.11) from Theorem 10.4 by means of the following "vector-valued" inequality: Let $\{I_j\}_{j\in\mathbb{Z}}$ be an arbitrary collection of intervals. Then for any $1 < p < \infty$

$$\begin{aligned}\big\|\{(\chi_{I_j}\hat{f}_j)^\vee\}_j\big\|_{L^p(\ell^2)} &:= \Big\|\Big(\sum_j |(\chi_{I_j}\hat{f}_j)^\vee|^2\Big)^{\frac{1}{2}}\Big\|_p \\ &\leqslant C_p \|\{f_j\}\|_{L^p(\ell^2)}\end{aligned} \tag{10.30}$$

where $\{f_j\}_{j\in\mathbb{Z}}$ are an arbitrary collection of functions in $\mathcal{S}(\mathbb{R}^1)$, say.

b) Now deduce (10.30) from Theorem 4.8 by means of Khinchin's inequality *Hint:* express the operator $f \mapsto (\chi_I \hat{f})^\vee$ by means of the Hilbert transform via the same procedure which was used in the proof of Theorem 4.11.

c) By similar means prove the following Littlewood-Paley theorem for functions in $L^p(\mathbb{T})$: For any $f \in L^1(\mathbb{T})$ let

$$S f = \Big(\sum_{j=0}^{\infty} |P_j f|^2\Big)^{\frac{1}{2}}$$

where

$$(P_j f)(\theta) = \sum_{2^{j-1} \leqslant |n| < 2^j} \hat{f}(n) e(n\theta)$$

for $j \geqslant 1$ and $\triangle_0 f(\theta) = \hat{f}(0)$. Show that, for any $1 < p < \infty$

$$C_p^{-1}\|f\|_{L^p(\mathbb{T})} \leqslant \|S f\|_{L^p(\mathbb{T})} \leqslant C_p \|f\|_{L^p(\mathbb{T})} \tag{10.31}$$

for all $f \in L^p(\mathbb{T})$.

d) As a consequence of (10.11) with $d = 1$ show the following multiplier theorem: Let $m : \mathbb{R}\backslash\{0\} \to \mathbb{C}$ have the property that, for each $j \in \mathbb{Z}$

$$m(\xi) = m_j = \text{ constant}$$

for all $2^{j-1} \leqslant |\xi| < 2^j$. Then, for any $1 < p < \infty$

$$\|(m\hat{f})^\vee\|_{L^p(\mathbb{R})} \leqslant C_p \sup_{j\in\mathbb{Z}} |m_j| \|f\|_p$$

for all $f \in \mathcal{S}(\mathbb{R})$. Prove a similar theorem for $L^p(\mathbb{T})$ using (10.31).

Problem 10.5. Prove the following stronger version of Problem 7.4. By a dyadic interval we mean an interval of the form $\pm[2^k, 2^{k+1})$ where $k \in \mathbb{Z}$. Suppose m is a bounded function on $\mathbb{R}$ which is locally of bounded variation. In fact, assume that for some finite B one has $\|m\|_\infty \leqslant B$ and

$$\int_I |dm| \leqslant B$$

for all dyadic intervals I. Show that m is bounded on $L^p(\mathbb{R})$ as a Fourier multiplier for any $1 < p < \infty$ with $\|(m\hat{f})^\vee\|_p \leqslant C(p)\, B\|f\|_p$. For an extension to higher dimensions see [**45**].

Problem 10.6. We extend Corollary 10.14 to estimates for elliptic equations on a region $\Omega \subset \mathbb{R}^d$ with variable coefficients, i.e.,

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) = f(x) \quad \text{in} \quad \Omega$$

where $\sum_{i,j} a_{ij}\xi^i\xi^j \geqslant \lambda|\xi|^2$ and $a_{ij} \in C^\alpha(\Omega)$. By "freezing x", prove the following apriori estimate from (10.29) for any $f \in C^{2,\alpha}(\Omega)$:

$$\sup_{1\leqslant i,j\leqslant d} \Big[\frac{\partial^2 f}{\partial x_i \partial x_j}\Big]_{C^\alpha(K)} \leqslant C(\alpha, d, K)\big([f]_{C^\alpha(\Omega)} + \|f\|_{L^\infty(\Omega)}\big)$$

for any compact $K \subset \Omega$. *Hint:* See Gilbarg-Trudinger [**22**] for this estimate and much more.

Problem 10.7. Under the same assumptions as in Theorem 10.1 one has

$$[(m\hat{f})^\vee]_\alpha \leqslant C_\alpha B\,[f]_\alpha$$

for any $0 < \alpha < 1$ and $f \in C^\alpha(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. *Hint:* This is a corollary of the proof of Theorem 10.1. Analyze which properties of K it establishes and what is needed in order for the proof of Theorem 10.11 to go through.

Problem 10.8. For this problem we need to recall the definition of a martingale difference sequence: let $\{\mathcal{F}_n\}_{n=1}^\infty$ be an increasing sequence of σ-algebras on some probability space $(\Omega, \mathbb{P})$ (this is called a filtration). Assume that the function $X_n : \Omega \to \mathbb{R}$ is measurable with respect to $\mathcal{F}_n$ and $\mathbb{E}\,[X_n|\mathcal{F}_{n-1}] = 0$. Then $\{X_n\}_{n=1}^\infty$ is called a martingale difference sequence. Suppose $\{X_n\}_{n=1}^\infty$ is such a martingale difference sequence adopted to some filtration $\{\mathcal{F}_n\}_{n=1}^\infty$. Show that

$$\mathbb{P}\Big[\big|\sum_{n=1}^{N} X_n\big| > \lambda\big(\sum_{n=1}^{N} \|X_n\|_\infty^2\big)^{\frac{1}{2}}\Big] \leqslant Ce^{-c\lambda^2}$$

for any $N = 1, 2, \ldots$, and $\lambda > 0$. Here $C, c > 0$ are absolute constants.

Problem 10.9. Let $\{r_j\}_1^\infty$ be a coin tossing sequence as before. Show that for any $\{a_j\} \in \mathbb{C}$, and $N \in \mathbb{Z}^+$

$$\mathbb{P}\Big[\sup_{0\leqslant\theta\leqslant 1} \big|\sum_{j=1}^{N} r_j\, a_j e^{2\pi i j\theta}\big| > C_0\big(\sum_{j=1}^{N} |a_j|^2\big)^{\frac{1}{2}} \sqrt{\log N}\Big] \leqslant C_0 N^{-2}$$

provided C_0 is a sufficiently large absolute constant.

Problem 10.10. Using the ideas of this chapter as well as the proof strategy of Theorem 8.6, try to prove (8.10) (but with an arbitrary multiplicative constant). For which p does this approach succeed?

Problem 10.11. This problem introduces Sobolev and Besov spaces and studies some embeddings. We remark that the spaces with dots are called "homogeneous", whereas those without are called "inhomogeneous".

- Let $1 \leqslant p < \infty$ and $s \geqslant 0$. Define the Sobolev spaces $\dot{W}^{s,p}(\mathbb{R}^d)$ and $W^{s,p}(\mathbb{R}^d)$ as completions of $\mathcal{S}(\mathbb{R}^d)$ under the norms
$$\|(-\Delta)^{\frac{s}{2}} f\|_p \ \text{ respectively } \|\langle\Delta\rangle^{\frac{s}{2}} f\|_p$$
where $\langle a\rangle = \sqrt{1+a^2}$. The operators here are interpreted as Fourier multipliers. Show that for any $1 < p < \infty$
$$\|f\|_{\dot{W}^{s,p}} \simeq \Big\| \Big(\sum_{j\in\mathbb{Z}} 2^{2js} |P_j f|^2 \Big)^{\frac{1}{2}} \Big\|_p$$
with implicit constants that only depend on s, p, d. Formulate and prove an analogous statement for $W^{s,p}(\mathbb{R}^d)$. Does anything change for $s < 0$? Note that $s < 0$ naturally arises by the duality relation $(\dot{W}^{s,p})^* = \dot{W}^{-s,p}$. For which s is this valid? What is the analogue for the inhomogeneous spaces?
- Define the Besov spaces $\dot{B}^s_{p,q}(\mathbb{R}^d)$ and $B^s_{p,q}(\mathbb{R}^d)$ by means of the norms (for q finite)
$$\Big(\sum_{j\in\mathbb{Z}} 2^{qjs} \|P_j f\|_p^q \Big)^{\frac{1}{q}} \ \text{ respectively } \Big(\sum_{j\geqslant 0} 2^{qjs} \|P_j f\|_p^q \Big)^{\frac{1}{q}}$$
where in the second sum P_0 is interpreted as the projection onto all frequencies $\leqslant C\,1$. By means of the results of this chapter obtain embedding theorems between Besov and Sobolev spaces.
- Prove the embedding $\dot{B}^\sigma_{q,2}(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$ for $2 \leqslant q \leqslant p < \infty$, $\sigma \geqslant 0$, and $\frac{1}{q} - \frac{1}{p} = \frac{\sigma}{d}$. Does the same hold for the inhomogeneous spaces? What if anything can one change in that case? By means of the previous item, deduce embedding theorems for the Sobolev spaces from this.

Problem 10.12. Prove that $[f]_\alpha \leqslant C(\alpha, d)\|f\|_{\dot{W}^{1,p}(\mathbb{R}^d)}$ with $d < p < \infty$, $\alpha = 1 - \frac{d}{p}$ and $f \in \mathcal{S}(\mathbb{R}^d)$. This is called *Morrey's estimate*. For the case $p = d$ see Problem 9.7.

Problem 10.13. Show that $\dot{B}^{\frac{d}{2}}_{2,1}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ but $\dot{B}^{\frac{d}{2}}_{2,q}(\mathbb{R}^d) \not\hookrightarrow L^\infty(\mathbb{R}^d)$ for any $1 < q \leqslant \infty$. Show the embedding $\dot{B}^{\frac{d}{2}}_{2,\infty}(\mathbb{R}^d) \hookrightarrow \mathrm{BMO}(\mathbb{R}^d)$. Conclude that if $\|f\|_{\dot{B}^{\frac{d}{2}}_{2,q}(\mathbb{R}^d)} \leqslant A$ for some $1 \leqslant q < \infty$, then $w := e^f$ satisfies
$$⨍_Q w(x)\,dx ⨍_Q w^{-1}(x)\,dx \leqslant C(A, d, q)$$
uniformly for all cubes $Q \subset \mathbb{R}^d$.

CHAPTER 11

Fourier Restriction and Applications

In the late 1960's, Elias M. Stein posed the following question: is it possible to restrict the Fourier transform $\hat{f}$ of a function $f \in (\mathbb{R}^d)$ with $1 \leqslant p \leqslant 2$ to the sphere S^{d-1} as a function in $L^q(S^{d-1})$ for some $1 \leqslant q \leqslant \infty$? By the uniform boundedness principle this is the same as asking about the validity of the inequality

$$\|\hat{f} \restriction S^{d-1}\|_{L^q(S^{d-1})} \leqslant C\|f\|_{L^p(\mathbb{R}^d)} \tag{11.1}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$, with a constant $C = C(d, p, q)$? As an example, take $p = 1$, $q = \infty$ and $C = 1$. On the other hand, $p = 2$ is impossible, as $\hat{f}$ is no better than a general L^2-function by Plancherel. Stein asked whether it is possible to find $1 < p < 2$ so that for some finite q one has the estimate (11.1).

We have encountered an estimate of this type in Chapter 6, under the name of *trace estimate*. In fact, by Lemma 6.11 one has

$$\|\hat{f} \restriction S^{d-1}\|_\infty \leqslant C\, \|\langle x\rangle^\sigma f\|_2$$

as long as $\sigma > \frac{1}{2}$ (some work is required to pass from the planar case to the curved one). However, (11.1) is translation invariant and therefore much more useful for applications to nonlinear partial differential equations. In addition, as we shall see, the answers to (11.1) crucially involve *curvature*, whereas the trace lemmas do not distinguish between flat and curved surfaces.

Exercise 11.1. Suppose S is a bounded subset of a hyper-plane in $\mathbb{R}^d$. Prove that if $\|\hat{f} \restriction S\|_1 \leqslant C\|f\|_{L^p(\mathbb{R}^d)}$ for all $f \in \mathcal{S}(\mathbb{R}^d)$, then necessarily $p = 1$. In other words, there can be no nontrivial restriction theorem for flat surfaces.

The following Tomas-Stein theorem settles the important case $q = 2$ of Stein's question. For simplicity, we state it for the sphere. But it equally well applies to any bounded subset of any smooth hyper-surface in $\mathbb{R}^d$ with nonvanishing Gaussian curvature.

Theorem 11.1. *For every dimension $d \geqslant 2$ there is a constant $C(d)$ such that for all $f \in L^p(\mathbb{R}^d)$*

$$\|\hat{f} \restriction S^{d-1}\|_{L^2(S^{d-1})} \leqslant C(d)\|f\|_{L^p(\mathbb{R}^d)} \tag{11.2}$$

with $p \leqslant p_d := \frac{2d+2}{d+3}$. Moreover, this bound fails for $p > p_d$.

The left-hand side in (11.2) is

$$\Big(\int_{S^{d-1}} |\hat{f}(w)|^2\, \sigma(dw)\Big)^{\frac{1}{2}}$$

where σ is the surface measure on S^{d-1}.

We shall prove this theorem below, after making some initial remarks. Firstly, there is nothing special about the sphere. In fact, if S_0 is a compact subset of a hypersurface S with nonvanishing Gaussian curvature, then

$$\|\hat{f} \restriction S_0\|_{L^2(S_0)} \leqslant C(d, S_0)\|f\|_{L^{\frac{2d+2}{d+3}}(\mathbb{R}^d)} \tag{11.3}$$

for any $f \in \mathcal{S}(\mathbb{R}^d)$. For example, take the truncated paraboloid

$$S_0 := \{(\xi', |\xi'|^2) \mid \xi' \in \mathbb{R}^{d-1}, |\xi'| \leqslant 1\}$$

which is important for the Schrödinger equation. On the other hand, (11.3) fails for

$$S_0 := \{(\xi', |\xi'|) \mid \xi' \in \mathbb{R}^{d-1}, 1 \leqslant |\xi'| \leqslant 2\}$$

since this piece of the cone has exactly one vanishing principal curvature, namely the one along a generator of the one. This latter example is relevant for the wave equation as the characteristic surface of the wave equation is a cone, and so we shall need to find a substitute of (11.3) for the wave equation. A much simpler remark concerns the range $1 \leqslant p \leqslant p_d$ in Theorem 11.1: for $p = 1$, one has

$$\|\hat{f} \restriction S^{d-1}\|_{L^2(S^{d-1})} \leqslant \|\hat{f} \restriction S^{d-1}\|_\infty |S^{d-1}|^{\frac{1}{2}} \leqslant \|f\|_{L^1(R^d)}|S^{d-1}|^{\frac{1}{2}} . \tag{11.4}$$

Hence, it suffices to prove Theorem 11.1 for $p = p_d$ since the cases $1 \leqslant p < p_d$ follows by interpolation with (11.4). The Stein-Tomas theorem is more accessible than the general restriction problem with $q > 2$; in fact, we shall see that the appearance of $L^2(S^{d-1})$ allows one to use duality in the proof. To do so, we need to identify the adjoint of the restriction operator

$$R : f \mapsto \hat{f} \restriction S^{d-1}$$

Lemma 11.2. *For any finite measure μ in R^d, and any $f, g \in \mathcal{S}(\mathbb{R}^d)$ one has the identity*

$$\int_{R^d} \hat{f}(\xi)\overline{\hat{g}}(\xi)\,\mu(d\xi) = \int_{R^d} f(x)(\overline{g} * \hat{\mu})(x)\,dx$$

Proof. We use the following elementary identity for tempered distributions: If μ is a finite measure, and $\phi \in \mathcal{S}$, then

$$\widehat{\phi\mu} = \hat{\phi} * \hat{\mu}$$

Therefore (and occasionally using $\mathcal{F}(f)$ synonymously with $\hat{f}$ for notational reasons),

$$\begin{aligned}\int_{R^d} \hat{f}(\xi)\overline{\hat{g}}(\xi)\,\mu(d\xi) &= \int_{R^d} f(x)\mathcal{F}\left(\overline{\hat{g}}\mu\right)(x)\,dx = \\ &= \int_{R^d} f(x)(\widehat{\overline{\hat{g}}} * \hat{\mu})(x)\,dx = \int_{R^d} f(x)(\overline{g} * \hat{\mu})(x)\,dx\end{aligned}$$

since $\widehat{\overline{\hat{g}}} = \overline{\check{\hat{g}}} = \overline{g}$. ∎

Lemma 11.3. *Let μ be a finite measure on $\mathbb{R}^d$, and $d \geqslant 2$. Then the following are equivalent:*

a) $\|\widehat{f\mu}\|_{L^q(\mathbb{R}^d)} \leqslant C\|f\|_{L^2(\mu)}$ *for all* $f \in \mathcal{S}(\mathbb{R}^d)$
b) $\|\hat{g}\|_{L^2(\mu)} \leqslant C\|g\|_{L^{q'}(\mathbb{R}^d)}$ *for all* $g \in \mathcal{S}(\mathbb{R}^d)$
c) $\|\hat{\mu} * f\|_{L^q(\mathbb{R}^d)} \leqslant C^2\|f\|_{L^{q'}(\mathbb{R}^d)}$ *for all* $f \in \mathcal{S}(\mathbb{R}^d)$.

Proof. By the previous lemma, for any $g \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}\|\hat{g}\|_{L^2(\mu)} &= \sup_{f\in\mathcal{S},\, \|f\|_{L^2(\mu)}=1} \Big| \int_{\mathbb{R}^d} \hat{g}(\xi) f(\xi)\mu(d\xi)\Big| \\ &= \sup_{f\in\mathcal{S},\, \|f\|_{L^2(\mu)}=1} \Big| \int_{\mathbb{R}^d} g(x)\widehat{f\mu}(x)\, dx\Big|\end{aligned} \tag{11.5}$$

Hence, if *a*) holds, then the right-hand side of (11.5) is no larger than $\|g\|_{L^{q'}(\mathbb{R}^d)}$ and *b*) follows. Conversely, if *b*) holds, then the entire expression in (11.5) is no larger than $C\|g\|_{L^{q'}(\mathbb{R}^d)}$, which implies *a*). Thus *a*) and *b*) are equivalent with the same choice of C. Clearly, applying first *b*) and then *a*) with $f = \hat{g}$ yields

$$\|\mathcal{F}(\hat{g}\mu)\|_{L^q(\mathbb{R}^d)} \leqslant C\|g\|_{L^{q'}(\mathbb{R}^d)}$$

for all $g \in \mathcal{S}(\mathbb{R}^d)$. Since $\mathcal{F}(\hat{g}\mu) = g(-\cdot) * \hat{\mu}$, part *c*) follows. Finally, we note the relation

$$\int g(x)(\hat{\mu} * f)(x)\, dx = \int g(x)\mathcal{F}(\mu\check{f})(x)\, dx = \int \hat{g}(\xi)\check{f}(\xi)\,\mu(d\xi)$$

for any $f, g \in \mathcal{S}(\mathbb{R}^d)$. Hence, if c) holds then

$$\Big| \int_{\mathbb{R}^d} \hat{g}(\xi)\check{f}(\xi)\,\mu(d\xi)\Big| \leqslant C^2\|g\|_{L^{q'}(\mathbb{R}^d)}\|f\|_{L^{q'}(\mathbb{R}^d)}$$

Now set $f(x) = g(-x)$. Then

$$\int_{\mathbb{R}^d} |\hat{g}(\xi)|^2\,\mu(d\xi) \leqslant C^2\|g\|^2_{L^{q'}(\mathbb{R}^d)}$$

which is *b*). □

Setting $\mu = \sigma = \sigma_{S^{d-1}}$, the surface measure of the unit sphere in $\mathbb{R}^d$, one now obtains the following:

Corollary 11.4. *The following assertions are equivalent*

a) The Stein-Tomas theorem in the "restriction form":

$$\|\hat{f} \restriction S^{d-1}\|_{L^2(\sigma)} \leqslant C\|f\|_{L^{q'}(\mathbb{R}^d)}$$

for $q' = \frac{2d+2}{d+3}$ *and all* $f \in \mathcal{S}(\mathbb{R}^d)$

b) the "extension form" of the Stein-Tomas theorem

$$\|\widehat{g\sigma_{S^{d-1}}}\|_{L^q(\mathbb{R}^d)} \leqslant C\|g\|_{L^2(\sigma)}$$

for $q = \frac{2d+2}{d-1}$ *and all* $g \in \mathcal{S}(\mathbb{R}^d)$.

c) The composition of a) and b): for all $f \in \mathcal{S}(\mathbb{R}^d)$

$$\|f * \widehat{\sigma_{S^{d-1}}}\|_{L^q(\mathbb{R}^d)} \leqslant C^2 \|f\|_{L^{q'}(\mathbb{R}^d)}$$

with $q = \frac{2d+2}{d-1}$.

Proof. Set $\mu = \sigma_{S^{d-1}} = \sigma$ in Lemma 10.2. □

Exercise 11.2. In general, *a*) and *b*) above remain true, whereas *c*) requires $L^2(\sigma)$. More precisely, show the following: The restriction estimate

$$\|\hat{f} \restriction S^{d-1}\|_{L^p(\sigma)} \leqslant C\|f\|_{L^{q'}(\mathbb{R}^d)} \qquad \forall\, g \in \mathcal{S}$$

is equivalent to the extension estimate

$$\|\widehat{g\sigma_{S^{d-1}}}\|_{L^q(\mathbb{R}^d)} \leqslant C\|g\|_{L^{p'}(\sigma)} \qquad \forall\, g \in \mathcal{S}$$

We shall now show via part *b*) of Corollary 11.4 that the Stein-Tomas theorem is optimal. This is the well-known Knapp example.

Lemma 11.5. *The exponent $p_d = \frac{2d+2}{d+3}$ in Theorem 11.1 is optimal.*

Proof. This is equivalent to saying that the exponent $q = \frac{2d+2}{d-1}$ in part *b*) of the previous corollary is optimal. Fix a small $\delta > 0$ and let $g \in \mathcal{S}$ such that $g = 1$ on $B(e_d; \sqrt{\delta})$, $g \geqslant 0$, and $\operatorname{supp}(g) \subset B(e_n; 2\sqrt{\delta})$ where $e_d = (0, \ldots, 0, 1)$ is a unit vector.

Then

$$\begin{aligned}
|\widehat{g\sigma}(\xi)| &= \Big| \int e^{-2\pi i[x'\cdot\xi' + \xi_d(\sqrt{1-|x'|^2}-1)]} \frac{g(x', \sqrt{1-|x'|^2})}{\sqrt{1-|x'|^2}}\, dx' \Big| \\
&\geqslant \Big| \int \cos(2\pi(x'\cdot\xi' + \xi_d(\sqrt{1-|x'|^2}-1))) \frac{g(x', \sqrt{1-|x'|^2})}{\sqrt{1-|x'|^2}}\, dx' \Big| \\
&\geqslant \cos\frac{\pi}{4} \int g\, d\sigma \geqslant C^{-1}\delta^{\frac{d-1}{2}}
\end{aligned} \tag{11.6}$$

provided $|\xi'| \leqslant \frac{(\sqrt{\delta})^{-1}}{100}\ |\xi_d| \leqslant \frac{\delta^{-1}}{100}$. Indeed, under these assumptions, and for $\delta > 0$ small,

$$\big|x'\cdot\xi' + \xi_d\big(\sqrt{1-|x'|^2} - 1\big)\big| \leqslant \sqrt{\delta}\cdot\frac{(\sqrt{\delta})^{-1}}{100} + \frac{\delta^{-1}}{100}\Big(\sqrt{\delta}\Big)^2 \leqslant \frac{1}{50}\,,$$

so that the argument of the cosine in (11.6) is smaller than $\frac{2\pi}{50} \leqslant \frac{\pi}{4}$ in absolute value, as claimed.

Hence,

$$\begin{aligned}
\|\widehat{g\sigma_{S^{d-1}}}\|_{L^q(\mathbb{R}^d)} &\geqslant C^{-1}\delta^{\frac{(d-1)}{2}} \cdot \Big(\delta^{-\frac{(d-1)}{2}}\cdot\delta^{-1}\Big)^{\frac{1}{q}} \\
&= C^{-1}\delta^{\frac{d-1}{2}}\delta^{-\frac{(d+1)}{2q}}
\end{aligned}$$

whereas $\|g\|_{L^2(\sigma)} \leqslant C\delta^{\frac{d-1}{4}}$. It is therefore necessary that

$$\frac{d-1}{4} \leqslant \frac{d-1}{2} - \frac{d+1}{2q}$$

or $q \geqslant \frac{2d+2}{d-1}$, as claimed. □

For the proof of Theorem 11.1 we need the following decay estimate for the Fourier transform of the surface measure $\sigma_{S^{d-1}}$, see Corollary 6.16:

$$\left|\widehat{\sigma_{S^{d-1}}}(\xi)\right| \leqslant C(1+|\xi|)^{-(\frac{d-1}{2})} \tag{11.7}$$

It is easy to see that (11.7) imposes a restriction on the possible exponents for an extension theorem of the form

$$\|\widehat{f\sigma_{S^{d-1}}}\|_{L^q(\mathbb{R}^d)} \leqslant C\|f\|_{L^p(S^{d-1})} \tag{11.8}$$

Indeed, setting $f = 1$ implies that one needs

$$q > \frac{2d}{d-1}$$

by (11.7). On the other hand, one has

EXERCISE 11.3. Check by means of Knapp's example from Lemma 10.3 that (11.8) can only hold for

$$q \geqslant \frac{d+1}{d-1}p' \tag{11.9}$$

The still unproved (in dimensions $d \geqslant 3$) *restriction conjecture* states that (11.8) holds under these conditions, i.e., provided

$$\infty \geqslant q > \frac{2d}{d-1} \;\text{ and }\; q \geqslant \frac{d+1}{d-1}p'$$

Observe that the Stein-Tomas theorem with $q = \frac{2d+2}{d-1}$ and $p = 2$ is a partial result in this direction. In dimension $d = 2$ the conjecture is true and can be proved by elementary means, see the following chapter.

It is known, see Wolff's notes, that the full restriction conjecture implies the *Kakeya conjecture* which states that Kakeya sets in dimension $d \geqslant 3$ have (Hausdorff) dimension d. This latter conjecture appears to be also very difficult.

PROOF OF THE STEIN-TOMAS THEOREM FOR $p < \frac{2d+2}{d+3}$. Let

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

for all $x \neq 0$ be the usual Littlewood-Paley partition of unity. By Lemma 10.2 and Corollary 11.4 it is necessary and sufficient to prove

$$\|f * \widehat{\sigma_{S^{d-1}}}\|_{L^{p'}(\mathbb{R}^d)} \leqslant C\|f\|_{L^p(\mathbb{R}^d)}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$. Firstly, let

$$\varphi(x) = 1 - \sum_{j\geqslant 0} \psi(2^{-j}x)$$

Clearly, $\varphi \in C_0^\infty(\mathbb{R}^d)$, and

$$1 = \varphi(x) + \sum_{j\geqslant 0} \psi(2^{-j}x) \quad \text{for all} \quad x \in \mathbb{R}^d$$

Now observe that $\varphi\,\widehat{\sigma_{S^{d-1}}} \in C_0^\infty$ so that

$$\|f * \varphi\,\widehat{\sigma_{S^{d-1}}}\|_{L^{p'}(\mathbb{R}^d)} \leqslant C\|f\|_{L^p(\mathbb{R}^d)} \tag{11.10}$$

with $C = \|\varphi\,\widehat{\sigma_{S^{d-1}}}\|_{L^r}$ where $1 + \frac{1}{p'} = \frac{1}{r} + \frac{1}{p}$, i.e., $\frac{2}{p'} = \frac{1}{r}$. It therefore remains to control

$$K_j := \psi(2^{-j}x)\,\widehat{\sigma_{S^{d-1}}}(x)$$

in the sense that we need to prove an estimate of the form

$$\|f * K_j\|_{L^{p'}(\mathbb{R}^d)} \leqslant C2^{-j\varepsilon}\|f\|_{L^p(\mathbb{R}^d)} \qquad \forall\ f \in \mathcal{S}(\mathbb{R}^d) \tag{11.11}$$

for all $j \geqslant 0$ and some small $\varepsilon > 0$. It is clear that the desired bound follows by summing (11.10) and (11.11) over $j \geqslant 0$. To prove (11.11) we interpolate a $2 \to 2$ and $1 \to \infty$ bound as follows:

$$\begin{aligned}\|f * K_j\|_{L^2} &= \|\hat{f}\|_{L^2}\|\hat{K}_j\|_{L^\infty} = \\ &= \|f\|_{L^2}\|2^{nj}\hat{\psi}(2^j\cdot) * \sigma_{S^{d-1}}\|_{L^\infty} \\ &\leqslant C\|f\|_{L^2}2^{dj}\cdot 2^{-j(d-1)} = C2^j\|f\|_{L^2}\end{aligned} \tag{11.12}$$

To pass to the estimate (11.12) one uses that

$$\sup_x \sigma_{S^{d-1}}(B(x,r)) \leqslant Cr^{d-1}$$

as well as the fact that $\hat{\psi}$ has rapidly decaying tails, which implies that the estimate is the same as for compactly supported $\hat{\psi}$.

On the other hand,

$$\|f * K_j\|_\infty \leqslant \|K_j\|_\infty\|f\|_1 \leqslant C2^{-j\frac{(d-1)}{2}}\|f\|_1 \tag{11.13}$$

since the size of K_j is controlled by (11.7). Interpolating (11.12) with (11.13) yields

$$\|f * K_j\|_{p'} \leqslant C2^{-j\frac{d-1}{2}\theta}2^{j(1-\theta)}\|f\|_p$$

where $\frac{1}{p'} = \frac{\theta}{\infty} + \frac{1-\theta}{2} = \frac{1-\theta}{2}$. We thus obtain (11.11) provided

$$\begin{aligned}0 <& \frac{d-1}{2}\theta - (1-\theta) = \frac{d+1}{2}\theta - 1 = \\ &= (d+1)\Big(\frac{1}{2} - \frac{1}{p'}\Big) - 1 = \frac{d-1}{2} - \frac{d+1}{p'}\end{aligned}$$

This is the same as $p' > \frac{2d+2}{d-1}$ or $p < \frac{2d+2}{d+3}$, as claimed. □

EXERCISE 11.4. Provide the details concerning the "rapidly decaying tails" in the previous proof.

The previous proof strategy of the Tomas-Stein theorem does not achieve the sharp exponent $p = \frac{2d+2}{d+3}$ because (11.11) leads to a divergent series with $\varepsilon = 0$. In order to achieve this endpoint exponent, one therefore has to avoid interpolating the operator bounds on each dyadic piece separately. The idea is basically to sum *first* and *then* to interpolate, rather than interpolating first and *then* summing. Of course, one needs to explain what it means to sum first: recall that the proof of the Riesz-Thorin interpolation theorem is based on the three lines theorem from complex analysis. The key idea in our context is to sum the dyadic pieces $T_j : f \mapsto f * K_j$ together with *complex weights* $w_j(z)$ in such a way that

$$T_z := \sum_{j \geqslant 0} w_j(z) T_j$$

converges on the strip $0 \leqslant \operatorname{Re} z \leqslant 1$ to an analytic operator-valued function with the property that

$$\begin{aligned} T_z &: L^1 \to L^\infty \text{ for } \operatorname{Re} z = 1 \text{ and} \\ T_z &: L^2 \to L^2 \text{ for } \operatorname{Re} z = 0 \end{aligned}$$

It then follows that $T_\theta = L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$ for $\frac{1}{p'} = \frac{1-\theta}{2}$. The "art" is then to choose the weights $w_j(z)$ so that T_θ at a specific θ is a prescribed operator such as convolution by $\widehat{\sigma_{S^{d-1}}}$ in our case.

While it is conceptually correct and helpful to describe complex interpolation as a method of summing divergent series, it is rarely implemented in this fashion. Rather, one tries to embed the desired operator under consideration into an analytic family from the start without ever breaking it up into dyadic pieces. There are certain standard ways of doing this which involve a little complex analysis and distribution theory (on the level of integration by parts and analytic continuation). We shall present such a standard approach now.

PROOF OF THE ENDPOINT FOR TOMAS-STEIN. We consider a surface of non-zero curvature which can be written locally as a graph: $\xi_d = h(\xi')$, $\xi' \in \mathbb{R}^{d-1}$. Define

$$M_z(\xi) = \frac{1}{\Gamma(z)} \big(\xi_d - h(\xi')\big)_+^{z-1} \chi_1(\xi')\chi_2\big(\xi_d - h(\xi')\big) \tag{11.14}$$

where $\chi_1 \in C_0^\infty(\mathbb{R}^{d-1})$, $\chi_2 \in C_0^\infty(\mathbb{R})$ are smooth cut-off-functions, Γ is the Gamma function, and $\operatorname{Re} z > 0$. Moreover, $(\cdot)_+$ refers to the positive part. We will show that

$$T_z f := (M_z \hat{f})^\vee \tag{11.15}$$

can be defined by means of analytic continuation to $\operatorname{Re} z \leqslant 0$. The main estimates are now

$$\|T_z\|_{2\to 2} \leqslant B(z) \text{ for } \operatorname{Re} z = 1 \tag{11.16}$$

$$\|T_z\|_{1\to\infty} \leqslant A(z) \quad \text{for } \operatorname{Re} z = -\frac{d-1}{2} \tag{11.17}$$

where $A(z)$, $B(z)$ grow now faster than $e^{C|z|^2}$ as $|\mathrm{Im}\, z| \to \infty$. Moreover, we will see that the singularity of $\big(\xi_d - h(\xi')\big)_+^{z-1}$ at $z = 0$ "cancels" out against the simple zero of $\frac{1}{\Gamma(z)}$ at $z = 0$ so as to produce

$$M_0(\xi) = \chi_1(\xi')\delta_0(\xi_d - h(\xi'))\, d\xi' \tag{11.18}$$

see (11.21); this means that $M_0(\xi)$ is proportional to surface measure on the graph. It then follows from Stein's complex interpolation theorem that

$$f \mapsto \widehat{M_0} * f$$

is bounded from $L^p \to L^{p'}$ where

$$\frac{1}{p'} = \frac{\theta}{\infty} + \frac{1-\theta}{2}, \quad 0 = -\theta\frac{d-1}{2} + 1 - \theta$$

which implies that

$$\frac{1}{p'} = \frac{d-1}{2d+2}$$

as desired. It remains to check (11.16)–(11.18). To do so, recall first that $\frac{1}{\Gamma(z)}$ is an entire function with simple zeros at $z = 0, -1, -2, \ldots$. It has the product representation

$$\frac{1}{\Gamma(z)} = ze^{\gamma z}\prod_{\nu=1}^{\infty}\Big(1 + \frac{z}{\nu}\Big)e^{-z/\nu}, \qquad z = x + iy$$

which converges everywhere in $\mathbb{C}$. Thus,

$$\begin{aligned}\left|\frac{1}{\Gamma(z)}\right|^2 &\leqslant |z|^2 e^{2\gamma x}\prod_{\nu=1}^{\infty}\left[\Big(1+\frac{x}{\nu}\Big)^2 + \frac{y^2}{\nu^2}\right]e^{-\frac{2}{\nu}x} \\ &\leqslant |z|^2 e^{2\gamma x}\prod_{\nu=1}^{\infty}\left[e^{\frac{2x}{\nu}+\frac{|z|^2}{\nu^2}}e^{\frac{-2x}{\nu}}\right] = |z|^2 e^{2\gamma x}e^{|z|^2\frac{\pi^2}{6}}\end{aligned} \tag{11.19}$$

In particular, if $\mathrm{Re}\, z = 1$, then for all $\xi \in \mathbb{R}^d$,

$$|M_z(\xi)| \leqslant (1+y^2)e^{2\gamma}e^{(1+y^2)\frac{\pi^2}{6}}\chi_1(\xi')\chi_2\big(\xi_d - h(\xi')\big) \leqslant Ce^{cy^2}$$

(11.16) holds with the stated bound on $B(z)$. We remark that the bound we just obtained is very wasteful. The true growth, as given by the Stirling formula, is of the form $|y|^{\frac{1}{2}}e^{\frac{\pi}{2}|y|}$ as $|y| \to \infty$ but this makes no difference in our particular case.

Now let $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Thus, for $\mathrm{Re}\, z > 0$,

$$\begin{aligned}&\int_{\mathbb{R}^d} M_z(\xi)\varphi(\xi)\, d\xi = \\ &= \frac{1}{\Gamma(z)}\int_{\mathbb{R}^{d-1}}\int_0^\infty \chi_2(t)\varphi(\xi', t + h(\xi'))t^{z-1}\, dt\, \chi_1(\xi')\, d\xi' \\ &= -\frac{1}{z\Gamma(z)}\int_{\mathbb{R}^{d-1}}\int_0^\infty \frac{d}{dt}\big[\chi_2(t)\varphi\big(\xi', t+h(\xi')\big)\big]\, t^z\, dt\, \chi_1(\xi')\, d\xi'\end{aligned} \tag{11.20}$$

Observe that the right-hand side is well-defined for $\operatorname{Re} z > -1$. Furthermore, at $z = 0$, using $z\Gamma(z)\big|_{z=0} = 1$

$$\int_{\mathbb{R}^d} M_0(\xi)\varphi(\xi)\,d\xi = \int_{\mathbb{R}^{d-1}} \chi_2(0)\varphi(\xi', h(\xi'))\chi_1(\xi')\,d\xi'$$

which shows that the analytic continuation of M_z to $z = 0$ is equal to (setting $\chi_2(0) = 1$)

$$M_0(\xi) = \chi_1(\xi')\,d\xi'\delta_0(\xi_d - h(\xi')) \tag{11.21}$$

Clearly, M_0 is proportional to surface measure on a piece of the surface

$$S = \left\{(\xi', h(\xi')) \mid \xi' \in \mathbb{R}^{d-1}\right\}$$

This is exactly what we want, since we need to bound $\widehat{\sigma_S} * f$.

Observe that (11.20) defines the analytic continuation to $\operatorname{Re} z > -1$. Integrating by parts again extends this to $\operatorname{Re} z > -2$ and so forth. Indeed, the right-hand side of

$$\int_{\mathbb{R}^d} M_z(\xi)\varphi(\xi)\,d\xi = \frac{(-1)^k}{z(z+1)\cdot\ldots\cdot(z+k-1)\Gamma(z)} \int_{\mathbb{R}^{d-1}} \chi_1(\xi') \int_0^\infty t^{z+k-1} \frac{d^k}{dt^k}\left[\chi_2(t)\varphi(\xi', \xi_d + k(\xi'))\right] dt\,d\xi'$$

is well-defined for all $\operatorname{Re} z > -k$. Next we prove (11.17) by means of an estimate on $\|\widehat{M_z}\|_\infty$. This requires the following preliminary calculation: let N be a positive integer such that $N > \operatorname{Re} z + 1 > 0$. Then we claim that

$$\Big|\int_0^\infty e^{-2\pi i t\tau} t^z \chi_2(t)\,dt\Big| \leqslant \frac{C_N(1+|z|)^N}{1+\operatorname{Re} z}(1+|\tau|)^{-\operatorname{Re} z-1} \tag{11.22}$$

To prove (11.22) we will distinguish large and small $t\tau$. Let $\psi \in C_0^\infty(\mathbb{R})$ be such that $\psi(t) = 1$ for $|t| \leqslant 1$ and $\psi(t) = 0$ for $|t| \geqslant 2$. Then, since $0 \leqslant \chi_2 \leqslant 1$, we have

$$\begin{aligned}\Big|\int_0^\infty e^{-2\pi i t\tau} t^z \psi(t\tau)\chi_2(t)\,dt\Big| &\leqslant \int_0^\infty t^{\operatorname{Re} z}\psi(t|\tau|)\,dt \\ &\leqslant |\tau|^{-\operatorname{Re} z-1}\int_0^2 t^{\operatorname{Re} z}\,dt \leqslant \frac{C}{\operatorname{Re} z+1}|\tau|^{-\operatorname{Re} z-1}\end{aligned} \tag{11.23}$$

If $|\tau| \leqslant 1$, then (11.23) is no larger than

$$\int_0^\infty t^{\operatorname{Re} z}\chi_2(t)\,dt \leqslant \frac{C}{\operatorname{Re} z+1}$$

Hence

$$(11.23) \leqslant \frac{C}{1+\operatorname{Re} z}(1+|\tau|)^{-\operatorname{Re} z-1} \tag{11.24}$$

in all cases. To treat the case $t\tau$ large, which implies that τ is large, we exploit cancellation in the phase. More precisely,

$$
\begin{aligned}
&\Big|\int_0^\infty e^{-2\pi it\tau}t^z(1-\psi(t\tau))\chi_2(t)\,dt\Big| \leqslant \Big(\frac{1}{2\pi|\tau|}\Big)^N\int_0^\infty \Big|\frac{d^N}{d\tau^N}\Big[t^z(1-\psi(t\tau))\chi_2(t)\Big]\Big|\,dt\\
&\leqslant C_N\Big(\frac{1}{2\pi|\tau|}\Big)^N\int_0^\infty dt\Big[|z(z-1)\cdot\ldots\cdot(z-N+1)|t^{\operatorname{Re}z-N}(1-\psi(t\tau))\chi_2(t)\\
&\qquad + t^{\operatorname{Re}z}|\psi^{(N)}(t\tau)|\tau^N\chi_2(t)+t^{\operatorname{Re}z}(1-\psi(t\tau))|\chi_2^{(N)}(t)|\Big]\\
&\leqslant C_N\Big(\frac{1}{2\pi|\tau|}\Big)^N|z(z-1)\cdot\ldots\cdot(z-N+1)|\int_{\frac{1}{\tau}}^\infty t^{\operatorname{Re}z-N}\,dt\\
&\qquad + \frac{C}{(2\pi|\tau|)^N}|\tau|^N|\tau|^{-\operatorname{Re}z-1}+C|\tau|^{-N}\int_0^1 t^{\operatorname{Re}z}|\chi_2^{(N)}(t)|\,dt\\
&\leqslant C_N\big[|z(z-1)\cdot\ldots\cdot(z-N+1)|+1\big]|\tau|^{-\operatorname{Re}z-1}+C_N|\tau|^{-N}
\end{aligned}\tag{11.25}
$$

Observe that the indefinite integrals here converge because of $\operatorname{Re}z-N<-1$. Moreover, the second term in (11.25) is $\leqslant|\tau|^{\operatorname{Re}z-1}$ by the same condition (recall that we are taking τ to be large). Hence (11.22) follows from(11.24) and (11.25). We now compute $\widehat{M_z}(x)$. Let k be a positive integer with $\operatorname{Re}z>-k$. Then $\widehat{M_z}(x)$ equals

$$
\begin{aligned}
&\int e^{-2\pi ix\cdot\xi}\frac{1}{\Gamma(z)}(\xi_d-h(\xi'))_+^{z-1}\chi_1(\xi')\chi_2(\xi_d-h(\xi'))\,d\xi'\,d\xi_d\\
&=\frac{1}{\Gamma(z)}\int_0^\infty e^{-2\pi ix_dt}\,t^{z-1}\chi_2(t)\,dt\int_{\mathbb{R}^{d-1}}e^{-2\pi i[x'\cdot\xi'+x_dh(\xi')]}\chi_1(\xi')\,d\xi'\\
&=\frac{(-1)^k}{\Gamma(z)z\cdot(z-1)\cdot\ldots\cdot(z-k+1)}\int_0^\infty\big(e^{-2\pi ix_dt}\chi_2(t)\big)^{(k)}t^{z+k-1}\,dt\cdot\\
&\qquad\cdot\int_{\mathbb{R}^{d-1}}e^{-2\pi i[x'\cdot\xi'+x_dh(\xi')]}\chi_1(\xi')\,d\xi'
\end{aligned}\tag{11.26}
$$

where the final expression is well-defined for $\operatorname{Re}z>-k$. Now suppose that $\operatorname{Re}z=-\frac{d-1}{2}$ and pick $k\in\mathbb{Z}^+$ such that $1-k\leqslant\operatorname{Re}z<-k+2$, i.e., $\frac{d+1}{2}\leqslant k<\frac{d+1}{2}+1$. Apply (11.22) with $z+k-1$ instead of z and with $N=2$. Then the first integral (11.26) is bounded by

$$
\begin{aligned}
&\Big|\int_0^\infty\big(e^{-2\pi ix_dt}\chi_2(t)\big)^{(k)}t^{z+k-1}\,dt\Big|\\
&\leqslant\frac{C(1+|z|)^2}{\operatorname{Re}z+k}(1+|x_d|)^{-\operatorname{Re}z-k}\cdot C_k(1+|x_d|)^k\\
&\leqslant C_k\,(1+|z|)^2(1+|x_d|)^{-\operatorname{Re}z}
\end{aligned}\tag{11.27}
$$

On the other hand, the second integral in (11.26) is controlled by the stationary phase estimate, cf. (11.7),

$$
\Big|\int_{\mathbb{R}^{d-1}}e^{-2\pi i[x'\cdot\xi'+x_dh(\xi')]}\chi_1(\xi')\,d\xi'\Big|\leqslant C(1+|x|)^{-\frac{d-1}{2}}\tag{11.28}
$$

Observe that the growth in $|x_d|$ for $\operatorname{Re} z = -\frac{d-1}{2}$ is exactly balanced by the decay in (11.28). One concludes that for $\operatorname{Re} z = -\frac{(d-1)}{2}$ and with k as in (11.27)

$$\|\widehat{M_z}\|_\infty \leqslant C_d \left|\frac{1}{\Gamma(z) z \cdot (z-1) \cdot \ldots \cdot (z-k+1)}\right| (1+|z|)^2$$

see (11.26)–(11.28). Thus (11.17) follows from the growth estimate (11.19) or Stirling's formula, and we are done. □

The endpoint of the Tomas-Stein theorem is so important because of *its scaling invariance*. To see this, let $u(t,x)$ be a smooth solution of the Schrödinger equation

$$\begin{cases} \frac{1}{i}\partial_t u + \frac{1}{2\pi} \triangle_{\mathbb{R}^d} u &= F \\ \hfill u|_{t=0} &= f \end{cases} \tag{11.29}$$

where $f \in \mathcal{S}(\mathbb{R}^d)$, say. The constant $\frac{1}{2\pi}$ is rather arbitrary and chosen for cosmetic reasons having to do with our choice of normalization of the Fourier transform. It could be replaced by any other positive constant by rescaling; one can even change the sign of the constant by passing to complex conjugates. We first treat the case $F = 0$. Then

$$u(t,x) = \int_{\mathbb{R}^d} e^{2\pi i (x\cdot\xi + t|\xi|^2)} \hat{f}(\xi)\, d\xi = (\hat{f}\mu)^\vee(t,x) \tag{11.30}$$

where μ is the measure in $\mathbb{R}^{d+1}$ defined by the integration

$$\int_{\mathbb{R}^{d+1}} F(\xi,\tau)\, \mu(d(\xi,\tau)) = \int_{\mathbb{R}^d} F(\xi, |\xi|^2)\, d\xi$$

for all $F \in C^0(\mathbb{R}^{d+1})$. Now let $\varphi \in C_0^\infty(\mathbb{R}^{d+1})$, $\varphi(\xi,\tau) = 1$ if $|\xi| + |\tau| \leqslant 1$. Then the endpoint of Stein-Tomas applies and one concludes that

$$\|(\hat{f}\varphi\mu)^\vee\|_{L^q(\mathbb{R}^{d+1})} \leqslant C\|\hat{f}\|_{L^2(\varphi\mu)}$$

where $q = \frac{2d+4}{d} = 2 + \frac{4}{d}$. In other words, if $\operatorname{supp}(\hat{f}) \subset B(0,1)$, then

$$\|(\hat{f}\mu)^\vee\|_{L^{2+\frac{4}{d}}(\mathbb{R}^{d+1})} \leqslant C\|\hat{f}\|_{L^2(\mathbb{R}^d)} = C\|f\|_{L^2(\mathbb{R}^d)} \tag{11.31}$$

To remove the condition $\operatorname{supp}(\hat{f}) \subset B(0,1)$ we rescale. Indeed, the rescaled functions

$$\begin{aligned} f_\lambda(x) &= f(x/\lambda) \\ u_\lambda(x,t) &= u(x/\lambda, t/\lambda^2) \end{aligned} \tag{11.32}$$

satisfy the equation

$$\begin{cases} \frac{1}{i}\partial_t u_\lambda + \frac{1}{2\pi} \triangle u_\lambda &= 0 \\ \hfill u_\lambda|_{t=0} &= f_\lambda \end{cases}$$

If $\operatorname{supp}(\hat{f})$ is compact, then $\operatorname{supp}(\hat{f_\lambda}) \subset B(0,1)$ if λ is large. Hence, in view of (11.30) and (11.31),

$$\|u_\lambda\|_{L^q(\mathbb{R}^{d+1})} \leqslant C\|f_\lambda\|_{L^2(\mathbb{R}^d)} \tag{11.33}$$

However,

$$\|f_\lambda\|_{L^2(\mathbb{R}^d)} = \lambda^{\frac{d}{2}}\|f\|_{L^2(\mathbb{R}^d)}$$

and

$$\|u_\lambda\|_{L^q(\mathbb{R}^{d+1})} = \lambda^{\frac{d+2}{q}} \|u\|_{L^q} = \lambda^{\frac{d}{2}} \|u\|_{L^q}$$

Thus, (11.33) is the same as

$$\|u\|_{L^q(\mathbb{R}^{d+1})} \leqslant C\|f\|_{L^2(\mathbb{R}^d)} \tag{11.34}$$

for all $f \in \mathcal{S}$ will $\mathrm{supp}(\hat{f})$ compact. These functions are dense in $L^2(\mathbb{R}^d)$. Therefore, (11.34), which is the original *Strichartz bound* for the Schrödinger equation in d spatial dimensions, follows for all $f \in L^2(\mathbb{R}^d)$. Note the main difference between the Strichartz estimate (11.33) and the Stein-Tomas endpoint estimate: while the latter is confined to *compact* surfaces, the former applies to functions which live on the non-compact characteristic variety of the Schrödinger equation, i.e., the paraboloid $\{(\xi, |\xi|^2) \mid \xi \in \mathbb{R}^d\}$. Therefore, the Strichartz estimate necessarily obeys the scaling symmetry of that variety which is also all we needed to pass from compact surfaces to this *specific* non-compact surface.

While the approach we followed to derive the estimate (11.34) for solutions of (11.29) with $F = 0$ is the original one used by Strichartz and instructive due to its reliance on the restriction property of the Fourier transform, it is technically advantageous for many reasons to use a shorter and somewhat different argument which was introduced and developed by Ginibre and Velo in a series of papers. We will present this argument now to derive the following theorem. We call a pair (p, q) *Strichartz admissible* if and only if

$$\frac{2}{p} + \frac{d}{q} = \frac{d}{2} \tag{11.35}$$

and $2 \leqslant p \leqslant \infty$ with $(p, q) \neq (2, \infty)$.

Theorem 11.6. *Let F be a space-time Schwartz function in $d+1$ dimensions, and f a spatial Schwartz function. Let $u(t,x)$ solve* (11.29)*. Then*

$$\|u\|_{L_t^p L_x^q(\mathbb{R}^{d+1})} \leqslant C\big(\|f\|_{L^2(\mathbb{R}^d)} + \|F\|_{L_t^{a'} L_x^{b'}(\mathbb{R}^{d+1})}\big) \tag{11.36}$$

where (p, q) and (a, b) are Strichartz admissible with $a > 2$ and $p > 2$. Finally, these estimates localize in time: if u is restricted to some time-interval I on the left-hand side, then F can be restricted to I on the right-hand side.

Some remarks are in order: first, the estimate (11.36) is scaling invariant under the law (11.32) if and only if (p, q) and (a, b) obey the relation (11.35). Hence the latter is *necessary* for the Strichartz estimates (11.36) to hold. Note that the pair $p = q = 2 + \frac{4}{d}$ which we derived above for $F = 0$ is of this type. On the other hand, the restriction $p > 2$ is technical and the important endpoint estimate for $p = 2$ and $d \geqslant 3$ was proved by Keel and Tao. Finally, under the strong conditions on F and f it is easy to actually solve (11.29) by means of an explicit expression:

$$u(t) = e^{-\frac{i}{2\pi}\triangle t} f + \int_0^t e^{-\frac{i}{2\pi}\triangle(t-s)} F(s)\, ds \tag{11.37}$$

with the understanding that the operator needs to be understood as a Fourier multiplier:

$$\left(e^{-\frac{i}{2\pi}\triangle t} f\right)(x) := \int e^{2\pi i(x\cdot\xi+|\xi|^2 t)}\,\hat{f}(\xi)\,d\xi$$

which is well-defined for $f \in \mathcal{S}(\mathbb{R}^d)$, say.

PROOF OF THEOREM 11.6. We again start with $F = 0$ and let $U(t)$ denote the propagator, i.e., $U(t)f$ is the solution of (11.29) as give by (11.37): $U(t) = e^{-\frac{i}{2\pi}\triangle t}$. By (6.14) we have the bound

$$\|U(t)f\|_{L^q(\mathbb{R}^d)} \leqslant C|t|^{-\frac{d}{2}(\frac{1}{q'}-\frac{1}{q})}\|f\|_{L^{q'}(\mathbb{R}^d)} \tag{11.38}$$

for $1 \leqslant p \leqslant 2$. Indeed, for $p = 1$ this follows by placing absolute values inside (6.14), whereas $p = 2$ is Plancherel's theorem. The intermediate p values then follow by interpolation. Our goal for now is to prove that for any (p, q) as in the theorem

$$\|Uf\|_{L^p_t L^q_x} \leqslant C\|f\|_2$$

where $(Uf)(t, x) = (U(t)f)(x)$, slightly abusing notation. In analogy with the "duality equivalence" Lemma 11.3 one now has that each of the following estimate implies the other two:

$$\|Uf\|_{L^p_t L^q_x} \leqslant C\|f\|_{L^2_x}$$
$$\|U^*F\|_{L^2_x} \leqslant C\|F\|_{L^{p'}_t L^{q'}_x}$$
$$\|U \circ U^*\|_{L^p_t L^q_x} \leqslant C^2\,\|F\|_{L^{p'}_t L^{q'}_x}$$

As in the case of the Tomas-Stein theorem, the key is to prove the third property. Note that one has

$$U^*F = \int_{-\infty}^{\infty} U(-s)F(s)\,ds$$

whence by the group property of $U(t)$

$$(U \circ U^*F)(t) = \int_{-\infty}^{\infty} U(t-s)F(s)\,ds$$

By (11.38),

$$\|(U \circ U^*F)(t)\|_{L^q_x} \leqslant C\int_{-\infty}^{\infty} |t-s|^{-\frac{d}{2}(\frac{1}{q'}-\frac{1}{q})}\|F(s)\|_{L^{q'}_x}\,ds$$

whence by fractional integration in time with

$$1 + \frac{1}{p} = \frac{1}{p'} + \frac{d}{2}(\frac{1}{q'} - \frac{1}{q}) \Longleftrightarrow (11.35)$$

provided

$$0 < \frac{d}{2}(\frac{1}{q'} - \frac{1}{q}) < 1 \Longleftrightarrow p > 2 \tag{11.39}$$

one has

$$\|U \circ U^*F\|_{L^p_t L^q_x} \leqslant C\|F\|_{L^{p'}_t L^{q'}_x}$$

which is the desired estimate for the case $F = 0$. Now suppose $f = 0$, and write the solution of (11.29) as

$$u(t) = \int_0^t U(t-s)F(s)\,ds = \int_{-\infty}^{\infty} \chi_{[0<s<t]}\, U(t-s)F(s)\,ds$$

We now claim two estimates on u as a space-time function:

$$\|u\|_{L_t^p L_x^q} \leqslant C\|F\|_{L_t^{p'} L_x^{q'}}$$
$$\|u\|_{L_t^p L_x^q} \leqslant C\|F\|_{L_t^1 L_x^2}$$

The first one is proved by the same argument as before, whereas the second one is proved as follows:

$$\begin{aligned}\|u(t)\|_{L_x^q} &\leqslant \int_{-\infty}^{\infty} \chi_{[0<s<t]}\, \|U(t-s)F(s)\|_{L_x^q}\,ds \\ &\leqslant \int_{-\infty}^{\infty} \|U(t-s)F(s)\|_{L_x^q}\,ds\end{aligned}$$

whence

$$\begin{aligned}\|u\|_{L_t^p L_x^q} &\leqslant C\int_{-\infty}^{\infty} \|U(t-s)F(s)\|_{L_t^p L_x^q}\,ds \\ &\leqslant C\int_{-\infty}^{\infty} \|F(s)\|_{L_x^2}\,ds = \|F\|_{L_s^1 L_x^2}\end{aligned}$$

By duality and interpolation one now concludes that

$$\|u\|_{L_t^p L_x^q} \leqslant C\|F\|_{L_s^{a'} L_x^{b'}}$$

for any admissible pair (a, b) and (p, q) with $a > 2$ and $p > 2$. This concludes the argument. The statement about time-localizations follows from the solution formula (11.37). □

EXERCISE 11.5. The previous argument of course offers an alternative to the complex interpolation approach that we used to obtain the Stein-Tomas endpoint. Provide the details of this alternative proof to Theorem 11.1 for general bounded subsets of surfaces of nonvanishing Gaussian curvature.

As an application of Theorem 11.6 we now solve the nonlinear equation

$$i\partial_t\psi + \Delta\psi = \lambda|\psi|^{\frac{4}{d}}\psi \tag{11.40}$$

in d spatial dimensions and with an arbitrary real-valued constant λ for small L^2-data $\psi|_{t=0} = \psi_0$. First, it is not apriori clear what we mean by "solve" since the data are L^2, nor is it immediately clear why we do not consider smoother data. Second, we remark that the choice of power in (11.40) is not accidental. In fact, it is the unique power for which the rescaled functions

$$\lambda^{\frac{d}{2}}\psi(\lambda x, \lambda^2 t) \tag{11.41}$$

are again solutions. The relevance of this scaling law is the fact that the L^2-invariant scaling of the data is $\lambda^{\frac{d}{2}}\psi_0(\lambda x)$, which therefore leaves the smallness condition unchanged. Moreover, the induced rescaling of the solution is then given by (11.41). Hence, the only setting in which we can hope for a meaningful global small L^2-data theory is (11.40).

As in the existence and uniqueness theory of ordinary differential equations, one reformulates (11.40) as an integral equation as follows (this is called Duhamel formula):

$$\psi(t) = e^{it\Delta}\psi_0 + i\lambda \int_0^t e^{i(t-s)\Delta}\, |\psi(s)|^{\frac{4}{d}}\psi(s)\, ds \tag{11.42}$$

and then seeks a solution of this integral equation by contraction or Picard iteration belonging to the space $C_t([0,\infty); L_x^2(\mathbb{R}^d))$. This is very natural, as (11.40) conserves the L^2-norm for real λ (see the problem section). However, this space is still too large to formulate a meaningful theory in, and so we restrict it further in a way that allows us to invoke Theorem 11.6.

Definition 11.7. By a global weak solution of (11.40) with data $\psi_0 \in L^2(\mathbb{R}^d)$ we mean a solution to the integral equation (11.42) which lies in the space

$$X := (C_t \cap L_t^\infty)([0,\infty); L_x^2(\mathbb{R}^d)) \cap L_t^{p_0}([0,\infty); L_x^{2p_0}(\mathbb{R}^d)) \tag{11.43}$$

with $p_0 := 1 + \frac{4}{d}$.

We remark that the power p_0 in the definition is very natural. In fact, passing L^2-norms onto (11.42) yields

$$\|\psi(t)\|_2 \leqslant C\Big(\|\psi_0\|_2 + |\lambda| \int_0^t \|\psi(s)\|_{2p_0}^{p_0}\, ds\Big)$$

which is at least finite for $\psi \in X$. Now we will show much more.

Corollary 11.8. *The nonlinear equation* (11.40) *in* $\mathbb{R}_{t,x}^{1+d}$ *admits a unique weak solution in the sense of Definition 11.7 for any data* $\psi_0 \in L^2(\mathbb{R}^d)$ *provided* $\|\psi_0\|_2 \leqslant \varepsilon_0(\lambda, d)$.

Proof. For the sake of simplicity, we set $d = 2$ which implies that the nonlinearity takes the form $\lambda|\psi|^2\psi$ and $p_0 = 3$. We leave the case of general dimensions to the reader. Let us first prove uniqueness. Thus, let ψ and φ be two weak solutions to (11.42). Taking differences yields

$$(\psi - \varphi)(t) = i\lambda \int_0^t e^{i(t-s)\Delta}\, (|\psi(s)|^2\psi(s) - |\varphi(s)|^2\varphi(s))\, ds$$

Therefore, applying the Strichartz estimates of the previous theorem locally in time on $I = [0, T)$ yields, with $X(I)$ being as in (11.43) localized to I,

$$\begin{aligned}\|\psi - \varphi\|_{X(I)} &\leqslant C|\lambda| \||\psi(s)|^2\psi(s) - |\varphi(s)|^2\varphi(s)\|_{L^1([0,T);L^2(\mathbb{R}^2)} \\ &\leqslant C|\lambda|(\|\psi\|_{X(I)} + \|\varphi\|_{X(I)})^2\, \|\psi - \varphi\|_{X(I)}\end{aligned}$$

However, if I is small enough then $C|\lambda|(\|\psi\|_{X(I)} + \|\varphi\|_{X(I)})^2 < 1$ whence $\|\psi - \varphi\|_{X(I)} = 0$. Thus, $\psi = \varphi$ on I. This argument shows that the set where $\psi = \varphi$ is both open and closed and since it is nonempty (it contains $t = 0$) therefore equals the whole time-line.

To prove existence, we set up a contraction argument in the space X for small data. Thus, we define and operator A on X by the formula

$$(A\psi)(t) := e^{it\Delta}\psi_0 + i\lambda \int_0^t e^{i(t-s)\Delta}\, |\psi(s)|^2 \psi(s)\, ds$$

and note that by Theorem 11.6

$$\|A\psi\|_X \leqslant C(\|\psi_0\|_2 + |\lambda| \|\psi\|_X^3)$$

Therefore, if $\|\psi\|_X \leqslant R$, and if

$$C(\|\psi_0\|_2 + |\lambda| R^3) < R \tag{11.44}$$

we conclude that A takes and R-ball in X into itself. To satisfy (11.44) we simply choose ε so small that $R = C\varepsilon$ will verify (11.44) for any $\|\psi_0\|_2 < \varepsilon$. Taking differences as in the uniqueness proof then shows that A is in fact a contraction. Therefore, there exists a unique fixed-point ψ in the R-ball in X. But $A\psi = \psi$ is precisely the definition of a weak solution and we are done. $\square$

Corollary 11.8 is only one of many results which can be obtained in this area. It is also clear that many interesting questions pose themselves now, such as the question of global existence for large data (which depends on the sign of λ) which is relatively easy for data in H^1 but hard for data in L^2, persistence of regularity, spatial decay of solutions etc. We present some of the easier results in this direction in the problem section. Showing that an equation with a derivative nonlinearity such as

$$i\partial_t \psi + \Delta\psi = \pm |\psi|^2\, \partial_1 \psi \tag{11.45}$$

has global solutions for small data in $\mathcal{S}$, say, *cannot be accomplished* by means of Strichartz estimates alone. Indeed, the Strichartz estimates are not able to make up for the loss of a derivative on the right-hand side.

We conclude this chapter with a sketch of Strichartz estimates for the wave equation in $\mathbb{R}^{1+d}_{t,x}$:

$$\begin{aligned} \Box u &= \partial_{tt} u - \Delta u = F \\ u\big|_{t=0} &= f, \quad \partial_t u\big|_{t=0} = g \end{aligned} \tag{11.46}$$

The solution to this Cauchy problem is

$$u(t) = \cos(t|\nabla|) f + \frac{\sin(t|\nabla|)}{|\nabla|} g + \int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} F(s)\, ds$$

at least for Schwartz data, say. The meaning of the operators appearing here is of course based on the Fourier transform. In other words,

$$[\cos(t|\nabla|)f](x) = \sum_{\pm} \frac{1}{2} \int_{\mathbb{R}^d} e^{2\pi i(x\cdot\xi \pm t|\xi|)}\, \hat{f}(\xi)\, d\xi$$

Proceeding analogously to the Schrödinger equation were are therefore lead to interpreting these integrals as Fourier transforms of measures living on the double *light-cone*

$$\Gamma := \big\{(\xi,\tau) \in \mathbb{R}^{d+1} \mid |\tau| = |\xi|\big\}$$

This is different from the paraboloid in the Schrödinger equation in two important ways: first, Γ has one vanishing principal curvature namely along the generators of the light cone. Second, Γ is singular at the vertex of the cone. To overcome the latter difficulty, we define

$$\Gamma_0 := \{(\tau,\xi) \in \mathbb{R}^{1+d} \mid 1 \leqslant |\xi| = \tau \leqslant 2\}$$

to be a section of the light cone in $\mathbb{R}^{1+d}$. Since the cone has one vanishing principal curvature, the Fourier transform of the surface measure on Γ_0 exhibits less decay as for the paraboloid which reflects itself in a different Stein-Tomas exponent (in fact, the numerology here is such that the dimension d is reduced to $d-1$ reflecting the loss of one principal curvature).

Exercise 11.6.

- By means of the method of stationary phase from Chapter 6 show that
$$\left|\widehat{\varphi\sigma_{\Gamma_0}}(t,x)\right| \leqslant C(1+|(t,x)|)^{-\frac{(d-1)}{2}} \tag{11.47}$$
for all $(t,x) \in \mathbb{R}^{1+d}$. Moreover, this is optimal for all directions (t,x) belonging to the dual cone Γ^* (which is equal to Γ if the opening angle is 90°).
- Check that the complex interpolation method from above therefore implies that there is the following restriction estimate for Γ_0:
$$\|\hat{f} \restriction \Gamma_0\|_{L^2(\sigma_\Gamma)} \leqslant C\|f\|_{L^p(\mathbb{R}^d)}$$
where $p = \frac{2d+2}{d+3}$ and $d \geqslant 2$ (for $d = 1$ this amounts to a trivial bound).

Via a rescaling and Littlewood-Paley theory we now arrive at the following analogue of the argument leading to (11.34) for the Schrödinger equation.

Theorem 11.9. *Let Γ be the double light-cone in $\mathbb{R}^{1+d}_{\tau,\xi}$ with $d \geqslant 2$, equipped with the measure $\mu(d(\tau,\xi)) = \frac{d\xi}{|\xi|}$. Then*

$$\Big(\int_\Gamma |\hat{f}(\tau,\xi)|^2\, \mu(d(\tau,\xi))\Big)^{\frac{1}{2}} \leqslant C\|f\|_{L^p(\mathbb{R}^d)} \tag{11.48}$$

with $p = \frac{2d+2}{d+3}$.

PROOF. Let Γ_0 be the cone restricted to $1 \leqslant |\tau| \leqslant 2$ as above. Then

$$\begin{aligned}
&\Big(\int_{\lambda \leqslant |(\tau,\xi)| \leqslant 2\lambda} |\hat{f}((\tau,\xi))|^2 \, \mu(d(\tau,\xi)) \Big)^{\frac{1}{2}} \\
&\leqslant \Big(\int_{\Gamma_0} |\hat{f}(\lambda(\tau,\xi))|^2 \mu(d(\lambda(\tau,\xi))) \Big)^{\frac{1}{2}} \\
&\leqslant C\lambda^{\frac{d-1}{2}} \Big\| \frac{1}{\lambda^{d+1}} f\big(\frac{1}{\lambda}\,\cdot\,\big) \Big\|_{L^p(\mathbb{R}^{1+d})} \\
&= C\lambda^{\frac{d-1}{2}} \cdot \lambda^{-d-1} \cdot \lambda^{\frac{d+1}{p}} \|f\|_{L^p(\mathbb{R}^{d+1})} = C\|f\|_{L^p(\mathbb{R}^{d+1})}
\end{aligned} \tag{11.49}$$

by choice of p. To sum up (11.49) for $\lambda = 2^j$ we use the Littlewood-Paley theorem together with Exercise 10.3. In fact, since $p < 2$, we conclude that

$$\|\hat{f}\|_{L^2(\mu)} \leqslant C\Big(\sum_{j\in\mathbb{Z}} \|P_j f\|_{L^p}^2 \Big)^{\frac{1}{2}} \leqslant C\Big\| \Big(\sum_{j\in\mathbb{Z}} \big|P_j f\big|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \leqslant C\|f\|_{L^p(\mathbb{R}^{d+1})}$$

which is (11.48). □

For the wave equation this means the following.

COROLLARY 11.10. *Let $u(t,x)$ be a solution of* (11.46) *with $f, g \in \mathcal{S}(\mathbb{R}^d)$, $d \geqslant 2$. Then*

$$\|u\|_{L^{\frac{2d+2}{d-1}}(\mathbb{R}^{d+1})} \leqslant C(\|f\|_{\dot{H}^{\frac{1}{2}}(\mathbb{R}^d)} + \|g\|_{\dot{H}^{-\frac{1}{2}}(\mathbb{R}^d)})$$

with a constant $C = C(d)$.

PROOF. For simplicity, set $f = 0$. Then

$$\begin{aligned}
u(t,x) &= \int_{\mathbb{R}^d} e^{2\pi i(t|\xi| + x\xi)} \hat{g}(\xi) \frac{d\xi}{4\pi i|\xi|} + \int_{\mathbb{R}^d} e^{2\pi i(-t|\xi| + x\cdot\xi)} \hat{g}(\xi) \frac{d\xi}{4\pi i|\xi|} \\
&= (F\mu)^\vee(x,t)
\end{aligned}$$

where $F(\xi, \pm|\xi|) = \hat{f}(\xi)$ and $d\mu(\xi, \pm|\xi|) = \frac{d\xi}{|\xi|}$. By the dual to Theorem 11.9,

$$\|(F\,\mu)^\vee\|_{L^{p'}(\mathbb{R}^{d+1})} \leqslant C\|F\|_{F^2(\mu)} \tag{11.50}$$

where $p' = \frac{2d+2}{d-1}$. Clearly,

$$\|F\|_{L^2(\mu)} = \Big(\int_{\mathbb{R}^d} |\hat{g}(\xi)|^2 \frac{d\xi}{|\xi|} \Big)^{\frac{1}{2}} = \|g\|_{\dot{H}^{-\frac{1}{2}}}$$

so that (11.50) implies that

$$\|u\|_{L^{\frac{2d+2}{d-1}}(\mathbb{R}^{d+1})} \leqslant C\|g\|_{\dot{H}^{-\frac{1}{2}}(\mathbb{R}^d)}$$

as claimed. □

EXERCISE 11.7. Derive estimates for the wave equation as in the previous corollary but with $\|f\|_{\dot{H}^1} + \|g\|_{L^2}$ on the right-hand side.

Notes

A standard introductory reference in this area are Stein's Beijing lectures [**48**] and Wolff's course notes [**56**]. The original reference by Strichartz is [**50**]. For an account summarizing more recent developments see Tao [**53**]. Strichartz estimates are a vast area, and continue to be developed, especially in the variable coefficient setting, see Bahouri, Chemin, and Danchin [**3**] for that aspect. A standard references on dispersive evolution equations is Tao [**52**]. Keel, Tao [**31**] prove the endpoint for the Strichartz estimates, i.e., L^2_t-type estimates. For more on the Schrödinger equation see Cazenave [**8**], as well as Sulem and Sulem [**51**]. For the wave equation, see Shatah and Struwe [**42**]. Strichartz esimates for the Klein-Gordon equation, which behaves like the wave equation for large frequencies and the Schrödinger equation for small frequencies, respectively, are derived in [**37**]. Some of the original work by Ginibre and Velo on Strichartz estimates for various dispersive equations can be found here [**23**]. Finally, for the solution theory of equations of the type (11.45) which require *smoothing estimates for the Schrödinger equation*, see Kenig, Ponce and Vega [**32**].

Problems

Problem 11.1. Suppose that ϕ is a smooth function on $(-1,1)$ with $\phi'(t) = 0$ if and only if $t = 0$. Assume further that $\phi''(0) = 0$ and $\phi'''(0) \neq 0$. Let $a \in C^\infty(-1,1)$ with $\mathrm{supp}(a) \subset (-1,1)$. Determine the sharp decay of

$$\int_{-1}^{1} e^{i\lambda\phi(t)}\, a(t)\, dt$$

as $\lambda \to \infty$.

Problem 11.2. Investigate the restriction of the Fourier transform of a function in $\mathbb{R}^3$ to a (section of a) smooth curve in $\mathbb{R}^3$. Assume that the curve has nonvanishing curvature and torsion. Try to generalize to other dimensions and codimensions.

Problem 11.3. This problem expands on Corollary 11.8.

- Show that if $\psi_0 \in H^k(\mathbb{R}^d)$ (one can again take $d = 1$ or $d = 2$ for simplicity) where k is a positive integer, then $\psi \in C_t([0,\infty); H^k(\mathbb{R}^d))$. Conclude via Sobolev imbedding that if $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$, say, then $\psi(t,x)$ is smooth in all variables and satisfies the nonlinear equation (11.40) in a pointwise sense.
- Show that if $\psi(t,x)$ is any solution (not necessarily small) of (11.40) with $\partial_t\psi \in C([0,T); L^2(\mathbb{R}^d))$ and $\psi \in C([0,T); H^2(\mathbb{R}^d))$, then one has *conservation of mass* (recall λ is real)

$$M(\psi) := \frac{1}{2}\|\psi(t)\|_2^2 = \frac{1}{2}\|\psi_0\|_2^2 \qquad \forall\ 0 \leqslant t < T$$

and *energy*

$$E(\psi) := \int_{\mathbb{R}^d} \Big(\frac{1}{2}|\nabla\psi(t,x)|^2 - \frac{\lambda}{2+\frac{4}{d}}|\psi(t,x)|^{2+\frac{4}{d}}\Big)\, dx$$

Hint: differentiate M and E with respect to time, and integrate by parts.

- Prove the relation
$$e^{it\Delta}x = (x + 2i\nabla)e^{it\Delta}$$
as operators acting on Schwartz functions and use this to show that if $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$, then the solution constructed in Corollary 11.8 has the property that for any time t one has $\psi(t) \in \mathcal{S}(\mathbb{R}^d)$.
- Show that the solution of Corollary 11.8 with $\psi_0 \in L^2(\mathbb{R}^d)$ preserves the mass $M(\psi)$ and the energy provided the data satisfy $\psi_0 \in H^1(\mathbb{R}^d)$.

Problem 11.4.

- For any power $p \geqslant 2$ show that the equation
$$i\partial_t\psi + \partial_{xx}\psi = \lambda|\psi|^{p-1}\psi \tag{11.51}$$
on the line $\mathbb{R}$, has local in-time unique weak solutions for any data $\psi_0 \in H^1(\mathbb{R})$. These solutions again need to be interpreted in the Duhamel integral equation sense, and the space in which to contract is $C([0,T); H^1(\mathbb{R}))$ where $T > 0$ can be chosen so that it is bounded below by a function of $\|\psi_0\|_{H^1(\mathbb{R})}$ alone. Show that these solutions conserve mass and energy.
- Show that if T_* is the maximal time so that a weak solution exists on $[0, T_*)$, and if $T_* < \infty$, then $\|\psi(t)\|_{H^1(\mathbb{R})} \to \infty$ as $t \to T_*$. Conclude that for $\lambda > 0$ one has global solutions, i.e., $T_* = \infty$. **Note:** for $\lambda < 0$ this is false, for example solutions of negative energy blow up in finite time in the sense that $T_* < \infty$, see [**8**] or [**51**].
- For small data in $H^1(\mathbb{R})$ and powers $p \geqslant 5$ show that one has global solutions irrespective of the sign of λ.

Problem 11.5. For the one-dimensional wave equation $u_{tt} - u_{xx} = 0$ with smooth data $u\big|_{t=0} = f$ and $\partial_t u\big|_{t=0} = g$ show that the solution is given by
$$u(t,x) = \frac{1}{2}\big(f(x+t) + f(x-t)\big) + \frac{1}{2}\int_{x-t}^{x+t} g(y)\,dy$$
for all times. Conclude that such waves do not decay, precluding any possibility of a Strichartz estimate other than one involving L_t^∞.

Problem 11.6. Carry out the Ginibre-Velo approach to Strichartz estimates for the wave equation, in analogy with the Schrödinger equation as in Theorem 11.6. This needs to be done for a fixed Littlewood-Paley piece, say P_0, and then rescaled and summed up using the results of the previous chapter. This approach leads to Besov-space estimates which are stronger than the Sobolev ones (one can switch to the latter by means of Exercise 10.3 as we did in the proof of Corollary 11.10). *Hint:* See [**42**] for details.

Bibliography

[1] Adams, D. R. *A note on Riesz potentials.* Duke Math. J. 42 (1975), no. 4, 765–778.

[2] Antonov, N. Yu. *Convergence of Fourier series.* Proceedings of the XX Workshop on Function Theory (Moscow, 1995). East J. Approx. 2 (1996), no. 2, 187–196.

[3] Bahouri, H., Chemin, J.-Y., Danchin, R. *Fourier analysis and nonlinear partial differential equations.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 343. Springer, Heidelberg, 2011.

[4] Bergh, J., Löfström, J. *Interpolation spaces.* An introduction. Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976.

[5] Bourgain, J. *Bounded orthogonal systems and the* $\Lambda(p)$*?(p)-set problem.* Acta Math. 162 (1989), no. 3-4, 227–245.

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

[7] Garca-Cuerva, J., Rubio de Francia, J., *Weighted norm inequalities and related topics.* North-Holland Mathematics Studies, 116. Notas de Matemática [Mathematical Notes], 104. North-Holland Publishing Co., Amsterdam, 1985.

[8] Cazenave, T. *Semilinear Schrödinger equations*. Courant Lecture Notes in Mathematics, 10. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003.

[9] Christ, M. *Lectures on singular integral operators.* CBMS Regional Conference Series in Mathematics, 77. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1990.

[10] Coifman, R., Meyer, Y. *Wavelets. Calderón-Zygmund and multilinear operators.* Translated from the 1990 and 1991 French originals by David Salinger. Cambridge Studies in Advanced Mathematics, 48. Cambridge University Press, Cambridge, 1997.

[11] Davis, K. M., Chang *Lectures on Bochner Riesz means.* London Mathematical Society, Lecture Note Series 114, 1987.

[12] Duoandikoetxea, J. *Fourier analysis.* (English summary) Translated and revised from the 1995 Spanish original by David Cruz-Uribe. Graduate Studies in Mathematics, 29. American Mathematical Society, Providence, RI, 2001.

[13] Evans, L. C. *Partial differential equations.* Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010.

[14] Fefferman, C. *The multiplier problem for the ball.* Ann. of Math. (2) 94 (1971), 330–336.

[15] Fefferman, CṪ*he uncertainty principle.* Bull. Amer. Math. Soc. (N.S.) 9 (1983), no. 2, 129–206.

[16] Fefferman, C., Stein, E. M. H^p*Hp spaces of several variables.* Acta Math. 129 (1972), no. 3-4, 137–193.

[17] Folland, Gerald B. *Introduction to partial differential equations.* Second edition. Princeton University Press, Princeton, NJ, 1995.

[18] Frank, R., Seiringer, R. *Non-linear ground state representations and sharp Hardy inequalities.* J. Funct. Anal. 255 (2008), no. 12, 3407–3430.

[19] Frazier, M., Jawerth, B., Weiss, G. *Littlewood-Paley theory and the study of function spaces.* CBMS Regional Conference Series in Mathematics, 79. Published for the Conference Board

of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1991.

[20] Garnett, J. *Bounded analytic functions.* Revised first edition. Graduate Texts in Mathematics, 236. Springer, New York, 2007.

[21] Garnett, J. B., Marshall, D. E. *Harmonic measure.* Reprint of the 2005 original. New Mathematical Monographs, 2. Cambridge University Press, Cambridge, 2008.

[22] Gilbarg, D., Trudinger, N. *Elliptic partial differential equations of second order.* Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 224. Springer-Verlag, Berlin, 1983.

[23] Ginibre, J., Velo, G. *On a class of nonlinear Schrödinger equation. I. The Cauchy problems; II. Scattering theory, general case*, J. Func. Anal. 32 (1979), 1-32, pp. 33–71; *Scattering theory in the energy space for a class of nonlinear Schrödinger equations*, J. Math. Pures Appl. (9) 64 (1985), no. 4, pp. 363–401; *The global Cauchy problem for the nonlinear Klein-Gordon equation.* Math. Z. **189** (1985), no. 4, 487–505.; *Time decay of finite energy solutions of the nonlinear Klein-Gordon and Schrödinger equations.* Ann. Inst. H. Poincaré Phys. Théor. **43** (1985), no. 4, 399–442.

[24] Havin, V., Jöricke, B. *The uncertainty principle in harmonic analysis.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 28. Springer-Verlag, Berlin, 1994.

[25] Hoffman, K. *Banach spaces of analytic functions.* Reprint of the 1962 original. Dover Publications, Inc., New York, 1988.

[26] Hörmander, L. *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Second edition. Springer Study Edition. Springer-Verlag, Berlin, 1990.

[27] Jaming, P. *Nazarov's uncertainty principles in higher dimension.* J. Approx. Theory 149 (2007), no. 1, 3041.

[28] Janson, S. *Mean oscillation and commutators of singular integral operators.* Ark. Mat. 16 (1978), no. 2, 263–270.

[29] Jerison, D. *An elementary approach to local solvability for constant coefficient partial differential equations.* Forum Math. 2 (1990), no. 1, 45–50.

[30] Katznelson, I. *An introduction to harmonic analysis.* Dover, New York, 1976.

[31] Keel, M., Tao, T. *Endpoint Strichartz estimates*, Amer. J. Math., 120 (1998), pp. 955–980.

[32] Kenig, C. E., Ponce, G., Vega, L. *The initial value problem for the general quasi-linear Schrödinger equation.* Recent developments in nonlinear partial differential equations, 101–115, Contemp. Math., 439, Amer. Math. Soc., Providence, RI, 2007.

[33] Kolmogorov, A. N. *Une série de Fourier-Lebesgue divergente presque partout.* Fundamenta Mathematicae (4), 1923, 324–328.

[34] Koosis, P. *Introduction to H_p spaces.* Second edition. With two appendices by V. P. Havin [Viktor Petrovich Khavin]. Cambridge Tracts in Mathematics, 115. Cambridge University Press, Cambridge, 1998.

[35] Konyagin, S. V. *On the divergence everywhere of trigonometric Fourier series.* (Russian) Mat. Sb. 191 (2000), no. 1, 103–126; translation in Sb. Math. 191 (2000), no. 1-2, 97–120.

[36] Lieb, E. H., Loss, M. *Analysis.* Second edition. Graduate Studies in Mathematics, 14. American Mathematical Society, Providence, RI, 2001.

[37] Nakanishi, K., Schlag, W. *Invariant Manifolds and Dispersive Hamiltonian Evolution Equations.*, EMS, 2011.

[38] Nazarov, F. L. *Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type.* Algebra i Analiz 5 (1993), no. 4, 3–66; translation in St. Petersburg Math. J. 5 (1994), no. 4, 663–717.

[39] Opic, B., Kufner, A. *Hardy-type inequalities.* Pitman Research Notes in Mathematics Series, 219. Longman Scientific & Technical,

[40] Rudin, W. *Functional analysis.* Second edition. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991.

[41] Rudin, W. *Trigonometric series with gaps.* J. Math. Mech. 9 1960, 203–227.

[42] Shatah, J., Struwe, M. *Geometric wave equations.* Courant Lecture Notes in Mathematics, 2. American Mathematical Society, Providence, RI, 1998.

[43] Sjölin, P. *An inequality of Paley and convergence a.e. of Walsh-Fourier series.* Ark. Mat. 7, 551–570 (1969).

[44] Stein, E. M. *On limits of seqences of operators.* Ann. of Math. (2) 74 (1961), 140–170.

[45] Stein, E. M. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J., 1970.

[46] Stein, E. M. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, N.J., 1993.

[47] Stein, E. M. *Topics in harmonic analysis related to the Littlewood-Paley theory.* Annals of Mathematics Studies, No. 63 Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo 1970.

[48] Stein, E. M. *Oscillatory integrals in Fourier analysis.* Beijing lectures in harmonic analysis (Beijing, 1984), 307355, Ann. of Math. Stud., 112, Princeton Univ. Press, Princeton, NJ, 1986.

[49] Stein, E. M., Weiss, G. *Introduction to Fourier analysis on Euclidean spaces.* Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.

[50] Strichartz, R. S. *Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations.* Duke Math. J. 44 (1977), no. 3, 705–714.

[51] Sulem, C., Sulem, P-L. *The nonlinear Schrödinger equation. Self-focusing and wave collapse*, Applied Mathematical Sciences, 139. Springer-Verlag, New York, 1999.

[52] Tao, T. *Nonlinear dispersive equations.* Local and global analysis. CBMS Regional Conference Series in Mathematics, 106. AMS Providence, RI, 2006.

[53] Tao, T. *Some recent progress on the restriction conjecture.* Fourier analysis and convexity, 217–243, Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 2004.

[54] Torchinsky, A. *Real-variable methods in harmonic analysis.* Pure and Applied Mathematics, 123. Academic Press, Inc., Orlando, FL, 1986.

[55] Uchiyama, A. *A constructive proof of the Fefferman-Stein decomposition of* $\mathrm{BMO}(\mathbf{R}^n)$. Acta Math. 148 (1982), 215–241.

[56] Wolff, T. H. *Lectures on harmonic analysis.* With a foreword by Charles Fefferman and preface by Izabella Łaba. Edited by Łaba and Carol Shubin. University Lecture Series, 29. American Mathematical Society, Providence, RI, 2003.

[57] Yafaev, D. *Sharp constants in the Hardy-Rellich inequalities.* J. Funct. Anal. 168 (1999), no. 1, 121–144.

[58] Zygmund, A. *Trigonometric series.* Vol. I, II. Third edition. With a foreword by Robert A. Fefferman. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2002.