Harmonic Analysis
Lecture Notes

University of Illinois
at Urbana–Champaign

Richard S. Laugesen *

January 9, 2009

*Copyright © 2009, Richard S. Laugesen (Laugesen@illinois.edu). This work is licensed under the Creative Commons Attribution–Noncommercial–Share Alike 3.0 Unported License. To view a copy of this license, visit `http://creativecommons.org/licenses/by-nc-sa/3.0/`.

Preface

A *textbook* presents more than any professor can cover in class. In contrast, these *lecture notes* present exactly* what I covered in Harmonic Analysis (Math 545) at the University of Illinois, Urbana–Champaign, in Fall 2008.

The first part of the course emphasizes Fourier series, since so many aspects of harmonic analysis arise already in that classical context. The Hilbert transform is treated on the circle, for example, where it is used to prove L^p convergence of Fourier series. Maximal functions and Calderón–Zygmund decompositions are treated in $\mathbb{R}^d$, so that they can be applied again in the second part of the course, where the Fourier transform is studied.

Real methods are used throughout. In particular, complex methods such as Poisson integrals and conjugate functions are *not* used to prove boundedness of the Hilbert transform.

Distribution functions and interpolation are covered in the Appendices. I inserted these topics at the appropriate places in my lectures (after Chapters 4 and 12, respectively).

The references at the beginning of each chapter provide guidance to students who wish to delve more deeply, or roam more widely, in the subject. Those references do not necessarily contain all the material in the chapter.

Finally, a word on personal taste: while I appreciate a good counterexample, I prefer spending class time on positive results. Thus I do not supply proofs of some prominent counterexamples (such as Kolmogorov's integrable function whose Fourier series diverges at every point).

I am grateful to Noel DeJarnette, Eunmi Kim, Aleksandra Kwiatkowska, Kostya Slutsky, Khang Tran and Ping Xu for TEXing parts of the document.

Please email me with corrections, and with suggested improvements of any kind.

Richard S. Laugesen
Department of Mathematics
University of Illinois
Urbana, IL 61801
U.S.A.

EMAIL: Laugesen@illinois.edu

*modulo some improvements after the fact

Introduction

Harmonic analysis began with Fourier's effort to analyze (extract information from) and synthesize (reconstruct) the solutions of the heat and wave equations, in terms of harmonics. Specifically, the computation of Fourier coefficients is analysis, while writing down the Fourier series is synthesis, and the harmonics in one dimension are $\sin(nt)$ and $\cos(nt)$. Immediately one asks: does the Fourier series converge? to the original function? In what sense does it converge: pointwise? mean-square? L^p? Do analogous results hold on $\mathbb{R}^d$ for the Fourier transform?

We will answer these classical qualitative questions (and more!) using modern quantitative estimates, involving tools such as summability methods (convolution), maximal operators, singular integrals and interpolation. These topics, which we address for both Fourier series and transforms, constitute the theoretical core of the course. We further cover the sampling theorem, Poisson summation formula and uncertainty principles.

This graduate course is *theoretical* in nature. Students who are intrigued by the fascinating *applications* of Fourier series and transforms are advised to browse [Dym and McKean], [Körner] and [Stein and Shakarchi], which are all wonderfully engaging books.

If more time (or a second semester) were available, I might cover additional topics such as: Littlewood–Paley theory for Fourier series and integrals, Fourier analysis on locally compact abelian groups [Rudin] (especially Bochner's theorem on Fourier transforms of nonnegative functions), short-time Fourier transforms [Gröchenig], discrete Fourier transforms, the Schwartz class and tempered distributions and applications in Fourier analysis [Strichartz], Fourier integral operators (including solutions of the wave and Schrödinger equations), Radon transforms, and some topics related to signal processing, such as maximum entropy, spectral estimation and prediction [Benedetto]. I might also cover multiplier theorems, ergodic theorems, and almost periodic functions.

Contents

Part I

Fourier series

Chapter 1

Fourier coefficients: basic properties

Goal

Derive basic properties of Fourier coefficients

Reference

[Katznelson] Section I.1

Notation

$\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ is the one dimensional **torus**
$L^p(\mathbb{T}) = \{$complex-valued, p-th power integrable, 2π-periodic functions$\}$
$\|f\|_{L^p(\mathbb{T})} = \left(\frac{1}{2\pi}\int_{\mathbb{T}} |f(t)|^p \, dt\right)^{1/p}$ where $\int_{\mathbb{T}}$ can be taken over any interval of length 2π
Nesting of L^p-spaces: $L^\infty(\mathbb{T}) \subset L^2(\mathbb{T}) \subset L^1(\mathbb{T})$
$C(\mathbb{T}) = \{$complex-valued, continuous, 2π-periodic functions$\}$, Banach space with norm $\|\cdot\|_{L^\infty(\mathbb{T})}$
Trigonometric polynomial $P(t) = \sum_{n=-N}^{N} a_n e^{int}$
Translation $f_\tau(t) = f(t-\tau)$

Definition 1.1. For $f \in L^1(\mathbb{T})$ and $n \in \mathbb{Z}$, define

$$\begin{aligned}\widehat{f}(n) &= n\text{-th } \textit{Fourier coefficient} \text{ of } f \\ &= \frac{1}{2\pi}\int_{\mathbb{T}} f(t)e^{-int}\,dt. \end{aligned} \tag{1.1}$$

The formal series $S[f] = \sum \widehat{f}(n)e^{int}$ is the *Fourier series* of f.

Aside. For $f \in L^2(\mathbb{T})$, note $\widehat{f}(n) = \langle f, e^{int}\rangle$ where $\langle f, g\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)\overline{g(t)}\,dt$ is that L^2 inner product. Thus $\widehat{f}(n)$ =amplitude of f in direction of e^{int}. See Chapter 5.

Theorem 1.2 (Basic properties)**.** *Let $f, g \in L^1(\mathbb{T}), j, n \in \mathbb{Z}, c \in \mathbb{C}, \tau \in \mathbb{T}$.*
Linearity $\widehat{(f+g)}(n) = \widehat{f}(n) + \widehat{g}(n)$ and $\widehat{(cf)}(n) = c\widehat{f}(n)$
Conjugation $\widehat{\overline{f}}(n) = \overline{\widehat{f}(-n)}$
Trigonometric polynomial $P(t) = \sum_{n=-N}^{N} a_n e^{int}$ has $\widehat{P}(n) = a_n$ for $|n| \le N$ and $\widehat{P}(n) = 0$ for $|n| > N$
$\widehat{\ }$ *takes translation to modulation, $\widehat{f_\tau}(n) = e^{-in\tau}\widehat{f}(n)$*
$\widehat{\ }$ *takes modulation to translation, $[f(t)e^{ijt}]\widehat{\ }(n) = \widehat{f}(n-j)$*
$\widehat{\ }: L^1(\mathbb{T}) \to \ell^\infty(\mathbb{Z})$ *is bounded, with $|\widehat{f}(n)| \le \|f\|_{L^1(\mathbb{T})}$*
Hence if $f_m \to f$ in $L^1(\mathbb{T})$ then $\widehat{f_m}(n) \to \widehat{f}(n)$ (uniformly in n) as $m \to \infty$.

Proof. Exercise. □

Lemma 1.3 (Difference formula)**.** *For $n \neq 0$,*

$$\widehat{f}(n) = \frac{1}{4\pi}\int_{\mathbb{T}} [f(t) - f(t - \pi/n)]\, e^{-int}\,dt.$$

Proof.

$$\begin{aligned}\widehat{f}(n) &= -\frac{1}{2\pi}\int_{\mathbb{T}} f(t)e^{-in(t+\pi/n)}\,dt \qquad\qquad \text{since } e^{-i\pi} = -1 \\ &= -\frac{1}{2\pi}\int_{\mathbb{T}} f(t-\pi/n)e^{-int}\,dt \end{aligned} \tag{1.2}$$

by $t \mapsto t - \pi/n$ and periodicity. By (1.2) and the definition (1.1),

$$\widehat{f}(n) = \frac{1}{2}\widehat{f}(n) + \frac{1}{2}\widehat{f}(n) = \frac{1}{4\pi}\int_{\mathbb{T}} [f(t) - f(t-\pi/n)]\, e^{-int}\,dt.$$

□

Lemma 1.4 (Continuity of translation)**.** *Fix $f \in L^p(\mathbb{T}), 1 \leq p < \infty$. The map*

$$\begin{aligned} \phi : \mathbb{T} &\to L^p(\mathbb{T}) \\ \tau &\mapsto f_\tau \end{aligned}$$

is continuous.

Proof. Let $\tau_0 \in \mathbb{T}$. Take $g \in C(\mathbb{T})$ and observe

$$\begin{aligned} \|f_\tau - f_{\tau_0}\|_{L^p(\mathbb{T})} &\leq \|f_\tau - g_\tau\|_{L^p(\mathbb{T})} + \|g_\tau - g_{\tau_0}\|_{L^p(\mathbb{T})} + \|g_{\tau_0} - f_{\tau_0}\|_{L^p(\mathbb{T})} \\ &= 2\|f - g\|_{L^p(\mathbb{T})} + \|g_\tau - g_{\tau_0}\|_{L^p(\mathbb{T})} \\ &\to 2\|f - g\|_{L^p(\mathbb{T})} \end{aligned}$$

as $\tau \to \tau_0$, by uniform continuity of g. By density of continuous functions in $L^p(\mathbb{T}), 1 \leq p < \infty$, the difference $f - g$ can be made arbitrarily small. Hence $\limsup_{\tau \to \tau_0} \|f_\tau - f_{\tau_0}\|_{L^p(\mathbb{T})} = 0$, as desired. □

Corollary 1.5 (Riemann–Lebesgue lemma)**.** *$\widehat{f}(n) \to 0$ as $|n| \to \infty$.*

Proof. Lemma 1.3 implies

$$|\widehat{f}(n)| \leq \frac{1}{2}\|f - f_{\pi/n}\|_{L^1(\mathbb{T})},$$

which tends to zero as $|n| \to \infty$ by the L^1-continuity of translation in Lemma 1.4, since $f = f_0$. □

Smoothness and decay

The Riemann–Lebesgue lemma says $\widehat{f}(n) = o(1)$, with $\widehat{f}(n) = O(1)$ explicitly by Theorem 1.2. We show the smoother f is, the faster its Fourier coefficients decay.

Theorem 1.6 (Less than one derivative)**.** *If $f \in C^\alpha(\mathbb{T}), 0 < \alpha \leq 1$, then $\widehat{f}(n) = O(1/|n|^\alpha)$.*

Here $C^\alpha(\mathbb{T})$ denotes the Hölder continuous functions: $f \in C^\alpha(\mathbb{T})$ if $f \in C(\mathbb{T})$ and there exists $A > 0$ such that $|f(t) - f(\tau)| \leq A|t - \tau|^\alpha$ whenever $|t - \tau| \leq 2\pi$.

Proof.

$$\widehat{f}(n) = \frac{1}{4\pi}\int_{\mathbb{T}} [f(t) - f(t - \pi/n)]e^{-int}\,dt$$

by the Difference Formula in Lemma 1.3. Therefore

$$|\widehat{f}(n)| \le \frac{1}{4\pi} A \left|\frac{\pi}{n}\right|^{\alpha} 2\pi = \frac{\text{const.}}{|n|^{\alpha}}.$$

□

Theorem 1.7 (One derivative)**.** *If f is 2π-periodic and absolutely continuous ($f \in W^{1,1}(\mathbb{T})$) then $\widehat{f}(n) = o(1/n)$ and $|\widehat{f}(n)| \le \|f'\|_{L^1(\mathbb{T})}/|n|$.*

Proof. Absolute continuity of f says

$$f(t) = f(0) + \int_0^t f'(\tau)\,d\tau,$$

where $f' \in L^1(\mathbb{T})$. Integrating by parts gives

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)e^{-int}\,dt = \frac{1}{2\pi}\int_{\mathbb{T}} \frac{e^{-int}}{in} f'(t)\,dt.$$

By Riemann-Lebesgue applied to f',

$$\widehat{f}(n) = \frac{1}{in}\widehat{f'}(n) = \frac{1}{in}o(1) = o(\frac{1}{n}),$$

with

$$|\widehat{f}(n)| \le \frac{1}{|n|}|\widehat{f'}(n)| \le \frac{1}{|n|}\|f'\|_{L^1(\mathbb{T})}.$$

□

Theorem 1.8 (Higher derivatives)**.** *If f is 2π-periodic and k times differentiable ($f \in W^{k,1}(\mathbb{T})$) then $\widehat{f}(n) = o(1/|n|^k)$ and $|\widehat{f}(n)| \le \|f^{(k)}\|_{L^1(\mathbb{T})}/|n|^k$.*

Proof. Integrate by parts k times. □

Remark 1.9. Similar decay results hold for functions of bounded variation, provided one integrates by parts using the Lebesgue–Stieltjes measure $df(t)$ instead of $f'(t)\,dt$.

Convolution

Definition 1.10. Given $f, g \in L^1(\mathbb{T})$, define their *convolution*

$$(f * g)(t) = \frac{1}{2\pi}\int_{\mathbb{T}} f(t-\tau)g(\tau)\,d\tau, \qquad t \in \mathbb{T}.$$

Theorem 1.11 (Convolution and Fourier coefficients)**.** *If $f \in L^p(\mathbb{T}), 1 \le p \le \infty$, and $g \in L^1(\mathbb{T})$, then $f * g \in L^p(\mathbb{T})$ with*

$$\|f * g\|_{L^p(\mathbb{T})} \le \|f\|_{L^p(\mathbb{T})}\|g\|_{L^1(\mathbb{T})} \tag{1.3}$$

and

$$\widehat{(f * g)}(n) = \widehat{f}(n)\widehat{g}(n), \qquad n \in \mathbb{Z}.$$

Further, if $f \in C(\mathbb{T})$ and $g \in L^1(\mathbb{T})$ then $f * g \in C(\mathbb{T})$.

Thus $\widehat{\ }$ takes convolution to multiplication.

Proof. That $f * g \in L^p(\mathbb{T})$ satisfies (1.3) is exactly Young's Theorem A.3. Then by Fubini's theorem,

$$\begin{aligned}
\widehat{(f * g)}(n) &= \frac{1}{2\pi}\int_{\mathbb{T}} \big(\frac{1}{2\pi}\int_{\mathbb{T}} f(t-\tau)g(\tau)\,d\tau\big)e^{-int}\,dt \\
&= \frac{1}{2\pi}\int_{\mathbb{T}} \big(\frac{1}{2\pi}\int_{\mathbb{T}} f(t-\tau)e^{-in(t-\tau)}\,dt\big)g(\tau)e^{-in\tau}\,d\tau \\
&= \widehat{f}(n)\widehat{g}(n).
\end{aligned}$$

Finally, if $f \in C(\mathbb{T})$ and $g \in L^1(\mathbb{T})$ then $f * g$ is continuous because $(f * g)(t+\delta) \to (f * g)(t)$ as $\delta \to 0$ by uniform continuity of f. □

Convolution facts [Katznelson, Section I.1.8]

1. Convolution is commutative:

$$\begin{aligned}
(f * g)(t) &= \frac{1}{2\pi}\int_{\mathbb{T}} f(t-\tau)g(\tau)\,d\tau \\
&= \frac{1}{2\pi}\int_{\mathbb{T}} f(\theta)g(t-\theta)\,d\theta \qquad \text{where } \tau = t-\theta, d\tau = -d\theta \\
&= (g * f)(t).
\end{aligned}$$

Convolution is also associative, and linear with respect to f and g.

2. Convolution is continuous on $L^p(\mathbb{T})$: if $f_m \to f \in L^p(\mathbb{T}), 1 \le p \le \infty$, and $g \in L^1(\mathbb{T})$ then $f_m * g \to f * g$ in $L^p(\mathbb{T})$.
Proof. Use linearity and (1.3), to prove $f_m * g \to f * g$ in $L^p(\mathbb{T})$.

3. Convolution with a trigonometric polynomial gives a trigonometric polynomial: if $f \in L^1(\mathbb{T})$ and $P(t) = \sum_{j=-n}^{n} a_j e^{ijt}$ then

$$(P * f)(t) = \sum_{j=-n}^{n} a_j \widehat{f}(j) e^{ijt}. \tag{1.4}$$

Proof.

$$\begin{aligned}(P * f)(t) &= \sum_{j=-n}^{n} a_j \frac{1}{2\pi} \int_{\mathbb{T}} e^{ij(t-\tau)} f(\tau)\, d\tau \\ &= \sum_{j=-n}^{n} a_j e^{ijt} \widehat{f}(j).\end{aligned}$$

[Sanity check: $\widehat{(P * f)}(j) = a_j \widehat{f}(j) = \widehat{P}(j)\widehat{f}(j)$ as expected.]
More generally, (1.4) holds for $P(t) = \sum_{j=-\infty}^{\infty} a_j e^{ijt}$ provided $\{a_j\} \in \ell^1(\mathbb{Z})$.

Chapter 2

Fourier series: summability in norm

Goal

Prove summability (averaged convergence) in norm of Fourier series

Reference

[Katznelson] Section I.2

Write

$$\begin{aligned}(S_n f)(t) &= \sum_{j=-n}^{n} \widehat{f}(j) e^{ijt} \\ &= n\text{-th partial sum of Fourier series of } f.\end{aligned}$$

In Chapter 9 we prove norm convergence of Fourier series: $S_n(f) \to f$ in $L^p(\mathbb{T})$, when $1 < p < \infty$. In this chapter we prove **summability** of Fourier series, meaning $\sigma_n(f) \to f$ in $L^p(\mathbb{T})$ when $1 \leq p < \infty$, where

$$\begin{aligned}\sigma_n(f) &= \frac{S_0(f) + \cdots + S_n(f)}{n+1} = \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j) \\ &= \text{arithmetic mean of partial sums.}\end{aligned}$$

Aside. Norm convergence is stronger than summability. Indeed, if a sequence $\{s_n\}$ in a Banach space converges to s, then the arithmetic means $(s_0 + \cdots + s_n)/(n+1)$ also converge to s (Exercise).

Definition 2.1. A *summability kernel* is a sequence $\{k_n\}$ in $L^1(\mathbb{T})$ satisfying:

$$\frac{1}{2\pi}\int_{\mathbb{T}} k_n(t)\,dt = 1 \qquad \text{(Normalization)} \tag{S1}$$

$$\sup_n \frac{1}{2\pi}\int_{\mathbb{T}} |k_n(t)|\,dt < \infty \qquad (L^1 \text{ bound}) \tag{S2}$$

$$\lim_{n\to\infty} \int_{\{\delta<|t|<\pi\}} |k_n(t)|\,dt = 0 \qquad (L^1 \text{ concentration}) \tag{S3}$$

for each $\delta \in (0,\pi)$.

Some kernels satisfy a stronger concentration property:

$$\lim_{n\to\infty} \sup_{\delta<|t|<\pi} |k_n(t)| = 0 \qquad (L^\infty \text{ concentration}) \tag{S4}$$

for each $\delta \in (0,\pi)$.

Call the kernel *positive* if $k_n \geq 0$ for each n.

Example 2.2. Define the **Dirichlet kernel**

$$D_n(t) = \sum_{j=-n}^{n} e^{ijt} \tag{2.1}$$

$$= \frac{e^{i(n+1)t} - e^{-int}}{e^{it}-1} \qquad \text{by geometric series} \tag{2.2}$$

$$= \frac{\sin\left((n+\frac{1}{2})t\right)}{\sin\left(\frac{1}{2}t\right)} \tag{2.3}$$

(S1) holds by (2.1). You can show (optional exercise) that $\|D_n\|_{L^1(\mathbb{T})} \sim (\text{const.})\log n$ as $n\to\infty$, so that (S2) fails.

$\therefore$ $\{D_n\}$ is **not** a summability kernel.

Example 2.3. Define the **Fejér kernel**

$$F_n(t) = \frac{D_0(t)+\cdots+D_n(t)}{n+1} \tag{2.4}$$

$$= \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) e^{ijt} \quad \text{by (2.4) and (2.1)} \tag{2.5}$$

$$= \frac{1}{n+1}\left(\frac{\sin\left(\frac{n+1}{2}t\right)}{\sin\left(\frac{1}{2}t\right)}\right)^2 \quad \text{by (2.4), (2.2) and geometric series} \tag{2.6}$$

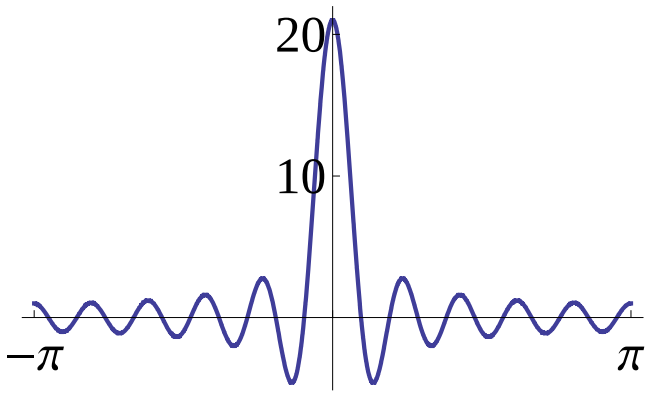

Figure 2.1: Dirichlet kernel with $n = 10$

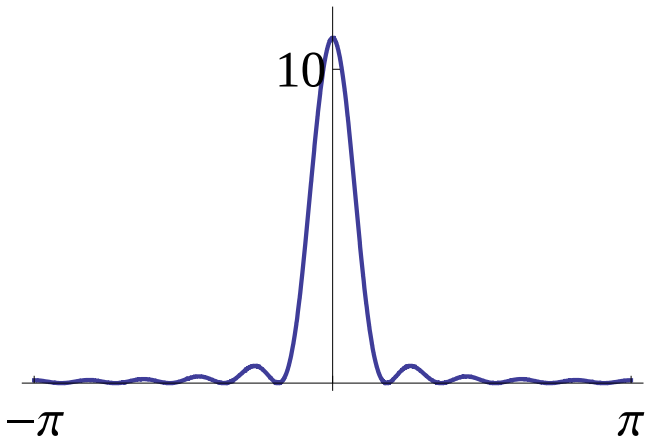

Figure 2.2: Fejér kernel with $n = 10$

(S1) holds by (2.5), and $F_n \geq 0$ so that (S2) holds also. For (S4),

$$\begin{aligned} \sup_{\delta<|t|<\pi} |F_n(t)| &\leq \frac{1}{n+1}\frac{1}{\sin^2\left(\frac{1}{2}\delta\right)} && \text{by (2.6)} \\ &\to 0 && \text{as } n \to \infty. \end{aligned}$$

$\therefore$ $\{F_n\}$ is a positive summability kernel.

Example 2.4. Define the **Poisson kernel**

$$P_r(t) = 1 + 2\sum_{j=1}^{\infty} r^j \cos(jt) \tag{2.7}$$

$$= \sum_{j=-\infty}^{\infty} r^{|j|} e^{ijt} \tag{2.8}$$

$$= \frac{1-r^2}{1-2r\cos t + r^2} \tag{2.9}$$

by summing two geometric series ($j < 0$ and $j \geq 0$) in (2.8) and simplifying.

The Poisson kernel is indexed by $r \in (0,1)$, with limiting process $r \nearrow 1$. After suitably modifying the definition of summability kernel, we see (S1) holds by (2.7), and $P_r \geq 0$ by (2.9) so that (S2) holds also. For (S4),

$$\begin{aligned} \sup_{\delta<|t|<\pi} |P_r(t)| &\leq \frac{1-r^2}{1-2r\cos\delta + r^2} && \text{by (2.9)} \\ &\to 0 && \text{as } r \nearrow 1. \end{aligned}$$

$\therefore$ $\{P_r\}$ is a positive summability kernel.

Example 2.5. Define the **Gauss kernel**

$$G_s(t) = \sum_{j=-\infty}^{\infty} e^{-j^2 s} e^{ijt} \tag{2.10}$$

$$= \frac{2\pi}{\sqrt{4\pi s}} \sum_{n=-\infty}^{\infty} e^{-(t+2\pi n)^2/4s} \tag{2.11}$$

by Example 23.7 later in the course.

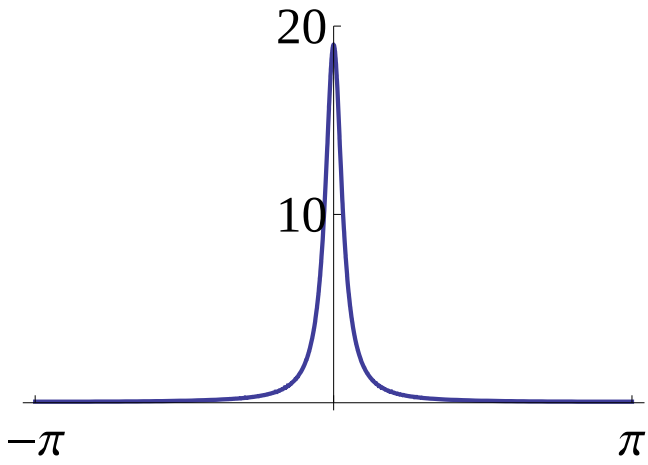

Figure 2.3: Poisson kernel with $r = 0.9$

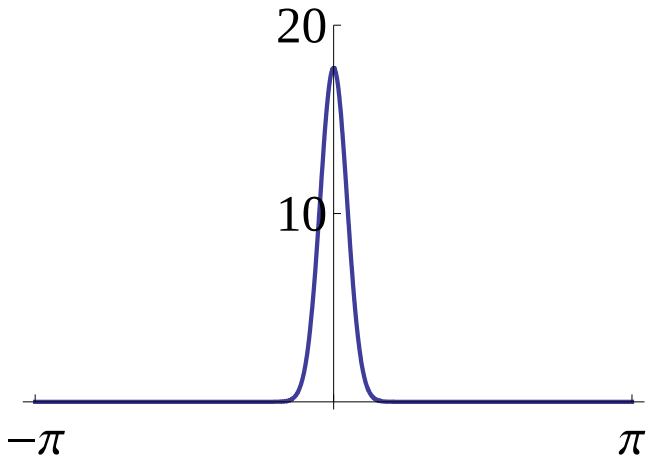

Figure 2.4: Gauss kernel with $s = 0.01$

The Gauss kernel is indexed by $s \in (0, \infty)$, with limiting process $s \searrow 0$. The analogue of (S1) holds by (2.10), and $G_s \geq 0$ by (2.11) so that (S2) holds also. For (S4),

$$\begin{aligned} \sup_{\delta<|t|<\pi} |G_s(t)| &\leq \frac{2\pi}{\sqrt{4\pi s}} \left[e^{-\delta^2/4s} + \sum_{n\neq 0} e^{-(\pi n)^2/4s} \right] && \text{by (2.11)} \\ &\to 0 && \text{as } s \searrow 0. \end{aligned}$$

$\therefore$ $\{G_s\}$ is a positive summability kernel.

Connection to Fourier series

$$\boxed{S_n(f) = D_n * f}$$

Proof. $D_n(t) \overset{(2.1)}{=} \sum_{j=-n}^{n} 1 e^{ijt}$ implies

$$(D_n * f)(t) = \sum_{j=-n}^{n} 1 \widehat{f}(j) e^{ijt} = S_n(f)$$

by Convolution Fact (1.4).

$$\boxed{\sigma_n(f) = F_n * f}$$

Proof. $F_n(t) \overset{(2.4)}{=} \sum_{j=-n}^{n} (1 - \frac{|j|}{n+1}) e^{ijt}$ implies

$$(F_n * f)(t) = \sum_{j=-n}^{n} \big(1 - \frac{|j|}{n+1}\big) \widehat{f}(j) e^{ijt} = \sigma_n(f)$$

by Convolution Fact (1.4). Alternatively, use that $\sigma_n(f) = [S_0(f) + \cdots + S_n(f)]/(n+1)$ and $F_n = [D_0 + \cdots + D_n]/(n+1)$.

Thus for summability of Fourier series, we want $F_n * f \to f$.

$$\boxed{\text{Abel mean of } S[f] = P_r * f}$$

Proof. $P_r(t) \overset{(2.8)}{=} \sum_{j=-\infty}^{\infty} r^{|j|} e^{ijt}$ implies

$$(P_r * f)(t) = \sum_{j=-\infty}^{\infty} r^{|j|} \widehat{f}(j) e^{ijt} \tag{2.12}$$

by Convolution Fact (1.4) (with the series converging absolutely and uniformly), and this last expression is the Abel mean of $S[f]$.

Summability in norm

Theorem 2.6 (Summability in $L^p(\mathbb{T})$ and $C(\mathbb{T})$)**.** *If $\{k_n\}$ is a summability kernel and $f \in L^p(\mathbb{T}), 1 \le p < \infty$, then*

$$k_n * f \to f \qquad \text{in } L^p(\mathbb{T}), \quad \text{as } n \to \infty.$$

Similarly, if $f \in C(\mathbb{T})$ then $k_n * f \to f$ in $C(\mathbb{T})$.

Proof. Let $\varepsilon > 0$. By (S2) and continuity of translation on $L^p(\mathbb{T})$ (Lemma 1.4), we can choose $0 < \delta < \pi$ such that

$$\max_{|\tau| \le \delta} \|f_\tau - f\|_{L^p(\mathbb{T})} \cdot \sup_n \|k_n\|_{L^1(\mathbb{T})} < \varepsilon. \tag{2.13}$$

Then

$$\begin{aligned}
&\big\|(k_n * f)(t) - f(t)\big\|_{L^p(\mathbb{T})} \\
&= \Big\|\frac{1}{2\pi}\int_{\mathbb{T}} k_n(\tau)[f_\tau(t) - f(t)]\, d\tau\Big\|_{L^p(\mathbb{T})} && \text{by (S1)} \\
&\le \frac{1}{2\pi}\int_{\mathbb{T}} |k_n(\tau)| \|f_\tau - f\|_{L^p(\mathbb{T})}\, d\tau \\
&\qquad\qquad \text{by Minkowski's Integral Inequality, Theorem A.1,} \\
&= \frac{1}{2\pi}\left(\int_{-\delta}^{\delta} + \int_{\{\delta<|\tau|<\pi\}}\right) |k_n(\tau)| \|f_\tau - f\|_{L^p(\mathbb{T})}\, d\tau \\
&\le \max_{|\tau|\le\delta} \|f_\tau - f\|_{L^p(\mathbb{T})} \frac{1}{2\pi}\int_{-\delta}^{\delta} |k_n(\tau)|\, d\tau \\
&\quad + \max_{|\tau|\le\pi} \|f_\tau - f\|_{L^p(\mathbb{T})} \frac{1}{2\pi}\int_{\{\delta<|\tau|<\pi\}} |k_n(\tau)|\, d\tau \\
&< \varepsilon + \varepsilon
\end{aligned}$$

by (2.13) and (S3), for all large n.

If $f \in C(\mathbb{T})$ then repeat the argument with $p = \infty$, using uniform continuity of f get that $f_\tau \to f$ in $L^\infty(\mathbb{T})$. □

Consequences

• Summability of Fourier series in $C(\mathbb{T}), L^p(\mathbb{T}), 1 \le p < \infty$:

$$\sigma_n(f) \to f$$

in norm.
Proof. Choose $k_n = F_n =$ Fejér kernel. Then $\sigma_n(f) = F_n * f \to f$ in norm by Theorem 2.6.

• Trigonometric polynomials are dense in $C(\mathbb{T}), L^p(\mathbb{T}), 1 \le p < \infty$.
Proof. $\sigma_n(f)$ is a trigonometric polynomial arbitrarily close to f.

Aside. Density of trigonometric polynomials in $C(\mathbb{T})$ proves the Weierstrass Trigonometric Approximation Theorem.

• Uniqueness theorem:

$$\text{if } f, g \in L^1(\mathbb{T}) \text{ with } \widehat{f}(n) = \widehat{g}(n) \text{ for all } n, \text{ then } f = g \text{ in } L^1(\mathbb{T}). \tag{2.14}$$

In other words, the map $\widehat{\ }: L^1(\mathbb{T}) \to \ell^\infty(\mathbb{Z})$ is injective.
Proof. $F_n * f = F_n * g$ by Convolution Fact (1.4), since $\widehat{f} = \widehat{g}$. Letting $n \to \infty$ gives $f = g$.

Connection to PDEs

To finish the section, we connect our summability kernels to some important partial differential equations. Fix $f \in L^1(\mathbb{T})$.

1. The Poisson kernel solves Laplace's equation in a disk:

$$v(re^{it}) = (P_r * f)(t) = \frac{1}{2\pi} \int_{\mathbb{T}} \frac{1 - r^2}{1 - 2r\cos(t - \tau) + r^2} f(\tau)\, d\tau$$

solves

$$\Delta v = v_{rr} + r^{-1} v_r + r^{-2} v_{tt} = 0$$

on the unit disk $\{r < 1\}$, with boundary value $v(1, t) = f(t)$ in the sense of Theorem 2.6.

That is, v is the harmonic extension of f from the boundary circle to the disk.

Proof. Differentiate through formula (2.12) for $P_r * f$ and note that

$$\Big(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial t^2}\Big)(r^{|j|}e^{ijt}) = 0.$$

2. The Gauss kernel solves the diffusion (heat) equation:

$$w(s,t) = (G_s * f)(t)$$

solves

$$w_s = w_{tt}$$

for $(s,t) \in (0,\infty) \times \mathbb{T}$, with initial value $w(0,t) = f(t)$ in the sense of Theorem 2.6.

Proof. $G_s(t) \overset{(2.10)}{=} \sum_{j=-\infty}^{\infty} e^{-j^2 s} e^{ijt}$ implies

$$(G_s * f)(t) = \sum_{j=-\infty}^{\infty} e^{-j^2 s}\widehat{f}(j)e^{ijt}$$

by Convolution Fact (1.4). Now differentiate through the sum and use that

$$\left(\frac{\partial}{\partial s} - \frac{\partial^2}{\partial t^2}\right)(e^{-j^2 s}e^{ijt}) = 0.$$

Chapter 3

Fourier series: summability at a point

Goal

Prove a sufficient condition for summability at a point

Reference

[Katznelson] Section I.3

By Chapter 2, if f is continuous then $\sigma_n(f) \to f$ in $C(\mathbb{T})$. That is, $\sigma_n(f) \to f$ uniformly. In particular, $\sigma_n(f)(t) \to f(t)$, for each $t \in \mathbb{T}$.

But what if f is merely continuous at a *point*?

Theorem 3.1 (Summability at a point)**.** *Assume $\{k_n\}$ is a summability kernel,$f \in L^1(\mathbb{T})$ and $t_0 \in \mathbb{T}$. Suppose either $\{k_n\}$ satisfies the L^∞ concentration hypothesis* (S4)*, or else $f \in L^\infty(\mathbb{T})$.*

(a) If f is continuous at t_0 then $(k_n * f)(t_0) \to f(t_0)$ as $n \to \infty$.

(b) If in addition the summability kernel is even ($k_n(-t) = k_n(t)$) and

$$L = \lim_{h \to 0} \frac{f(t_0 + h) + f(t_0 - h)}{2}$$

exists (or equals $\pm\infty$), then

$$(k_n * f)(t_0) \to L \qquad \text{as } n \to \infty.$$

Note if f has limits from the left and right at t_0, then the quantity L equals the average of those limits.

The Fejér and Poisson kernels satisfy (SR4), and so Theorem 3.1 applies in particular to summability at a point for $\sigma_n(f) = F_n * f$ and for the Abel mean $P_r * f$.

Proof. (a) Let $\varepsilon > 0$ and choose $0 < \delta < \pi$ such that

$$\sup_n \|k_n\|_{L^1(\mathbb{T})} \cdot \max_{|\tau| \le \delta} |f(t_0 - \tau) - f(t_0)| < \varepsilon, \tag{3.1}$$

using here (S2) and continuity of f at t_0. Then as $n \to \infty$,

$$\begin{aligned}
&|(k_n * f)(t_0) - f(t_0)| \\
&= \Big| \int_{\{|\tau|<\delta\}} k_n(\tau)[f(t_0-\tau) - f(t_0)]\, d\tau \\
&\qquad - \int_{\{\delta<|\tau|<\pi\}} k_n(\tau)\, d\tau \cdot f(t_0) + \int_{\{\delta<|\tau|<\pi\}} k_n(\tau) f(t_0-\tau)\, d\tau \Big| \qquad \text{using (S1)} \\
&< \begin{cases} \varepsilon + o(1) + o(1) \cdot \|f\|_{L^1(\mathbb{T})} & \text{by (3.1), (S3) and (S4), or else} \\ \varepsilon + o(1) + o(1) \cdot \|f\|_{L^\infty(\mathbb{T})} & \text{by (3.1), (S3) and (S1),} \end{cases} \\
&< \varepsilon
\end{aligned}$$

for all large n.

(b) The proof is similar to (a), but uses symmetry of the kernel. □

Remark 3.2.

1. How does the proof of summability at a point, in Theorem 3.1(a), differ from the proof of summability in norm, in Theorem 2.6?
2. Theorem 3.1 treats summability at a single point t_0 at which f is continuous. Chapter 7 will prove $k_n * f \to f$ at almost every point, for each integrable f.

Chapter 4

Fourier coefficients in $\ell^1(\mathbb{Z})$ (or, $f \in A(\mathbb{T})$)

Goal

Establish the algebra structure of $A(\mathbb{T})$

Reference

[Katznelson] Section I.6

Define

$$\begin{aligned} A(\mathbb{T}) &= \{f \in L^1(\mathbb{T}) : \sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty\} \\ &= \text{functions with Fourier coefficients in } \ell^1(\mathbb{Z}). \end{aligned}$$

The map $\widehat{\ }: A(\mathbb{T}) \to \ell^1(\mathbb{Z})$ is a linear bijection.

Proof. Injectivity follows from the uniqueness result (2.14). To prove surjectivity, let $\{c_n\} \in \ell^1(\mathbb{Z})$ and define $g(t) = \sum_{n\in\mathbb{Z}} c_n e^{int}$. The series for g converges uniformly since

$$\sup_{t\in\mathbb{T}} \Big| \sum_{|n|>N} c_n e^{int} \Big| \le \sum_{|n|>N} |c_n| \to 0$$

as $N \to \infty$. (Hence g is continuous.) We have $\widehat{g}(m) = c_m$ for every m, and so $\widehat{g} = \{c_m\}$ as desired.

Our proof has shown each $f \in A(\mathbb{T})$ is represented by its Fourier series:

$$f(t) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{int} \quad \text{a.e.} \tag{4.1}$$

so that f is continuous (after redefinition on a set of measure zero). This Fourier series converges absolutely and uniformly.

Definition 4.1. Define a norm on $A(\mathbb{T})$ by

$$\|f\|_{A(\mathbb{T})} = \|\widehat{f}\|_{\ell^1(\mathbb{Z})} = \sum_n |\widehat{f}(n)|.$$

$A(\mathbb{T})$ is a Banach space under this norm (because $\ell^1(\mathbb{Z})$ is one).

Define the convolution of sequences $a, b \in \ell^1(\mathbb{Z})$ by

$$(a * b)(n) = \sum_{m\in\mathbb{Z}} a(m)b(n-m).$$

Clearly

$$\|a * b\|_{\ell^1(\mathbb{Z})} \le \|a\|_{\ell^1(\mathbb{Z})}\|b\|_{\ell^1(\mathbb{Z})} \tag{4.2}$$

because

$$\sum_n |(a*b)(n)| \le \sum_m |a(m)| \sum_n |b(n-m)| = \|a\|_{\ell^1(\mathbb{Z})}\|b\|_{\ell^1(\mathbb{Z})}.$$

Theorem 4.2 ($\widehat{\ }$ takes multiplication to convolution)**.** *$A(\mathbb{T})$ is an algebra, meaning that if $f, g \in A(\mathbb{T})$ then $fg \in A(\mathbb{T})$. Indeed*

$$\widehat{fg} = \widehat{f} * \widehat{g}$$

and $\|fg\|_{A(\mathbb{T})} \le \|f\|_{A(\mathbb{T})}\|g\|_{A(\mathbb{T})}$.

Proof. fg is continuous, and hence integrable, with

$$\begin{aligned}
\widehat{(fg)}(n) &= \frac{1}{2\pi}\int_{\mathbb{T}} f(t)g(t)e^{-int}dt \\
&= \sum_m \widehat{f}(m)\frac{1}{2\pi}\int_{\mathbb{T}} g(t)e^{-i(n-m)t}\,dt \qquad \text{by (4.1)} \\
&= \sum_m \widehat{f}(m)\widehat{g}(n-m) \\
&= (\widehat{f} * \widehat{g})(n).
\end{aligned}$$

So $\|\widehat{(fg)}\|_{\ell^1(\mathbb{Z})} \le \|\widehat{f}\|_{\ell^1(\mathbb{Z})}\|\widehat{g}\|_{\ell^1(\mathbb{Z})}$ by (4.2). □

Sufficient conditions for membership in $A(\mathbb{T})$ are discussed in [Katznelson, Section I.6], for example, Hölder continuity: $C^\alpha(\mathbb{T}) \subset A(\mathbb{T})$ when $\alpha > \frac{1}{2}$.

Theorem 4.3 (Wiener's Inversion Theorem)**.** *If $f \in A(\mathbb{T})$ and $f(t) \neq 0$ for every $t \in \mathbb{T}$ then $1/f \in A(\mathbb{T})$.*

We omit the proof. Clearly $1/f$ is continuous, but it is *not* clear that $\widehat{(1/f)}$ belongs to $\ell^1(\mathbb{Z})$.

Chapter 5

Fourier coefficients in $\ell^2(\mathbb{Z})$ (or, $f \in L^2(\mathbb{T})$)

Goal

Study the Fourier ONB for $L^2(\mathbb{T})$, using analysis and synthesis operators

Notation and definitions

Let H be a Hilbert space with inner product $\langle u, v\rangle$ and norm $\|u\| = \sqrt{\langle u, u\rangle}$.

Given a sequence $\{u_n\}_{n\in\mathbb{Z}}$ in H, define the

$$\textbf{synthesis operator} \quad S : \ell^2(\mathbb{Z}) \to H$$
$$\{c_n\}_{n\in\mathbb{Z}} \mapsto \sum_n c_n u_n$$

and

$$\textbf{analysis operator} \quad T : H \to \ell^2(\mathbb{Z})$$
$$u \mapsto \{\langle u, u_n\rangle\}_{n\in\mathbb{Z}}.$$

Theorem 5.1. *If analysis is bounded ($\sum_n |\langle u, u_n\rangle|^2 \le (const.)\|u\|^2$ for all $u \in H$), then so is synthesis, and the series $S(\{c_n\}) = \sum c_n u_n$ converges unconditionally.*

Proof. Since T is bounded, the adjoint $T^* : \ell^2(\mathbb{Z}) \to H$ is bounded, and for each sequence $\{c_n\}, u \in H, N \geq 1$, we have

$$\begin{aligned}\langle T^*(\{c_n\}_{n=-N}^N), u\rangle &= \langle \{c_n\}_{n=-N}^N, Tu\rangle_{\ell^2} \\ &= \sum_{n=-N}^{N} c_n \overline{\langle u, u_n\rangle} \qquad \text{by definition of } Tu \\ &= \langle \sum_{n=-N}^{N} c_n u_n, u\rangle.\end{aligned}$$

Hence $T^*(\{c_n\}_{n=-N}^N) = \sum_{n=-N}^N c_n u_n$. The limit as $N \to \infty$ exists on the left side, and hence on the right side; therefore $T^*(\{c_n\}) = \sum_{n=-\infty}^{\infty} c_n u_n$, so that $T^* = S$. Hence S is bounded.

Convergence of the synthesis series is unconditional, because if $A \subset \mathbb{Z}$ then

$$\begin{aligned}\big\|S(\{c_n\}_{n\in\mathbb{Z}}) - S(\{c_n\}_{n\in A})\big\| &= \big\|S(\{c_n\}_{n\in\mathbb{Z}\setminus A})\big\| \\ &\leq \|S\|\|\{c_n\}\|_{\ell^2(\mathbb{Z}\setminus A)},\end{aligned}$$

which tends to 0 as A expands to fill $\mathbb{Z}$, regardless of the order in which A expands. □

Remark 5.2. The last proof shows $S = T^*$, meaning

analysis and synthesis are adjoint operations.

Theorem 5.3 (Fourier coefficients on $L^2(\mathbb{T})$)**.** *The Fourier coefficient (or analysis) operator* $\widehat{\ }: L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ *is an isometry, with*

$$\begin{aligned}\|f\|_{L^2(\mathbb{T})} &= \|\widehat{f}\|_{\ell^2(\mathbb{Z})} && \text{(Plancherel)} \\ \langle f, g\rangle_{L^2(\mathbb{T})} &= \langle \widehat{f}, \widehat{g}\rangle_{\ell^2(\mathbb{Z})} && \text{(Parseval)}\end{aligned}$$

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

Proof. First we prove Plancherel's identity: since $P_r * f \to f$ in $L^2(\mathbb{T})$ by

Theorem 2.6, we have

$$\begin{aligned}\frac{1}{2\pi}\int_{\mathbb{T}} |f(t)|^2\,dt &= \lim_{r\to 1}\frac{1}{2\pi}\int_{\mathbb{T}} f(t)\overline{(P_r * f)(t)}\,dt \\ &= \lim_{r\to 1}\frac{1}{2\pi}\int_{\mathbb{T}} f(t)\sum_{n\in\mathbb{Z}}\overline{r^{|n|}\widehat{f}(n)e^{int}}\,dt \qquad \text{by (2.12) for } P_r * f \\ &= \lim_{r\to 1}\sum_{n\in\mathbb{Z}} r^{|n|}|\widehat{f}(n)|^2 \\ &= \sum_{n\in\mathbb{Z}}|\widehat{f}(n)|^2\end{aligned}$$

by monotone convergence.

Parseval follows from Plancherel by polarization, or by repeating the argument for Plancherel with $\langle f, f\rangle$ changed to $\langle f, g\rangle$ (and using dominated instead of monotone convergence). □

Since the Fourier analysis operator is bounded, so is its adjoint, the Fourier synthesis operator

$$\begin{aligned}\check{}\,: \ell^2(\mathbb{Z}) &\to L^2(\mathbb{T}) \\ \{c_n\}_{n\in\mathbb{Z}} &\mapsto \sum_n c_n e^{int}\end{aligned}$$

Theorem 5.4 (Fourier ONB)**.**
(a) If $f \in L^2(\mathbb{T})$ then $\sum_n \widehat{f}(n)e^{int} = f$ with unconditional convergence in $L^2(\mathbb{T})$. That is, $(\hat{f})\check{} = f$.
(b) If $c = \{c_n\} \in \ell^2(\mathbb{Z})$ then $(\sum_{n\in\mathbb{Z}} c_n e^{int})\widehat{\ }(j) = c_j$. That is, $(\check{c})\hat{} = c$.
(c) $\{e^{int}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{T})$.

Part (a) says Fourier series converge in $L^2(\mathbb{T})$. Parts (a) and (b) together show that Fourier analysis and synthesis are inverse operations.

Proof. Fourier analysis and synthesis are bounded operators, and analysis followed by synthesis equals the identity ($\sum_n \widehat{f}(n)e^{int} = f$) on the class of trigonometric polynomials. That class is dense in $L^2(\mathbb{T})$, and so by continuity, analysis followed by synthesis equals the identity on $L^2(\mathbb{T})$.

Argue similarly for part (b), using the dense class of finite sequences in $\ell^2(\mathbb{Z})$.

For orthonormality in part (c), observe

$$\langle e^{int}, e^{imt} \rangle = \frac{1}{2\pi} \int_{\mathbb{T}} e^{int} e^{-imt}\, dt = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{if } n \neq m. \end{cases}$$

The basis property follows from part (a), noting $\widehat{f}(n) = \langle f, e^{int} \rangle_{L^2(\mathbb{T})}$. □

Remark 5.5. Fourier analysis satisfies

$$\begin{aligned} &\widehat{\ }\colon L^1(\mathbb{T}) \to \ell^\infty(\mathbb{Z}) && \text{by Theorem 1.2,} \\ &\widehat{\ }\colon L^2(\mathbb{T}) \to \ell^2(\mathbb{Z}) && \text{(isometrically) by Theorem 5.3.} \end{aligned}$$

Further, $\widehat{\ }\colon L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ is a linear bijection by Theorem 5.4.

In Chapter 13 we will interpolate to show

$$\widehat{\ }\colon L^p(\mathbb{T}) \to \ell^{p'}(\mathbb{Z}), \qquad \text{whenever } 1 \leq p \leq 2, \quad \frac{1}{p} + \frac{1}{p'} = 1.$$

Chapter 6

Maximal functions

Goals

Connect abstract maximal functions to convergence a.e.
Prove weak and strong bounds on the Hardy–Littlewood maximal function
Prepare for sumability pointwise a.e. in next Chapter

References

[Duoandikoetxea] Section 2.2
[Grafakos] Section 2.1
[Stein] Section 1.1

Definition 6.1 (Weak and strong operators)**.** Let (X, μ) and (Y, ν) be measure spaces, and $1 \le p, q \le \infty$. Suppose

$$T : L^p(X) \to \{\text{measurable functions on } Y\}.$$

(We do not assume T is linear.)

Call T **strong** $(\mathbf{p}, \mathbf{q})$ if T is bounded from $L^p(X)$ to $L^q(Y)$, meaning a constant $C > 0$ exists such that

$$\|Tf\|_{L^q(Y)} \le C\|f\|_{L^p(X)}, \qquad f \in L^p(X).$$

When $q < \infty$, we call T **weak** $(\mathbf{p}, \mathbf{q})$ if $C > 0$ exists such that

$$\nu\big(\{y \in Y : |(Tf)(y)| > \lambda\}\big)^{1/q} \le \frac{C\|f\|_{L^p(X)}}{\lambda} \qquad \forall\ \lambda > 0, \quad f \in L^p(X).$$

When $q = \infty$, we call T **weak** $(\mathbf{p}, \infty)$ if it is strong (p, ∞):

$$\|Tf\|_{L^\infty(Y)} \le C\|f\|_{L^p(X)}, \qquad f \in L^p(X).$$

Lemma 6.2. *Strong* (p, q) $\Rightarrow$ *weak* (p, q).

Proof. When $q = \infty$ the result is immediate by definition. Suppose $q < \infty$. Write

$$E(\lambda) = \{y \in Y : |(Tf)(y)| > \lambda\}$$

for the level set of Tf above height λ. Then

$$\begin{aligned}
\lambda^q \, \nu\big(E(\lambda)\big) &= \int_{E(\lambda)} \lambda^q \, d\nu(y) \\
&\le \int_{E(\lambda)} |(Tf)(y)|^q \, d\nu(y) \qquad \text{since } \lambda < |Tf| \text{ on } E(\lambda) \\
&\le \|Tf\|^q_{L^q(Y)}
\end{aligned}$$

and so

$$\begin{aligned}
\nu\big(E(\lambda)\big)^{1/q} &\le \frac{\|Tf\|_{L^q(Y)}}{\lambda} \\
&\le \frac{C\|f\|_{L^p(X)}}{\lambda}
\end{aligned}$$

if T is strong (p, q). □

Lemma 6.3. *If* T *is weak* (p, q) *then* $Tf \in L^r_{loc}(Y)$ *for all* $0 < r < q$.

Thus intuitively, T "almost" maps L^p into L^q, locally.

Proof. Let $f \in L^p(X)$ and suppose $Z \subset Y$ with $\nu(Z) < \infty$. We will show $Tf \in L^r(Z)$.

Write $g = Tf$. Then

$$\begin{aligned}
&\int_Z |g(y)|^r \, d\nu(y) \\
&= \int_0^\infty r\lambda^{r-1} \nu\big(\{y \in Z : |g(y)| > \lambda\}\big) \, d\lambda \qquad \text{by AppendixB} \\
&\le \int_0^1 r\lambda^{r-1} \nu(Z) \, d\lambda + \int_1^\infty r\lambda^{r-1} \Big(\frac{C\|f\|_{L^p(X)}}{\lambda}\Big)^q \, d\lambda \qquad \text{by weak } (p, q) \\
&< \infty
\end{aligned}$$

since $\nu(Z) < \infty$ and $\int_1^\infty \lambda^{-1-q+r} \, d\lambda < \infty$ (using that $-q + r < 0$). □

Theorem 6.4 (Maximal functions and convergence a.e.)**.** *Assume*

$$T_n : L^p(X) \to \{\textit{measurable functions on } X\}$$

for $n = 1, 2, 3, \ldots$. *Define*

$$T^* : L^p(X) \to \{\textit{measurable functions on } X\}$$

by

$$(T^* f)(x) = \sup_n |(T_n f)|(x)|, \qquad x \in X.$$

If T^* *is weak* (p, q) *and each* T_n *is linear, then the collection*

$$\mathcal{C} = \{f \in L^p(X) : \lim_n (T_n f)(x) = f(x) \ a.e\}$$

is closed in $L^p(X)$.

T^* is called the **maximal operator** for the family $\{T_n\}$. Clearly it takes values in $[0, \infty]$. Note T^* is not linear, in general.

Remark 6.5. In this theorem a *quantitative* hypothesis (weak (p, q)) implies a *qualitative* conclusion (closure of the collection $\mathcal{C}$ where $T_n f \to f$ a.e.).

Proof. Let $f_k \in \mathcal{C}$ with $f_k \to f$ in $L^p(X)$. We show $f \in \mathcal{C}$.

Suppose $q < \infty$. For any $\lambda > 0$,

$$\begin{aligned}
&\mu\big(\{x \in X : \limsup_n |(T_n f(x) - f(x)| > 2\lambda\}\big) \\
&= \mu\big(\{x \in X : \limsup_n |T_n(f - f_k)(x) - (f - f_k)(x)| > 2\lambda\}\big) \\
&\qquad \text{by linearity and the pointwise convergence } T_n f_k \to f_k \text{ a.e.} \\
&\le \mu\big(\{x \in X : T^*(f - f_k)(x) + |(f - f_k)(x)| > 2\lambda\}\big) \qquad \text{by triangle inequality} \\
&\le \mu\big(\{x \in X : T^*(f - f_k)(x) > \lambda\}\big) + \mu\big(\{x \in X : |(f - f_k)(x)| > \lambda\}\big) \\
&\le \Big(\frac{C\|f - f_k\|_{L^p(X)}}{\lambda}\Big)^q + \Big(\frac{\|f - f_k\|_{L^p(X)}}{\lambda}\Big)^p \qquad \text{by weak } (p, q) \text{ on } T^* \\
&\to 0
\end{aligned}$$

as $k \to \infty$.

Therefore $\limsup_n |(T_n f)(x) - f(x)| \le 2\lambda$ a.e. Taking a countable sequence of $\lambda \searrow 0$, we conclude $\limsup_n |(T_n f)(x) - f(x)| = 0$ a.e. Therefore $\lim_n (T_n f)(x) = f(x)$ a.e., so that $f \in \mathcal{C}$.

The case $q = \infty$ is left to the reader. □

To apply maximal functions on $\mathbb{R}^d$ and $\mathbb{T}$, we will need:

Lemma 6.6 (Covering)**.** *Let $\{B_i\}_{i=1}^k$ be a finite collection of open balls in $\mathbb{R}^d$. Then there exists a pairwise disjoint subcollection $\{B_{i_j}\}_{j=1}^l$ of balls such that*

$$|\bigcup_{i=1}^{k} B_i| \leq 3^d |\bigcup_{j=1}^{l} B_{i_j}| = 3^d \sum_{j=1}^{l} |B_{i_j}|.$$

Thus the subcollection covers at least $1/3^d$ of the total volume of the balls.

Proof. Re-label the balls in decreasing order of size: $|B_1| \geq |B_2| \geq \cdots \geq |B_k|$. Choose $i_1 = 1$ and employ the following greedy algorithm. After choosing i_j, choose i_{j+1} to be the smallest index $i > i_j$ such that B_i is disjoint from $B_{i_1}, \ldots, B_{i_j}$. Continue until no such ball B_i exists.

The B_{i_j} are pairwise disjoint, by construction.

Let $i \in \{1, \ldots, k\}$. If B_i is *not* one of the B_{i_j} chosen, then B_i must intersect one of the B_{i_j} and be smaller than it, so that

$$\text{radius}\,(B_i) \leq \text{radius}(B_{i_j}).$$

Hence $B_i \subset 3B_{i_j}$ (where we mean the ball with the same center and three times the radius). Thus

$$\begin{aligned} |\bigcup_{i=1}^{k} B_i| &\leq |\bigcup_{j=1}^{l} (3B_{i_j})| \\ &\leq \sum_{j=1}^{l} |3B_{i_j}| \\ &= 3^d \sum_{j=1}^{l} |B_{i_j}| \\ &= 3^d |\bigcup_{j=1}^{l} B_{i_j}| \end{aligned}$$

by disjointness of the B_{i_j}. □

Definition 6.7. The *Hardy–Littlewood (H-L) maximal function* of a locally integrable function f on $\mathbb{R}^d$ is

$$\begin{aligned}(Mf)(x) &= \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy \\ &= \text{“largest local average” of } |f| \text{ around } x.\end{aligned}$$

Properties

$$\begin{aligned} & Mf \geq 0 \\ |f| \leq |g| \;\Rightarrow\; & Mf \leq Mg \\ & M(f+g) \leq Mf + Mg \quad \text{(sub-linearity)} \\ & Mc = c \quad \text{if } c = \text{(const.)} \geq 0 \end{aligned}$$

Theorem 6.8 (H-L maximal operator)**.** *M is weak $(1,1)$ and strong (p,p) for $1 < p \leq \infty$.*

Proof. For weak $(1,1)$ we show

$$|E(\lambda)| \leq \frac{3^d \|f\|_{L^1(\mathbb{R}^d)}}{\lambda} \tag{6.1}$$

where $E(\lambda) = \{x \in \mathbb{R}^d : Mf(x) > \lambda\}$. If $x \in E(\lambda)$ then

$$\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy > \lambda$$

for some $r > 0$. The same inequality holds for all x' close to x, so that $x' \in E(\lambda)$. Thus $E(\lambda)$ is open (and measurable), and Mf is lower semicontinuous (and measurable).

Let $F \subset E(\lambda)$ be compact. Each $x \in F$ is the center of some ball B such that

$$|B| < \frac{1}{\lambda} \int_B |f(y)|\, dy. \tag{6.2}$$

By compactness, F is covered by finitely many such balls, say $B_1, \ldots, B_k$.

The Covering Lemma 6.6 yields a subcollection $B_{i_1}, \ldots, B_{i_l}$. Then

$$\begin{aligned}
|F| &\le \Big|\bigcup_{i=1}^{k} B_i\Big| \\
&\le 3^d \sum_{j=1}^{l} |B_{i_j}| && \text{by Covering Lemma 6.6} \\
&\le \frac{3^d}{\lambda} \sum_{j=1}^{l} \int_{B_{i_j}} |f(y)|\, dy && \text{by (6.2)} \\
&\le \frac{3^d}{\lambda} \int_{\mathbb{R}^d} |f(y)|\, dy && \text{by disjointness} \\
&= \frac{3^d}{\lambda} \|f\|_{L^1(\mathbb{R}^d)}.
\end{aligned}$$

Taking the supremum over all compact $F \subset E(\lambda)$ gives (6.1).

For strong (∞, ∞), note $Mf(x) \le \|f\|_{L^\infty(\mathbb{R}^d)}$ for all $x \in \mathbb{R}^d$, by definition of Mf. Hence $\|Mf\|_{L^\infty(\mathbb{R}^d)} \le \|f\|_{L^\infty(\mathbb{R}^d)}$.

For strong (p, p) when $1 < p < \infty$, let $\lambda > 0$ and define

$$g(x) = \begin{cases} f(x) & \text{if } |f(x)| > \lambda/2 \\ 0 & \text{otherwise} \end{cases} = \text{“large” part of } f,$$

$$h(x) = \begin{cases} f(x) & \text{if } |f(x)| \le \lambda/2 \\ 0 & \text{otherwise} \end{cases} = \text{“small” part of } f.$$

Then $f = g + h$ and $|h| \le \lambda/2$, so that $Mf \le Mg + \lambda/2$. Hence

$$\begin{aligned}
|E(\lambda)| &= \big|\{x : Mf(x) > \lambda\}\big| \\
&\le \big|\{x : Mg(x) > \lambda/2\}\big| \\
&\le \frac{3^d \|g\|_{L^1(\mathbb{R}^d)}}{\lambda/2} && \text{by the above weak } (1,1) \text{ result} \\
&= \frac{2 \cdot 3^d}{\lambda} \int_{\{x : |f(x)| > \lambda/2\}} |f(x)|\, dx. && (6.3)
\end{aligned}$$

Therefore

$$\begin{aligned}
&\int_{\mathbb{R}^d} |Mf(x)|^p \, dx \\
&= \int_0^\infty p\lambda^{p-1} |E(\lambda)| \, d\lambda && \text{by Appendix B} \\
&\le 2 \cdot 3^d p \int_0^\infty \lambda^{p-2} \int_{\{x:|f(x)|>\lambda/2\}} |f(x)| \, dx d\lambda && \text{by (6.3)} \\
&= \left(2^p 3^d \frac{p}{p-1}\right) \int_{\mathbb{R}^d} |f(x)|^p \, dx
\end{aligned}$$

by Lemma B.1 with $r = 1, \alpha = 2$. We have proved the strong (p, p) bound. □

Notice the constant in the strong (p, p) bound blows up as $p \searrow 1$. As this observation suggests, the Hardy–Littlewood maximal operator is not strong $(1, 1)$. For example, the indicator function $f = \mathbb{1}_{[-1,1]}$ in 1 dimension has $Mf(x) \sim c/|x|$ when $|x|$ is large, so that $Mf \notin L^1(\mathbb{R})$.

The maximal function is *locally* integrable provided $f \in L \log L(\mathbb{R}^d)$; see Problem 9.

Chapter 7

Fourier series: summability pointwise a.e.

Goal

Prove summability a.e. using Fejér and Poisson maximal functions

Definition 7.1.

$$\begin{aligned}
\text{Dirichlet maximal function} \quad & (D^*f)(t) = \sup_n |(D_n * f)(t)| = \sup_n |S_n(f)(t)| \\
\text{Fejér maximal function} \quad & (F^*f)(t) = \sup_n |(F_n * f)(t)| = \sup_n |\sigma_n(f)(t)| \\
\text{Poisson maximal function} \quad & (P^*f)(t) = \sup_{0<r<1} |(P_r * f)(t)| \\
\text{Gauss maximal function} \quad & (G^*f)(t) = \sup_{0<s<\infty} |(G_s * f)(t)| \\
\text{Lebesgue maximal function} \quad & (L^*f)(t) = \sup_{0<h<\pi} |(L_h * f)(t)|
\end{aligned}$$

where the Lebesgue kernel is $L_h(t) = 2\pi \frac{1}{2h} \mathbb{1}_{[-h,h]}(t)$, extended 2π-periodically. Notice $(L_h * f)(t) = \frac{1}{2h} \int_{t-h}^{t+h} f(\tau)\, d\tau$ is a local average of f around t.

Lemma 7.2 (Majorization)**.** *If $k \in L^1(\mathbb{T})$ is nonnegative and symmetric ($k(-t) = k(t)$), and decreasing on $[0, \pi]$, then*

$$|(k * f)(t)| \le \|k\|_{L^1(\mathbb{T})} (L^*f)(t) \qquad \text{for all } t \in \mathbb{T}, \quad f \in L^1(\mathbb{T}).$$

Thus convolution with a symmetric decreasing kernel is majorized by the Hardy–Littlewood maximal function.

Proof. Assume k is absolutely continuous, for simplicity. We first establish a "layer cake" decomposition of k, representing it as a linear combination of kernels L_h:

$$\begin{aligned} k(t) = k(|t|) &= k(\pi) - \int_{|t|}^{\pi} k'(h)\,dh \\ &= k(\pi) - \frac{1}{2\pi}\int_0^{\pi} 2hL_h(t)k'(h)\,dh, \end{aligned}$$

since

$$\frac{1}{2\pi}2hL_h(t) = \begin{cases} 1, & \text{if } h \geq |t|, \\ 0, & \text{if } h < |t|. \end{cases}$$

Hence

$$\begin{aligned} (k*f)(t) &= k(\pi)\frac{1}{2\pi}\int_{\mathbb{T}} f(\tau)\,d\tau + \frac{1}{2\pi}\int_0^{\pi} 2h(L_h * f)(t)\big(-k'(h)\big)\,dh \\ |(k*f)(t)| &\leq k(\pi)|(L_\pi * f)(t)| + \frac{1}{2\pi}\int_0^{\pi} 2h\big(-k'(h)\big)\,dh\,(L^*f)(t) \\ &\qquad\qquad \text{using } k(\pi) \geq 0 \text{ and } k' \leq 0 \\ &\leq \frac{2}{2\pi}\int_0^{\pi} k(h)\,dh\,(L^*f)(t) \qquad \text{by parts} \\ &= \|k\|_{L^1(\mathbb{T})}(L^*f)(t) \end{aligned}$$

by symmetry of k. □

Theorem 7.3 (Lebesgue dominates Fejér and Poisson)**.** *For all $f \in L^1(\mathbb{T})$,*

$$\begin{aligned} F^*f &\leq 2L^*|f| \\ P^*f &\leq \ L^*f \end{aligned}$$

Proof. $P_r(t)$ is nonnegative, symmetric, and decreasing on $[0,\pi]$ (exercise), with $\|P_r\|_{L^1(\mathbb{T})} = 1$. Hence $|P_r * f| \leq L^*f$ by Majorization Lemma 7.2, so that $P^*f \leq L^*f$.

The Dirichlet kernel is *not* decreasing on $[0,\pi]$, but it is bounded by a symmetric decreasing kernel, as follows:

$$\begin{aligned} F_n(t) &= \frac{1}{n+1}\left(\frac{\sin\left(\frac{n+1}{2}t\right)}{\sin\left(\frac{1}{2}t\right)}\right)^2 \\ &\le k(t) \stackrel{\text{def}}{=} \frac{1}{n+1}\begin{cases}(n+1)^2 & \text{if } |t| \le \pi/(n+1),\\ \pi^2/t^2 & \text{if } \pi/(n+1) \le |t| \le \pi,\end{cases} \end{aligned}$$

since

$$\begin{aligned} \sin(n+1)\theta &\le (n+1)\sin\theta, && 0 \le \theta \le \frac{\pi}{2},\\ \sin\left(\frac{1}{2}t\right) &\ge \frac{t}{\pi}, && 0 \le t \le \pi. \end{aligned}$$

Note the kernel k is nonnegative, symmetric, and decreasing on $[0,\pi]$, with

$$\|k\|_{L^1(\mathbb{T})} = \frac{1}{2\pi}\left(4\pi - \frac{2\pi}{n+1}\right) < 2.$$

Hence $|F_n * f| \le k * |f| \le 2L^*|f|$ by Majorization Lemma 7.2, so that $F^*f \le 2L^*|f|$. □

The Gauss kernel can be shown to be symmetric decreasing, so that $G^*f \le L^*f$, but we omit the proof.

Corollary 7.4. *F^*, P^* and L^* are weak $(1,1)$ on $\mathbb{T}$.*

Proof.

$$\left|\{t \in \mathbb{T} : (L^*f)(t) > 2\pi\lambda\}\right| \le \left|\{t \in \mathbb{T} : (L^*|f|)(t) > 2\pi\lambda\}\right| \le \frac{3}{\lambda}\int_{\mathbb{T}} |f(t)|\,dt$$

by repeating the weak $(1,1)$ proof for the Hardy–Littlewood maximal function. These weak $(1,1)$ estimates for L^*f and $L^*|f|$ imply weak $(1,1)$ for F^*f, since if $(F^*f)(t) > \lambda$ then $(L^*|f|)(t) > \lambda/2$ by Theorem 7.3. Argue similarly for P^*f. □

Theorem 7.5 (Summability a.e.)**.** *If $f \in L^1(\mathbb{T})$ then*

$$\begin{aligned} \sigma_n(f) = F_n * f \to f \ \text{a.e.} \quad &\text{as } n \to \infty \quad \textit{(Fejér summability)}\\ P_r * f \to f \ \text{a.e.} \quad &\text{as } r \nearrow 1 \quad \textit{(Abel summability)}\\ L_h * f \to f \ \text{a.e.} \quad &\text{as } h \searrow 0 \quad \textit{(Lebesgue differentiation theorem)} \end{aligned}$$

Proof. By the weak $(1,1)$ estimate in Corollary 7.4 and the abstract convergence result in Theorem 6.4, the set

$$\mathcal{C} = \{f \in L^1(\mathbb{T}) : \lim_n (F_n * f)(t) = f(t) \text{ a.e.}\}$$

is closed in $L^1(\mathbb{T})$.

Obviously $\mathcal{C}$ contains the continuous functions on $\mathbb{T}$, since $F_n * f \to f$ uniformly when f is continuous. Thus $\mathcal{C}$ is dense in $L^1(\mathbb{T})$. Because $\mathcal{C}$ is also closed, it must equal $L^1(\mathbb{T})$, thus proving Fejér summability a.e. for each $f \in L^1(\mathbb{T})$.

Argue similarly for $P_r * f$ and $L_h * f$. □

The result that $L_h * f \to f$ a.e. means

$$\frac{1}{2h} \int_{t-h}^{t+h} f(\tau)\, d\tau \to f(t) \quad \text{a.e.},$$

which is the Lebesgue differentiation theorem on $\mathbb{T}$.

Chapter 8

Fourier series: convergence at a point

Goals

State divergence pointwise can occur for $L^1(\mathbb{T})$
Show divergence pointwise can occur for $C(\mathbb{T})$
Prove convergence pointwise for $C^\alpha(\mathbb{T})$ and $BV(\mathbb{T})$

References

[Katznelson] Section II.2, II.3
[Duoandikoetxea] Section 1.1

Fourier series can behave badly for integrable functions.

Theorem 8.1 (Kolmogorov)**.** *There exists $f \in L^1(\mathbb{T})$ whose Fourier series diverges unboundedly at* every *point. That is,*

$$\sup_n |S_n(f)(t)| = \infty \qquad \text{for all } t \in \mathbb{T},$$

so that $D^* f \equiv \infty$.

Recall $S_n(f) = D_n * f$ and $D^* f$ is the maximal function for the Dirichlet kernel.

Proof. [Katznelson, Section II.3]. □

Even continuous functions can behave badly.

Theorem 8.2. *There exists a continuous function whose Fourier series diverges unboundedly at $t = 0$. That is,*

$$\sup_n |S_n(f)(0)| = \infty.$$

Proof. Define

$$\begin{aligned} T_n : C(\mathbb{T}) &\to \mathbb{C} \\ f &\mapsto S_n(f)(0) = (n\text{th partial sum of } f \text{ at } t = 0). \end{aligned}$$

Then T_n is linear. Each T_n is bounded since

$$\begin{aligned} |T_n(f)| &= |S_n(f)(0)| \\ &= |(D_n * f)(0)| \\ &= \Big|\frac{1}{2\pi}\int_{\mathbb{T}} D_n(\tau) f(0-\tau)\, d\tau\Big| \\ &\le \|D_n\|_{L^1(\mathbb{T})}\|f\|_{L^\infty(\mathbb{T})}. \end{aligned}$$

Thus $\|T_n\| \le \|D_n\|_{L^1(\mathbb{T})}$. We show $\|T_n\| = \|D_n\|_{L^1(\mathbb{T})}$. Let $\varepsilon > 0$ and choose $g \in C(\mathbb{T})$ with $\|g\|_{L^\infty(\mathbb{T})} = 1$ and g even and

$$g(t) = \begin{cases} 1 & \text{if } D_n(t) > 0, \\ -1 & \text{if } D_n(t) < 0, \\ & \text{except for small intervals around the zeros of } D_n, \\ & \text{with total length of those intervals} < \varepsilon/(2n+1). \end{cases}$$

Then

$$\begin{aligned} |T_n(g)| &= \Big|\frac{1}{2\pi}\int_{\mathbb{T}} D_n(\tau) g(\tau)\, d\tau\Big| \\ &\ge \frac{1}{2\pi}\int_{\mathbb{T}\setminus\{\text{intervals}\}} |D_n(\tau)|\, d\tau - \frac{1}{2\pi}\int_{\{\text{intervals}\}} |D_n(\tau)|\, d\tau \\ &= \frac{1}{2\pi}\int_{\mathbb{T}} |D_n(\tau)|\, d\tau - \frac{2}{2\pi}\int_{\{\text{intervals}\}} |D_n(\tau)|\, d\tau \\ &\ge \|D_n\|_{L^1(\mathbb{T})} - \frac{1}{\pi}\frac{\varepsilon}{2n+1}(2n+1) \qquad \text{using definition (2.1) of } D_n \\ &= \|D_n\|_{L^1(\mathbb{T})} - \frac{\varepsilon}{\pi} \\ &= \big(\|D_n\|_{L^1(\mathbb{T})} - \frac{\varepsilon}{\pi}\big)\|g\|_{L^\infty(\mathbb{T})}. \end{aligned}$$

Thus $\|T_n\| \geq \|D_n\|_{L^1(\mathbb{T})} - \varepsilon/\pi$ for all $\varepsilon > 0$, and so $\|T_n\| = \|D_n\|_{L^1(\mathbb{T})}$.

Recalling that $\|D_n\|_{L^1(\mathbb{T})} \to \infty$ as $n \to \infty$ (in fact, $\|D_n\| \sim c \log n$ by [Katznelson, Ex. II.1.1]) we conclude from the Uniform Bounded Principle (Banach–Steinhaus) that there exists $f \in C(\mathbb{T})$ with $\sup_n |T_n(f)| = \infty$, as desired. □

Another proof. [Katznelson, Sec II.2] gives an explicit construction of f, proving divergence not only at $t = 0$ but on a dense set of t-values.

Now we prove convergence results.

Theorem 8.3 (Dini's Convergence Test)**.** *Let $f \in L^1(\mathbb{T}), t \in \mathbb{T}$. If*

$$\int_{-\pi}^{\pi} \left| \frac{f(t-\tau) - f(t)}{\tau} \right| d\tau < \infty$$

then the Fourier series of f converges at t to $f(t)$.

Proof.

$$\begin{aligned} S_n(f)(t) - f(t) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(t-\tau) - f(t)] \frac{\sin\left((n+\frac{1}{2})\tau\right)}{\sin\left(\frac{1}{2}\tau\right)} \, d\tau \\ &\qquad \text{using that } \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(\tau) \, d\tau = 1 \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{f(t-\tau) - f(t)}{\tau} \frac{\tau}{\sin\left(\frac{1}{2}\tau\right)} \cos\left(\frac{1}{2}\tau\right) \right\} \sin(n\tau) \, d\tau \\ &\qquad + \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(t-\tau) - f(t)] \cos(n\tau) \, d\tau \end{aligned} \tag{8.1}$$

by expanding $\sin\left((n+\frac{1}{2})\tau\right)$ with a trigonometric identity.

Notice the factor $\{\cdots\}$ is integrable with respect to τ, by the Dini hypothesis. And $\tau \mapsto [f(t-\tau) - f(t)]$ is integrable too. Hence both integrals in (8.1) tend to 0 as $n \to \infty$, by the Riemann–Lebesgue Corollary 1.5 (after expressing $\sin(n\tau)$ and $\cos(n\tau)$ in terms of $e^{\pm in\tau}$). □

Corollary 8.4 (Convergence for Hölder continuous f)**.** *If $f \in C^\alpha(\mathbb{T}), 0 < \alpha \leq 1$, then the Fourier series of f converges to $f(t)$, for every $t \in \mathbb{T}$.*

Proof. Put Hölder into Dini:

$$\int_{-\pi}^{\pi} \left| \frac{f(t-\tau) - f(t)}{\tau} \right| d\tau \leq \int_{-\pi}^{\pi} \frac{(\text{const.})|\tau|^{\alpha}}{|\tau|}\, d\tau < \infty.$$

Now apply Dini's Theorem 8.3.

(Exercise. Prove the Fourier series in fact converges uniformly.) □

Corollary 8.5 (Localization Principle)**.** *Let $f \in L^1(\mathbb{T}), t \in \mathbb{T}$. If f vanishes on a neighborhood of t, then $S_n(f)(t) \to 0$ as $n \to \infty$.*

Proof. Apply Dini's Theorem 8.3. □

In particular, if two functions agree on a neighborhood of t and the Fourier series of one of them converges at t, then the Fourier series of the other function converges at t to the same value. Thus Fourier series depend only on local information.

Theorem 8.6 (Convergence for bounded variation f)**.** *If $f \in BV(\mathbb{T})$ then the Fourier series converges everywhere to $\frac{1}{2}[f(t+)+f(t-)]$, and hence converges to $f(t)$ at every point of continuity.*

Proof. Let $t \in \mathbb{T}$. On the interval $(t-\pi, t+\pi)$, express f as the difference of two bounded increasing functions, say $f = g - h$. It suffices to prove the theorem for g and h individually.

We have

$$S_n(g)(t) - \frac{1}{2}[g(t+) + g(t-)] = \frac{1}{2\pi}\int_0^{\pi} \big(g(t-\tau) - g(t-)\big) D_n(\tau)\, d\tau \tag{8.2}$$

$$+ \frac{1}{2\pi}\int_0^{\pi} \big(g(t+\tau) - g(t+)\big) D_n(\tau)\, d\tau \tag{8.3}$$

since $D_n(\tau)$ is even and $\frac{1}{2\pi}\int_0^{\pi} D_n(\tau)\, d\tau = \frac{1}{2}$.

Let $G(\tau) = g(t+\tau) - g(t+)$ for $\tau \in (0, \pi)$, so that G is increasing with $G(0+) = 0$. Write

$$H_n(\tau) = \int_0^{\tau} D_n(\sigma)\, d\sigma$$

so that $H_n' = D_n$. Let $0 < \delta < \pi$. Then

$$\begin{aligned}(8.3) &= \frac{1}{2\pi}\int_0^{\delta} G(\tau) H_n'(\tau)\, d\tau + \frac{1}{2\pi}\int_{\delta}^{\pi} G(\tau) D_n(\tau)\, d\tau \\ &= \frac{1}{2\pi} G(\delta) H_n(\delta) - \frac{1}{2\pi}\int_{(0,\delta]} H_n(\tau)\, dG(\tau) + o(1)\end{aligned}$$

as $n \to \infty$, by parts in the first term and by the Localization Principle in the last term, since the function

$$\begin{cases} G(\tau), & \delta < \tau < \pi, \\ 0, & -\delta < \tau < \delta, \\ G(-\tau), & -\pi < \tau < -\delta, \end{cases}$$

vanishes near the origin. Hence

$$\begin{aligned} \limsup_n |(8.3)| &\le \frac{1}{2\pi} \sup_n \|H_n\|_{L^\infty(\mathbb{T})} \left(G(\delta) + \int_{(0,\delta]} dG(\tau) \right) \\ &= \frac{1}{2\pi} \sup_n \|H_n\|_{L^\infty(\mathbb{T})} \cdot 2G(\delta) \qquad \text{since } G(0+) = 0 \\ &\to 0 \end{aligned}$$

as $\delta \to 0$. Therefore (8.3) $\to 0$ as $n \to \infty$. Argue similarly for (8.2), and for h.

Thus we are done, provided we show

$$\sup_n \|H_n\|_{L^\infty(\mathbb{T})} < \infty.$$

We have

$$\begin{aligned} |H_n(\tau)| &\le \Big| \int_0^\tau \frac{\sin\big((n+\frac{1}{2})\sigma\big)}{\frac{1}{2}\sigma}\, d\sigma \Big| + \Big| \int_0^\tau \sin\big((n+\frac{1}{2})\sigma\big) \Big(\frac{1}{\sin\big(\frac{1}{2}\sigma\big)} - \frac{1}{\frac{1}{2}\sigma} \Big) d\sigma \Big| \\ &\le 2 \Big| \int_0^{(n+\frac{1}{2})\tau} \frac{\sin\sigma}{\sigma}\, d\sigma \Big| + \int_0^\pi (\text{const.}) \frac{\sigma^3}{\sigma^2}\, d\sigma \qquad \text{by a change of variable} \\ &\le 2 \sup_{\rho > 0} \Big| \int_0^\rho \frac{\sin\sigma}{\sigma}\, d\sigma \Big| + (\text{const.}) \\ &< \infty \end{aligned}$$

since $\lim_{\rho\to\infty} \int_0^\rho \frac{\sin\sigma}{\sigma}\, d\sigma$ exists.

□

The convergence results so far in this chapter rely just on Riemann–Lebsgue and direct estimates. A much deeper result is:

Theorem 8.7 (Carleson–Hunt)**.** *If $f \in L^p(\mathbb{T}), 1 < p < \infty$ then the Fourier series of f converges to $f(t)$ for almost every $t \in \mathbb{T}$.*

For $p = 1$, the result is spectacularly false by Kolmogorov's Theorem 8.1.

Proof. Omitted. The idea is to prove that the Dirichlet maximal operator $(D^* f)(t) = \sup_n |(D_n * f)(t)|$ is strong (p, p) for $1 < p < \infty$. Then it is weak (p, p), and so convergence a.e. follows from Chapter 6.

Thus one wants

$$\big\| \sup_n |D_n * f| \big\|_{L^p(\mathbb{T})} \leq C_p \|f\|_{L^p(\mathbb{T})}$$

for $1 < p < \infty$. The next Chapters show

$$\sup_n \|D_n * f\|_{L^p(\mathbb{T})} \leq C_p \|f\|_{L^p(\mathbb{T})},$$

but that is not good enough to prove Carleson–Hunt! □

Chapter 9

Fourier series: norm convergence

Goals

Characterize norm convergence in terms of uniform norm bounds
Show norm divergence can occur for $L^1(\mathbb{T})$ and $C(\mathbb{T})$
Show norm convergence for $L^p(\mathbb{T})$ follows from boundedness of the Hilbert transform

Reference

[Katznelson] Section II.1

Theorem 9.1. *Let B be one of the spaces $C(\mathbb{T})$ or $L^p(\mathbb{T}), 1 \le p < \infty$.*
(a) If $\sup_n \|S_n\|_{B\to B} < \infty$ then Fourier series converge in B:

$$\lim_{n\to\infty} \|S_n(f) - f\|_B = 0 \qquad \text{for each } f \in B.$$

(b) If $\sup_n \|S_n\|_{B\to B} = \infty$ then there exists $f \in B$ whose Fourier series diverges unboundedly: $\sup_n \|S_n(f)\|_B = \infty$.

Proof.
(b) This part follows immediately from the Uniform Boundedness Principle in functional analysis.

(a) The collection of trigonometric polynomials is dense in B (as remarked after Theorem 2.6). Further, if g is a trigonometric polynomial then $S_n(g) = g$ whenever n exceeds the degree of g. Hence the set

$$\mathcal{C} = \{f \in B : \lim_{n\to\infty} S_n(f) = f \text{ in } B\}$$

is dense in B. The set $\mathcal{C}$ is also closed, by the following proposition, and so $\mathcal{C} = B$, which proves part (a). □

Proposition 9.2. *Let B be any Banach space and assume the $T_n : B \to B$ are bounded linear operators.*

If $\sup_n \|T_n\|_{B\to B} < \infty$ then

$$\mathcal{C} = \{f \in B : \lim_{n\to\infty} T_n f = f \text{ in } B\}$$

is closed.

Proof. Let $A = \sup_n \|T_n\|_{B\to B}$. Consider a sequence $f_m \in \mathcal{C}$ with $f_m \to f$. We must show $f \in \mathcal{C}$, so that $\mathcal{C}$ is closed.

Choose $\varepsilon > 0$ and fix m such that $\|f_m - f\| < \varepsilon/2(A+1)$. Since $f_m \in \mathcal{C}$ there exists N such that $\|T_n f_m - f_m\| < \varepsilon/2$ whenever $n > N$. Then

$$\begin{aligned}\|T_n f - f\| &\le \|T_n f - T_n f_m\| + \|T_n f_m - f_m\| + \|f_m - f\| \\ &\le (A+1)\|f - f_m\| + \|T_n f_m - f_m\| < \varepsilon\end{aligned}$$

whenever $n > N$, as desired. □

Norm Estimates

$$\|S_n\|_{B\to B} \le \|D_n\|_{L^1(\mathbb{T})}$$

when B is $C(\mathbb{T})$ or $L^p(\mathbb{T}), 1 \le p < \infty$, since

$$\|S_n(f)\|_B = \big\|D_n * f\big\|_B \le \|D_n\|_{L^1(\mathbb{T})}\|f\|_B.$$

This upper estimate is not useful, since we know $\|D_n\|_{L^1(\mathbb{T})} \to \infty$.

Example 9.3 (Divergence in $C(\mathbb{T})$)**.** For $B = C(\mathbb{T})$ we have

$$\|S_n\|_{C(\mathbb{T})\to C(\mathbb{T})} = \|D_n\|_{L^1(\mathbb{T})}.$$

Indeed, for each $\varepsilon > 0$ one can construct $g \in C(\mathbb{T})$ that approximates $\operatorname{sign}(D_n)$ (like in Chapter 8), so that

$$\|S_n(g)\|_{C(\mathbb{T})} \geq |S_n(g)(0)| \geq \big(\|D_n\|_{L^1(\mathbb{T})} - \varepsilon\big)\|g\|_{C(\mathbb{T})}.$$

Therefore $\sup_n \|S_n\|_{C(\mathbb{T})\to C(\mathbb{T})} = \infty$, so that (by Theorem 9.1(b)) there exists a continuous function $f \in C(\mathbb{T})$ whose Fourier series diverges unboundedly in the uniform norm: $\sup_n \|S_n(f)\|_{C(\mathbb{T})} = \infty$.

Of course, this result follows already from the pointwise divergence in Chapter 8.

Example 9.4 (Divergence in $L^1(\mathbb{T})$)**.** For $B = L^1(\mathbb{T})$ we have

$$\|S_n\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})} = \|D_n\|_{L^1(\mathbb{T})}.$$

Proof. Fix n. Then $S_n(F_N) = F_N * D_n \to D_n$ in $L^1(\mathbb{T})$ as $N \to \infty$, and so

$$\begin{aligned}\|D_n\|_{L^1(\mathbb{T})} &= \lim_{N\to\infty} \|S_n(F_N)\|_{L^1(\mathbb{T})} \\ &\leq \|S_n\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})}\|F_N\|_{L^1(\mathbb{T})} \\ &= \|S_n\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})}.\end{aligned}$$

Therefore $\sup_n \|S_n\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})} = \infty$, so that (by Theorem 9.1(b)) there exists an integrable function $f \in L^1(\mathbb{T})$ whose Fourier series diverges unboundedly in the L^1 norm: $\sup_n \|S_n(f)\|_{L^1(\mathbb{T})} = \infty$.

Aside. For an explicit example of L^1 divergence, see [Grafakos, Exercise 3.5.9].

Convergence in $L^p(\mathbb{T}), 1 < p < \infty$

1. We shall prove (in Chapters 10–12) the existence of a bounded linear operator

$$H : L^p(\mathbb{T}) \to L^p(\mathbb{T}), \qquad 1 < p < \infty,$$

called the **Hilbert transform** on $\mathbb{T}$, with the property

$$\widehat{(Hf)}(n) = -i \operatorname{sign}(n) \widehat{f}(n).$$

(Thus H is a *Fourier multiplier* operator.) That is

$$Hf \sim \sum_{n=-\infty}^{\infty} (-i) \operatorname{sign}(n) \widehat{f}(n) e^{int}.$$

2. Then the *Riesz projection* $P : L^p(\mathbb{T}) \to L^p(\mathbb{T})$ defined by

$$Pf = \frac{1}{2}\widehat{f}(0) + \frac{1}{2}(f + iHf)$$

is also bounded, when $1 < p < \infty$. (Note the constant term $\widehat{f}(0)$ is bounded by $\|f\|_{L^p(\mathbb{T})}$, by Hölder's inequality.)

Observe P projects onto the nonnegative frequencies:

$$Pf \sim \sum_{n \geq 0} \widehat{f}(n) e^{int}$$

since $i(-i \operatorname{sign}(n)) = \operatorname{sign}(n)$.

3. The following formula expresses the Fourier partial sum operator in terms of the Riesz projection and some modulations:

$$e^{-imt} P(e^{imt} f) - e^{i(m+1)t} P(e^{-i(m+1)t} f) = S_m(f). \tag{9.1}$$

Proof.

$$\begin{aligned} e^{imt} f &\sim \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{i(m+n)t} \\ P(e^{imt} f) &\sim \sum_{n \geq -m} \widehat{f}(n) e^{i(m+n)t} \\ e^{-imt} P(e^{imt} f) &\sim \sum_{n \geq -m} \widehat{f}(n) e^{int} \\ e^{i(m+1)t} P(e^{-i(m+1)t} f) &\sim \sum_{n \geq m+1} \widehat{f}(n) e^{int} \end{aligned}$$

Subtracting the last two formulas gives $S_m(f)$, on the right side, and we conclude that the left side of (9.1) has the same Fourier coefficients as $S_m(f)$. By the uniqueness result (2.14), the left side of (9.1) must equal $S_m(f)$.

4. From (9.1) and boundedness of the Riesz projection it follows that

$$\sup_m \|S_m\|_{L^p(\mathbb{T}) \to L^p(\mathbb{T})} \leq 2\|P\|_{L^p(\mathbb{T}) \to L^p(\mathbb{T})} < \infty$$

when $1 < p < \infty$. Hence from Theorem 9.1 we conclude:

Theorem 9.5 (Fourier series converge in $L^p(\mathbb{T})$)**.** *Let $1 < p < \infty$. Then*

$$\lim_{n \to \infty} \|S_n(f) - f\|_{L^p(\mathbb{T})} = 0 \qquad \textit{for each } f \in L^p(\mathbb{T}).$$

It remains to prove L^p boundedness of the Hilbert transform.

Chapter 10

Hilbert transform on $L^2(\mathbb{T})$

Goal

Obtain time and frequency representations of the Hilbert transform

Reference

[Edwards and Gaudry] Section 6.3

Definition 10.1. The *Hilbert transform* on $L^2(\mathbb{T})$ is

$$\begin{aligned} H : L^2(\mathbb{T}) &\to L^2(\mathbb{T}) \\ f &\mapsto \sum_{n=-\infty}^{\infty} \big(-i \operatorname{sign}(n) \widehat{f}(n) \big) e^{int}. \end{aligned}$$

We call $\{-i \operatorname{sign}(n)\}$ the *multiplier sequence* of H.
Since $|\operatorname{sign}(n)| \leq 1$, the definition indeed yields $Hf \in L^2(\mathbb{T})$, with

$$\|Hf\|^2_{L^2(\mathbb{T})} = \sum_{n\in\mathbb{Z}} |\widehat{(Hf)}(n)|^2 = \sum_{n\neq 0} |\widehat{f}(n)|^2 \leq \|\widehat{f}\|^2_{\ell^2(\mathbb{Z})} = \|f\|^2_{L^2(\mathbb{T})}$$

by Plancherel in Chapter 5. Hence $\|H\|_{L^2\to L^2} = 1$. Observe also $H^2(f) = H(Hf) = -\sum_{n\neq 0} \widehat{f}(n) e^{int} = -f + \widehat{f}(0)$.

Lemma 10.2 (Adjoint of Hilbert transform)**.** $H^* = -H$

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

$$\begin{aligned}\langle Hf, g\rangle_{L^2(\mathbb{T})} &= \langle \widehat{Hf}, \widehat{g}\rangle_{\ell^2(\mathbb{Z})} \\ &= \langle -i\,\mathrm{sign}(n)\widehat{f}(n), \widehat{g}(n)\rangle_{\ell^2(\mathbb{Z})} \\ &= \langle \widehat{f}(n), i\,\mathrm{sign}(n)\widehat{g}(n)\rangle_{\ell^2(\mathbb{Z})} \\ &= \langle \widehat{f}, -\widehat{Hg}\rangle_{\ell^2(\mathbb{Z})} \\ &= \langle f, -Hg\rangle_{L^2(\mathbb{T})}.\end{aligned}$$

□

Proposition 10.3. *If $f \in L^2(\mathbb{T})$ is C^1-smooth on an open interval $I \subset \mathbb{T}$, then*

$$(Hf)(t) = \frac{1}{2\pi}\int_0^\pi [f(t-\tau) - f(t+\tau)]\cot\left(\frac{\tau}{2}\right) d\tau \tag{10.1}$$

$$= \lim_{\varepsilon\to 0}\frac{1}{2\pi}\int_{\varepsilon<|\tau|<\pi} f(t-\tau)\cot\left(\frac{\tau}{2}\right) d\tau \tag{10.2}$$

for almost every $t \in I$.

Remark 10.4. Formally (10.2) says that

$$Hf = f * \cot\left(\frac{t}{2}\right).$$

But the convolution is ill-defined because the Hilbert kernel $\cot(t/2)$ is not integrable. That is why (10.2) evaluates the convolution in the *principal valued* sense, taking the limit of integrals over $\mathbb{T}\setminus[-\varepsilon,\varepsilon]$.

Proof. First, geometric series calculations show that

$$\begin{aligned}\sum_{n=-N}^{N}\big(-i\,\mathrm{sign}(n)e^{in\tau}\big) &= i\sum_{n=-N}^{-1} e^{in\tau} - i\sum_{n=1}^{N} e^{in\tau} \\ &= i\frac{e^{-i(N+1)\tau} - e^{-i\tau}}{e^{-i\tau}-1} - i\frac{e^{i(N+1)\tau}-e^{i\tau}}{e^{i\tau}-1} \\ &= i\frac{e^{-i(N+1/2)\tau} - e^{-i\tau/2} + e^{i(N+1/2)\tau} - e^{i\tau/2}}{e^{-i\tau/2}-e^{i\tau/2}} \\ &= \frac{\cos\left(\frac{\tau}{2}\right) - \cos\left((N+\frac{1}{2})\tau\right)}{\sin(\frac{\tau}{2})}.\end{aligned} \tag{10.3}$$

Second, the Nth partial sum of Hf is

$$\begin{aligned}
&\sum_{n=-N}^{N} \big(-i\,\mathrm{sign}(n)\widehat{f}(n)\big)e^{in\tau} \\
&= \frac{1}{2\pi}\int_{\mathbb{T}} f(\tau) \sum_{n=-N}^{N} (-i)\,\mathrm{sign}(n)e^{in(t-\tau)}\,d\tau \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-\tau) \sum_{n=-N}^{N} (-i)\,\mathrm{sign}(n)e^{in\tau}\,d\tau \qquad \text{by } \tau \mapsto t-\tau \\
&= \frac{1}{2\pi}\int_{0}^{\pi} [f(t-\tau)-f(t+\tau)]\frac{\cos\left(\frac{\tau}{2}\right)-\cos\left((N+\frac{1}{2})\tau\right)}{\sin(\frac{\tau}{2})}\,d\tau \\
&= \frac{1}{2\pi}\int_{0}^{\pi} [f(t-\tau)-f(t+\tau)]\cot\left(\frac{\tau}{2}\right)\,d\tau \qquad \text{by (10.3)} \\
&\quad - \frac{1}{2\pi}\int_{0}^{\pi} \frac{f(t-\tau)-f(t+\tau)}{\sin(\frac{\tau}{2})}\cos\left((N+\frac{1}{2})\tau\right)\,d\tau.
\end{aligned}$$

If $t \in I$ then the second integrand belongs to $L^1(\mathbb{T})$ since it is bounded for τ near 0, by the C^1-smoothness of f. Hence the second integral tends to 0 as $N \to \infty$ by the Riemann-Lebesgue Corollary 1.5. Formula (10.1) now follows, because the partial sum

$$\sum_{n=-N}^{N} \left(-i\,\mathrm{sign}(n)\widehat{f}(n)\right)e^{in\tau}$$

converges to $Hf(t)$ in $L^2(\mathbb{T})$ and hence some subsequence of the partial sums converges to $(Hf)(t)$ a.e.

Now write (10.1) as

$$(Hf)(t) = \lim_{\varepsilon\to 0}\frac{1}{2\pi}\int_{\varepsilon}^{\pi} [f(t-\tau)-f(t+\tau)]\cot\left(\frac{\tau}{2}\right)\,d\tau$$

and use oddness of $\cot(\tau/2)$ to obtain (10.2).

Chapter 11

Calderón–Zygmund decompositions

Goal

Decompose a function into good and bad parts, preparing for a weak $(1,1)$ estimate on the Hilbert transform

References

[Duoandikoetxea] Section 2.5
[Grafakos] Section 4.3

Definition 11.1. For $k \in \mathbb{Z}$, let

$$Q_k = \{2^{-k}\big([0,1)^d + m\big) : m \in \mathbb{Z}^d\}.$$

Notice the cubes in Q_k are small when k is large.

Call $\cup_k Q_k$ the *collection of dyadic cubes.*

Facts (exercise)

1. For all $x \in \mathbb{R}^d$ and $k \in \mathbb{Z}$, there exists a unique $Q \in Q_k$ such that $x \in Q$. That is, there exists a unique $m \in \mathbb{Z}^d$ with $x \in 2^{-k}\big([0,1)^d + m\big)$.

2. Given $Q \in Q_k$ and $j < k$, there exists a unique $\widetilde{Q} \in Q_j$ with $Q \subset \widetilde{Q}$.

3. Each cube in Q_k contains exactly 2^d cubes in Q_{k+1}.

4. Given two dyadic cubes, either one of them is contained in the other, or else the cubes are disjoint.

Definition 11.2. For $f \in L^1_{loc}(\mathbb{R}^d)$, let

$$(E_k f)(x) = \sum_{Q \in Q_k} \Big(\frac{1}{|Q|} \int_Q f(y)\, dy \Big) \mathbb{1}_Q(x).$$

Then $E_k f$ is constant on each cube in Q_k (equalling there the average of f over that cube), and

$$\int_\Omega E_k f\, dx = \int_\Omega f\, dx \tag{11.1}$$

whenever Ω is a finite union of cubes in Q_k.

Define the *dyadic maximal function*

$$\begin{aligned}(M_d f)(x) &= \sup_k |(E_k f)(x)| \\ &= \sup \Big\{ \Big| \frac{1}{|Q|} \int_Q f(y)\, dy \Big| : Q \text{ is a dyadic cube containing } x \Big\}.\end{aligned}$$

Theorem 11.3.
(a) M_d is weak $(1,1)$.
(b) If $f \in L^1_{loc}(\mathbb{R}^d)$ then $\lim_{k\to\infty}(E_k f)(x) = f(x)$ a.e.

Proof. We employ a "stopping time" argument like in probability theory for martingales.

For part (a), let $f \in L^1(\mathbb{R}^d), \lambda > 0$. Since $M_d f \le M_d |f|$, we can assume $f \ge 0$. Let

$$\begin{aligned}\Omega &= \{x \in \mathbb{R}^d : (M_d f)(x) > \lambda\}, \\ \Omega_k &= \{x \in \mathbb{R}^d : (E_k f)(x) > \lambda \text{ and } (E_j f)(x) \le \lambda \text{ for all } j < k\}.\end{aligned}$$

Clearly $\Omega_k \subset \Omega$. And if $x \in \Omega$ then $(E_k f)(x) > \lambda$ for some k; a smallest such k exists, because

$$\begin{aligned}\lim_{j\to-\infty} (E_j f)(x) &\le \lim_{j\to-\infty} \frac{1}{(2^{-j})^d} \int_{\mathbb{R}^d} f(y)\, dy \\ &= 0 \\ &< \lambda.\end{aligned}$$

Choosing the smallest k implies $(E_j f)(x) \le \lambda$ for all $j < k$, and so $x \in \Omega_k$. Hence $\Omega = \cup_k \Omega_k$, so that

$$\begin{aligned}
|\Omega| &= \sum_k |\Omega_k| && \text{by disjointness of the } \Omega_k \\
&\le \frac{1}{\lambda} \sum_k \int_{\Omega_k} E_k f \, dx && \text{since } E_k f > \lambda \text{ on } \Omega_k \\
&= \frac{1}{\lambda} \sum_k \int_{\Omega_k} f \, dx && \text{by (11.1), since } \Omega_k \text{ equals a union of cubes in } Q_k \\
& && \text{(recall } E_k f \text{ is constant on each cube in } Q_k\text{)} \\
&\le \frac{1}{\lambda} \int_{\mathbb{R}^d} f \, dx.
\end{aligned}$$

Therefore M_d is weak $(1,1)$.

Part (b) holds if f is continuous, and hence if $f \in L^1_{loc}(\mathbb{R}^d)$ by Theorem 6.4 (exercise), using that the dyadic maximal operator M_d is weak $(1,1)$. □

Note we did not need a covering lemma, when proving the dyadic maximal function is weak $(1,1)$, because disjointness of the cubes is built into the construction.

Theorem 11.4 (Calderón–Zygmund decomposition at level λ)**.** *Let $f \in L^1(\mathbb{R}^d), \lambda > 0$. Then there exists a "good' function $g \in L^1 \cap L^\infty(\mathbb{R}^d)$ and a "bad" function $b \in L^1(\mathbb{R}^d)$ such that*

i. $f = g + b$

ii. $\|g\|_{L^1(\mathbb{R}^d)} \le \|f\|_{L^1(\mathbb{R}^d)}, \quad \|g\|_{L^\infty(\mathbb{R}^d)} \le 2^d \lambda, \quad \|b\|_{L^1(\mathbb{R}^d)} \le 2\|f\|_{L^1(\mathbb{R}^d)}$

iii. $b = \sum_l b_l$ *where b_l is supported in a dyadic cube $Q(l)$ and the $\{Q(l)\}$ are disjoint; we do* not *assume $Q(l) \in Q_l$, just $Q(l) \in Q_k$ for some k.*

iv. $\int_{Q(l)} b_l(x) \, dx = 0$

v. $\|b_l\|_{L^1(\mathbb{R}^d)} \le 2^{d+1} \lambda |Q(l)|$

vi. $\sum_l |Q(l)| \le \frac{1}{\lambda} \|f\|_{L^1(\mathbb{R}^d)}$

Proof. Apply the proof of Theorem 11.3 to $|f|$, and decompose the disjoint sets Ω_k into dyadic cubes in Q_k. Together, these cubes form the collection $\{Q(l)\}$. Property (vi) is just the weak $(1,1)$ estimate that we proved.

For (i), (iii), (iv), argue as follows. Let

$$b_l(x) = \Big(f(x) - \frac{1}{|Q(l)|}\int_{Q(l)} f(y)\,dy\Big)\mathbb{1}_{Q(l)}(x)$$

so that b_l integrates to 0. Define

$$b(x) = \sum_l b_l(x) = \begin{cases} f(x) - \frac{1}{|Q(l)|}\int_{Q(l)} f(y)\,dy & \text{on } Q(l), \text{ for each } l, \\ 0 & \text{on } \mathbb{R}^d \setminus \cup_l Q(l). \end{cases}$$

Then let

$$\begin{aligned} g &= f - b \\ &= \begin{cases} \frac{1}{|Q(l)|}\int_{Q(l)} f(y)\,dy & \text{on } Q(l), \text{ for each } l, \\ f(x) & \text{on } \mathbb{R}^d \setminus \cup_l Q(l). \end{cases} \end{aligned}$$

For (ii), note $\|g\|_{L^1(\mathbb{R}^d)} \le \|f\|_{L^1(\mathbb{R}^d)}$, since $g = f$ off $\cup_l Q(l)$ and on $Q(l)$ we have

$$\int_{Q(l)} |g(x)|\,dx \le \int_{Q(l)} |f(x)|\,dx.$$

Hence $\|b\|_{L^1(\mathbb{R}^d)} = \|f - g\|_{L^1(\mathbb{R}^d)} \le 2\|f\|_{L^1(\mathbb{R}^d)}$.

Next we show $\|g\|_{L^\infty(\mathbb{R}^d)} \le 2^d\lambda$. Suppose $x \in \mathbb{R}^d \setminus \cup_l Q(l)$. Then $g(x) = f(x)$. Since $x \notin \Omega_k$ for all k we have $(E_k|f|)(x) \le \lambda$ for all k. Hence $|f(x)| \le \lambda$ (for almost every such x) by Theorem 11.3(b), so that $|g(x)| \le \lambda$.

Next suppose $x \in Q(l)$ for some l, so that $x \in \Omega_k$ for some k. Then $(E_{k-1}|f|)(x) \le \lambda$, which means

$$\frac{1}{|Q|}\int_Q |f(y)|\,dy \le \lambda$$

for some cube $Q \in Q_{k-1}$ with $x \in Q(l) \subset Q$. Hence

$$\frac{1}{2^d|Q(l)|}\int_{Q(l)} |f(y)|\,dy \le \lambda \tag{11.2}$$

since $Q(l) \subset Q$ and $\operatorname{side}(Q) = 2\operatorname{side}(Q(l))$. Therefore $|g(x)| \le 2^d\lambda$, by definition of g.

For (v), just note

$$\begin{aligned}\int_{Q(l)} |b_l(x)|\,dx &\le 2\int_{Q(l)} |f(x)|\,dx \qquad \text{by definition of } b_l \\ &\le 2^{d+1}\lambda|Q(l)|\end{aligned}$$

by (11.2). □

Now we adapt the theorem to $\mathbb{T}$. We will restrict to "large" λ values, so that the dyadic intervals have length at most 2π and thus fit into $\mathbb{T}$.

Corollary 11.5 (Calderón–Zygmund decomposition on $\mathbb{T}$)**.** *Let $f \in L^1(\mathbb{T}), \lambda > \|f\|_{L^1(\mathbb{T})}$. Then there exists a "good' function $g \in L^\infty(\mathbb{T})$ and a "bad" function $b \in L^1(\mathbb{T})$ such that*

i. $f = g + b$

ii. $\|g\|_{L^1(\mathbb{T})} \le \|f\|_{L^1(\mathbb{T})}, \quad \|g\|_{L^\infty(\mathbb{T})} \le 2\lambda, \quad \|b\|_{L^1(\mathbb{T})} \le 2\|f\|_{L^1(\mathbb{T})}$

iii. $b = \sum_l b_l$ *where b_l is supported in some interval $I(l)$ of the form $2\pi \cdot 2^{-k}\big([0,1) + m\big)$ where $k \ge 1, 0 \le m \le 2^k - 1$, and where the $\{I(l)\}$ are disjoint.*

iv. $\int_{I(l)} b_l(t)\,dt = 0$

v. $\|b_l\|_{L^1(\mathbb{T})} \le \frac{4}{2\pi}\lambda|I(l)|$

vi. $\sum_l |I(l)| \le \frac{2\pi}{\lambda}\|f\|_{L^1(\mathbb{T})}$

Proof. Let $d = 1$. Apply the Calderón–Zygmund Theorem 11.4 to

$$\widetilde{f}(t) = \begin{cases} f(2\pi t), & 0 \le t < 1, \\ 0, & \text{otherwise,} \end{cases}$$

to get $\widetilde{f} = \widetilde{g} + \widetilde{b}$. Note Ω_k is empty for $k \le 0$, since

$$\begin{aligned}(E_k|\widetilde{f}|)(t) &\le \frac{1}{2^{-k}}\int_0^1 |\widetilde{f}(\tau)|\,d\tau \\ &= 2^k\frac{1}{2\pi}\int_0^{2\pi} |f(\tau)|\,d\tau \\ &\le \|f\|_{L^1(\mathbb{T})} \qquad \text{since } k \le 0 \\ &< \lambda\end{aligned}$$

by assumption on λ.

Further, $\Omega_k \subset [0,1]$ for $k \geq 1$, since $E_k|\widetilde{f}| = 0$ outside $[0,1]$. Thus $I(l) = 2\pi Q(l)$ has the form stated in the Corollary.

The Corollary now follows from Theorem 11.4, with $\widetilde{f} = \widetilde{g} + \widetilde{b}$ yielding $f = g + b$. □

Chapter 12

Hilbert transform on $L^p(\mathbb{T})$

Goals

Prove a weak $(1,1)$ estimate on the Hilbert transform on $\mathbb{T}$
Deduce strong (p,p) estimates by interpolation and duality

Reference

[Duoandikoetxea] Section 3.3

Theorem 12.1 (weak $(1,1)$ on $L^2(\mathbb{T})$)**.** *There exists $A > 0$ such that*

$$|\{t \in \mathbb{T} : |(Hf)(t)| > \lambda\}| \leq \frac{A}{\lambda}\|f\|_{L^1(\mathbb{T})}$$

for all $\lambda > 0$ and $f \in L^2(\mathbb{T})$.

Proof. If $\lambda \leq \|f\|_{L^1(\mathbb{T})}$ then $A = 2\pi$ works. So suppose $\lambda > \|f\|_{L^1(\mathbb{T})}$. Apply the Calderón–Zygmund Corollary 11.5 to get $f = g + b$. Note $g \in L^\infty(\mathbb{T})$ and so $g \in L^2(\mathbb{T})$, hence $Hg \in L^2(\mathbb{T})$ by Chapter 10. And $b = f - g \in L^2(\mathbb{T})$ so that $Hb \in L^2(\mathbb{T})$. Further, $b_l \in L^2(\mathbb{T})$ and $b = \sum_l b_l$ with convergence in $L^2(\mathbb{T})$, using disjointness of the supports of the b_l. Hence $Hb = \sum_l Hb_l$ with convergence in $L^2(\mathbb{T})$.

Since $Hf = Hg + Hb$, we have

$$\begin{aligned}
&|\{t \in \mathbb{T} : |(Hf)(t)| > \lambda\}| \\
&\leq |\{t \in \mathbb{T} : |(Hg)(t)| > \lambda/2\}| + |\{t \in \mathbb{T} : |(Hb)(t)| > \lambda/2\}| \\
&= \gamma + \beta,
\end{aligned}$$

say. First, use the L^2 theory on g:

$$\begin{aligned}
\gamma &\le \int_{\mathbb{T}} \frac{|(Hg)(t)|^2}{(\lambda/2)^2}\,dt \\
&\le \frac{4}{\lambda^2}\int_{\mathbb{T}} |g(t)|^2\,dt && \text{since } \|H\|_{L^2(\mathbb{T})\to L^2(\mathbb{T})} = 1 \text{ by Chapter 10} \\
&\le \frac{8}{\lambda}\int_{\mathbb{T}} |g(t)|\,dt && \text{since } \|g\|_{L^\infty(\mathbb{T})} \le 2\lambda \\
&\le \frac{8\cdot 2\pi}{\lambda}\|f\|_{L^1(\mathbb{T})} && \text{since } \|g\|_{L^1(\mathbb{T})} \le \|f\|_{L^1(\mathbb{T})}.
\end{aligned}$$

Second, use L^1 estimates on b, as follows:

$$\begin{aligned}
\beta &\le \Big|\bigcup_l 2I(l)\Big| + \Big|\{t\in\mathbb{T}\setminus\bigcup_l 2I(l) : |(Hb)(t)| > \lambda/2\}\Big| \\
&\le \frac{4\pi}{\lambda}\|f\|_{L^1(\mathbb{T})} + \int_{\mathbb{T}\setminus\cup_l 2I(l)} \frac{|(Hb)(t)|}{\lambda/2}\,dt \\
&\qquad\qquad \text{by the Calderón–Zygmund Corollary 11.5(vi)} \\
&\le \frac{4\pi}{\lambda}\|f\|_{L^1(\mathbb{T})} + \frac{2}{\lambda}\sum_l \int_{\mathbb{T}\setminus 2I(l)} |(Hb_l)(t)|\,dt
\end{aligned}$$

since $|Hb| \le \sum_l |Hb_l|$ a.e.

To finish the proof, we show

$$\sum_l \int_{\mathbb{T}\setminus 2I(l)} |(Hb_l)(t)|\,dt \le (\text{const.})\|f\|_{L^1(\mathbb{T})}. \tag{12.1}$$

By Proposition 10.3 on the interval $\mathbb{T} \setminus 2I(l)$, we have

$$\begin{aligned}
&\int_{\mathbb{T}\setminus 2I(l)} |Hb_l(t)|\,dt \\
&= \int_{\mathbb{T}\setminus 2I(l)} \Big| \frac{1}{2\pi} \int_{I(l)} b_l(\tau) \cot\big(\frac{1}{2}(t-\tau)\big)\,d\tau \Big|\,dt \\
&\qquad \text{noting } t-\tau \text{ is bounded away from 0, since } \tau \in I(l) \text{ and } t \notin 2I(l), \\
&= \int_{\mathbb{T}\setminus 2I(l)} \Big| \frac{1}{2\pi} \int_{I(l)} b_l(\tau) \big[\cot\big(\frac{1}{2}(t-\tau)\big) - \cot\big(\frac{1}{2}(t-c_l)\big) \big]\,d\tau \Big|\,dt \\
&\qquad \text{where } c_l \text{ is the center of } I(l), \text{ using here that } \int_{I(l)} b_l(\tau)\,d\tau = 0, \\
&= \int_{\mathbb{T}\setminus 2I(l)} \Big| \frac{1}{2\pi} \int_{I(l)} b_l(\tau) \frac{\sin\big(\frac{1}{2}(\tau-c_l)\big)}{\sin\big(\frac{1}{2}(t-\tau)\big)\sin\big(\frac{1}{2}(t-c_l)\big)}\,d\tau \Big|\,dt \\
&\le (\text{const.}) \int_{I(l)} |b_l(\tau)| \int_{\mathbb{R}\setminus 2I(l)} \frac{|I(l)|}{|t-\tau||t-c_l|}\,dt d\tau.
\end{aligned}$$

Note that

$$\begin{aligned}
|t-c_l| &\le |t-\tau| + |\tau - c_l| \\
&\le |t-\tau| + \frac{1}{2}|I(l)| && \text{when } \tau \in I(l) \\
&\le 2|t-\tau| && \text{when } t \in \mathbb{R} \setminus 2I(l).
\end{aligned}$$

Hence

$$\begin{aligned}
\int_{\mathbb{R}\setminus 2I(l)} \frac{|I(l)|}{|t-\tau||t-c_l|}\,dt &\le 2 \int_{\mathbb{R}\setminus 2I(l)} \frac{|I(l)|}{|t-c_l|^2}\,dt \\
&= 4 \int_{2r}^{\infty} \frac{2r}{t^2}\,dt && \text{where } 2r = |I(l)| \\
&= 4.
\end{aligned}$$

Thus

$$\begin{aligned}
\text{the left side of (12.1)} &\le (\text{const.}) \sum_l \int_{I(l)} |b_l(\tau)|\,d\tau \\
&= (\text{const.}) \|b\|_{L^1(\mathbb{T})} \\
&\le (\text{const.}) \|f\|_{L^1(\mathbb{T})}
\end{aligned}$$

by the Calderón–Zygmund Corollary 11.5.

We have proved (12.1), and thus the theorem. □

Corollary 12.2. *The Hilbert transform is strong (p,p) for $1<p<\infty$, with $\widehat{(Hf)}(n) = -i\,\mathrm{sign}(n)\widehat{f}(n)$ for all $f \in L^p(\mathbb{T}), n \in \mathbb{Z}$.*

Proof. H is strong $(2,2)$ and linear, by definition in Chapter 10, and H is weak $(1,1)$ on $L^2(\mathbb{T})$ (and hence on the simple functions on $\mathbb{T}$) by Theorem 12.1. So H is strong (p,p) for $1<p<2$ by Remark C.4 after Marcinkiewicz Interpolation (in Appendix C). That is, $H : L^p(\mathbb{T}) \to L^p(\mathbb{T})$ is bounded and linear for $1<p<2$.

For $2<p<\infty$ we will use duality and anti-selfadjointness $H^* = -H$ on $L^2(\mathbb{T})$ (see Lemma 10.2) to reduce to the case $1<p<2$. Write $\frac{1}{p}+\frac{1}{p'}=1$. If $f \in L^p \cap L^2(\mathbb{T})$ then

$$\begin{aligned}
\|Hf\|_p &= \sup\big\{\big|\frac{1}{2\pi}\int_{\mathbb{T}}(Hf)\overline{g}\,dt\big| : g \in L^{p'}(\mathbb{T}) \text{ with norm } 1\big\}\\
&= \sup\{\big|\frac{1}{2\pi}\int_{\mathbb{T}}(Hf)\overline{g}\,dt\big| : g \in L^{p'}\cap L^2(\mathbb{T}) \text{ with norm } 1\}\\
&\qquad\qquad \text{by density of } L^{p'}\cap L^2 \text{ in } L^{p'}\\
&= \sup\{\big|\frac{1}{2\pi}\int_{\mathbb{T}} f(\overline{Hg})\,dt\big| : g \in L^{p'}\cap L^2(\mathbb{T}) \text{ with norm } 1\}\\
&\qquad\qquad \text{since } H^* = -H \text{ on } L^2(\mathbb{T})\\
&\le \|f\|_{L^p(\mathbb{T})}\sup\{\|Hg\|_{L^{p'}(\mathbb{T})} : g \in L^{p'}\cap L^2(\mathbb{T}) \text{ with norm } 1\} \quad \text{by Holder}\\
&\le (\text{const.})_{p'}\|f\|_{L^p(\mathbb{T})}
\end{aligned}$$

by the strong (p',p') bound proved above, noting $1<p'<2$. Thus H is bounded and linear on the dense subset $L^p \cap L^2(\mathbb{T})$ of $L^p(\mathbb{T})$. Hence H extends to a bounded operator on $L^p(\mathbb{T})$.

Finally, for $f \in L^p(\mathbb{T}), 1<p<\infty$, let $f_m \in L^p \cap L^2(\mathbb{T})$ with $f_m \to f$ in $L^p(\mathbb{T})$. Boundedness of H on L^p implies $Hf_m \to Hf$ in L^p. Hence $f_m \to f$ and $Hf_m \to Hf$ in $L^1(\mathbb{T})$. Thus passing to the limit in $\widehat{(Hf_m)}(n) = -i\,\mathrm{sign}(n)\widehat{f_m}(n)$ yields $\widehat{(Hf)}(n) = -i\,\mathrm{sign}(n)\widehat{f}(n)$, as desired. □

Chapter 13

Applications of interpolation

Goal

Apply Marcinkiewicz and Riesz–Thorin interpolation to the Hilbert transform, maximal operator, Fourier analysis and convolution

The Marcinkiewicz and Riesz–Thorin interpolation theorems are covered in Appendix C. Some important applications are:

Hilbert transform.

$$H : L^p(\mathbb{T}) \to L^p(\mathbb{T}) \text{ is bounded, for } 1 < p < \infty,$$

by the Marcinkiewicz interpolation and duality argument in Corollary 12.2.

Hardy–Littlewood maximal operator. M is weak $(1,1)$ and strong (∞,∞) by Chapter 6, and hence M is strong (p,p) for $1 < p < \infty$ by the Marcinkiewicz Interpolation Theorem C.2. (Note M is sublinear.)

Strong (p,p) was proved directly, already, in Chapter 6.

Fourier analysis. The *Hausdorff–Young* theorem says

$$\widehat{\ }: L^p(\mathbb{T}) \to \ell^{p'}(\mathbb{Z}), \qquad 1 \le p \le 2, \quad \frac{1}{p} + \frac{1}{p'} = 1.$$

It fails for $p > 2$ [Katznelson, Section IV.2.3].

To interpret the theorem, note $L^p(\mathbb{T})$ gets smaller as p increases, and so does $\ell^{p'}(\mathbb{Z})$.

Proof. The analysis operators $\widehat{\ }\colon L^1(\mathbb{T}) \to \ell^\infty(\mathbb{Z})$ and $\widehat{\ }\colon L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ are bounded. Observe

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2} \quad \Longleftrightarrow \quad \frac{\theta}{2} = 1 - \frac{1}{p} \quad \Longleftrightarrow \quad \frac{1}{p'} = \frac{1-\theta}{\infty} + \frac{\theta}{2}.$$

Now apply the Riesz–Thorin Interpolation Theorem C.6.

Convolution. The *Generalized Young's* theorem says

$$\|f*g\|_{L^r(\mathbb{R}^d)} \le \|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^q(\mathbb{R}^d)} \qquad \text{when} \quad \frac{1}{p}+\frac{1}{q} = \frac{1}{r}+1, \quad 1 \le p,q,r \le \infty.$$

Proof. Fix $g \in L^q(\mathbb{R}^d)$ and define $Tf = f * g$. Then T is strong $(1,q)$ since

$$\|f * g\|_{L^q(\mathbb{R}^d)} \le \|f\|_{L^1(\mathbb{R}^d)}\|g\|_{L^q(\mathbb{R}^d)}$$

by Young's Theorem A.3, and T is strong (q',∞) since

$$\|f * g\|_{L^\infty(\mathbb{R}^d)} \le \|f\|_{L^{q'}(\mathbb{R}^d)}\|g\|_{L^q(\mathbb{R}^d)}$$

by Hölder's inequality. In both cases, $\|T\| \le \|g\|_{L^q(\mathbb{R}^d)}$. Observe

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{q'} \quad \Longleftrightarrow \quad \frac{\theta}{q} = 1 - \frac{1}{p} = \frac{1}{q} - \frac{1}{r} \quad \Longleftrightarrow \quad \frac{1}{r} = \frac{1-\theta}{q} + \frac{\theta}{\infty}.$$

Now apply the Riesz–Thorin Interpolation Theorem C.6.

Epilogue: Fourier series in higher dimensions

We have studied Fourier series only on the one dimensional torus $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. The theory extends readily to the higher dimensional torus $\mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d$.

Summability kernels can be obtained by taking products of one dimensional kernels. Thus the higher dimensional Dirichlet kernel is

$$\begin{aligned}\mathcal{D}_n(t) &= D_n(t_1)\cdots D_n(t_d)\\ &= \sum_{j_1,\ldots,j_d=-n}^{n} e^{ijt},\end{aligned}$$

where $j = (j_1, \ldots, j_d), t = (t_1, \ldots, t_d)^\dagger$ and $\dagger$ denotes the transpose operation.

The Dirichlet kernel corresponds to "cubical" partial sums of multiple Fourier series, because

$$\begin{aligned}(\mathcal{D}_n * f)(t) &= \frac{1}{(2\pi)^d}\int_{\mathbb{T}}\cdots\int_{\mathbb{T}} \mathcal{D}_n(t-\tau)f(\tau)\,d\tau_1\cdots d\tau_d\\ &= \sum_{j_1,\ldots,j_d=-n}^{n} \widehat{f}(j)e^{ijt}.\end{aligned}$$

"Spherical" partial sums of the form $\sum_{|j|\le n}\widehat{f}(j)e^{ijt}$ can be badly behaved. For example, they can fail to converge for $f \in L^p(\mathbb{T}^d)$ when $p \neq 2$. See [Grafakos] for this theorem and more on Fourier series in higher dimensions.

Part II

Fourier integrals

Prologue: Fourier series converge to Fourier integrals

Fourier series do not apply to a function $g \in L^1(\mathbb{R})$, since g is not periodic. Instead we take a large *piece* of g and look at its Fourier series: for $\rho > 0$, let

$$f(t) = g(\rho t), \qquad t \in [-\pi, \pi),$$

and extend f to be 2π-periodic. Then

$$\begin{aligned} \widehat{f}(j) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\rho t) e^{-ijt}\, dt \\ &= \frac{1}{2\pi\rho} \int_{-\rho\pi}^{\rho\pi} g(y) e^{-i(j/\rho)y}\, dy \end{aligned}$$

by changing variable. Formally, for $|x| < \rho\pi$ we have

$$\begin{aligned} g(x) = f(\rho^{-1}x) &= \sum_{j=-\infty}^{\infty} \widehat{f}(j) e^{ij(\rho^{-1}x)} \\ &= \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \Big(\int_{-\rho\pi}^{\rho\pi} g(y) e^{-i(j/\rho)y}\, dy \Big) e^{i(j/\rho)x} \cdot \frac{1}{\rho} \\ &\to \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} g(y) e^{-i\xi y}\, dy \Big) e^{i\xi x}\, d\xi \end{aligned}$$

as $\rho \to \infty$, by using Riemann sums on the ξ-integral.

The inner integral ("Fourier transform") is analogous to a Fourier coefficient.

The outer integral ("Fourier inverse") is analogous to a Fourier series.

We aim to develop a Fourier integral theory that is analogous to the theory of Fourier series.

Chapter 14

Fourier transforms: basic properties

Goal

Derive basic properties of Fourier transforms

Reference

[Katznelson] Section VI.1

Notation

$\|f\|_{L^p(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(x)|^p \, dx \right)^{1/p}$
Nesting of L^p-spaces fails: $L^\infty(\mathbb{R}^d) \not\subset L^2(\mathbb{R}^d) \not\subset L^1(\mathbb{R}^d)$ due to behavior at infinity *e.g.* $1/(1+|x|)$ is in $L^2(\mathbb{R})$ but not $L^1(\mathbb{R})$
$C_c(\mathbb{R}^d) = \{$complex-valued, continuous functions with compact support$\}$
$C_0(\mathbb{R}^d) = \{$complex-valued, continuous functions with $f(x) \to 0$ as $|x| \to \infty\}$, Banach space with norm $\|\cdot\|_{L^\infty(\mathbb{R}^d)}$
Translation $f_y(x) = f(x-y)$

Definition 14.1. For $f \in L^1(\mathbb{R}^d)$ and $\xi \in \mathbb{R}^d$, define

$$\begin{aligned} \widehat{f}(\xi) &= \textit{Fourier transform} \text{ of } f \\ &= \int_{\mathbb{R}^d} f(x) e^{-i\xi x} \, dx. \end{aligned} \tag{14.1}$$

Here ξ is a row vector, x is a column vector, and so $\xi x = \xi_1 x_1 + \cdots + \xi_d x_d$ equals the dot product.

Theorem 14.2 (Basic properties)**.** *Let $f, g \in L^1(\mathbb{R}^d), \xi, \omega \in \mathbb{R}^d, c \in \mathbb{C}, y \in \mathbb{R}^d, A \in GL(\mathbb{R}, d)$.*
Linearity $\widehat{(f+g)}(\xi) = \widehat{f}(\xi) + \widehat{g}(\xi)$ and $\widehat{(cf)}(\xi) = c\widehat{f}(\xi)$
Conjugation $\overline{\widehat{f}}(\xi) = \widehat{f}(-\xi)$
$\widehat{\ }$ *takes translation to modulation,* $\widehat{f_y}(\xi) = e^{-i\xi y}\widehat{f}(\xi)$
$\widehat{\ }$ *takes modulation to translation,* $[f(x)e^{i\omega x}]\widehat{\ }(\xi) = \widehat{f}(\xi - \omega)$
$\widehat{\ }$ *takes matrix dilation to its inverse,* $[\,|\det A| f(Ax)]\widehat{\ }(\xi) = \widehat{f}(\xi A^{-1})$
$\widehat{\ }: L^1(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d)$ *is bounded, with* $\|\widehat{f}\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}$
$\widehat{f}$ *is uniformly continuous*
If $f_m \to f$ *in* $L^1(\mathbb{R}^d)$ *then* $\widehat{f_m} \to \widehat{f}$ *in* $L^\infty(\mathbb{R}^d)$.

Proof. Exercise. For continuity, observe

$$|\widehat{f}(\xi+\omega) - \widehat{f}(\xi)| \leq \int_{\mathbb{R}^d} |f(x)||e^{-i\xi x}||e^{-i\omega x} - 1|\, dx$$
$$\to 0$$

as $\omega \to 0$, by dominated convergence. The convergence is independent of ξ, and so $\widehat{f}$ is uniformly continuous. □

Corollary 14.3 (Transform of a radial function)**.** *If $f \in L^1(\mathbb{R}^d)$ is radial then $\widehat{f}$ is radial.*

Recall that f is *radial* if it depends only on the distance to the origin: $f(x) = F(|x|)$ for some function F. Equivalently, f is radial if $f(Ax) = f(x)$ for every x and every orthogonal ("rotation and reflection") matrix A.

Proof. Suppose A is orthogonal. Then $f(Ax) = f(x)$ (since f is radial) and so

$$\widehat{f}(\xi A^{-1}) = [\,|\det A| f(Ax)]\widehat{\ }(\xi) = \widehat{f}(\xi),$$

using Theorem 14.2 and that $|\det A| = 1$. □

Lemma 14.4 (Transform of a product)**.** *If $f_1, \ldots, f_d \in L^1(\mathbb{R})$ then $f(x) = \prod_{j=1}^d f_j(x_j)$ has transform $\widehat{f}(\xi) = \prod_{j=1}^d \widehat{f_j}(\xi_j)$.*

Proof. Use Fubini and the homomorphism property of the exponential: $e^{-i\xi x} = \prod_{j=1}^d e^{-i\xi_j x_j}$. □

Lemma 14.5 (Difference formula)**.** *For $\xi \neq 0$,*

$$\widehat{f}(\xi) = \frac{1}{2}\int_{\mathbb{R}^d} [f(x) - f(x - \pi\xi^\dagger/|\xi|^2)]\, e^{-i\xi x}\, dx,$$

where $\xi^\dagger$ is the column vector transpose of ξ.

Proof. Like Lemma 1.3. □

Lemma 14.6 (Continuity of translation)**.** *Fix $f \in L^p(\mathbb{R}^d), 1 \leq p < \infty$. The map*

$$\begin{aligned} \phi : \mathbb{R}^d &\to L^p(\mathbb{R}^d) \\ y &\mapsto f_y \end{aligned}$$

is continuous.

Proof. Like Lemma 1.4 except using $C_c(\mathbb{R}^d)$, which is dense in $L^p(\mathbb{R}^d)$. □

Corollary 14.7 (Riemann–Lebesgue lemma)**.** *$\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$. Thus $\widehat{f} \in C_0(\mathbb{R}^d)$.*

Proof. Lemma 14.5 implies

$$|\widehat{f}(\xi)| \leq \frac{1}{2}\|f - f_{\pi\xi^\dagger/|\xi|^2}\|_{L^1(\mathbb{R}^d)},$$

which tends to zero as $|\xi| \to \infty$ by the L^1-continuity of translation in Lemma 14.6, since $\xi^\dagger/|\xi|^2$ has magnitude $1/|\xi| \to 0$. □

Example 14.8. We compute the Fourier transforms in Table 14.1.

1. $\int_{\mathbb{R}} \mathbb{1}_{[-1,1]}(x) e^{-i\xi x}\, dx = \int_{-1}^{1} e^{-i\xi x}\, dx = 2\sin(\xi)/\xi$

2. $\int_{\mathbb{R}} (1 - |x|)\mathbb{1}_{[-1,1]}(x) e^{-i\xi x}\, dx = 2\int_0^1 (1-x)\cos(\xi x)\, dx = 2\xi^{-2}(1 - \cos\xi)$, and $1 - \cos\xi = 2\sin^2(\xi/2)$

4. Next we compute for the fourth example, the Gaussian $e^{-|x|^2/2}$, so that we can use it later for the third example $e^{-|x|}$.

For $d = 1$, let $g(\xi) = \int_{\mathbb{R}} e^{-x^2/2} e^{-i\xi x}\, dx$ be the transform we want. Note $g(0) = \sqrt{2\pi}$. Differentiating,

$$g'(\xi) = \int_{\mathbb{R}} e^{-x^2/2}(-ix) e^{-i\xi x}\, dx,$$

dimension	$f(x)$	$\widehat{f}(\xi)$
1	$\mathbb{1}_{[-1,1]}(x)$	$2\frac{\sin\xi}{\xi} = 2\,\mathrm{sinc}\,\xi$
1	$(1-\lvert x\rvert)\mathbb{1}_{[-1,1]}(x)$	$\left(\frac{\sin(\xi/2)}{\xi/2}\right)^2 = \mathrm{sinc}^2(\xi/2)$
d	$e^{-\lvert x\rvert}$	$\frac{(2\pi)^d c_d}{(1+\lvert\xi\rvert^2)^{(d+1)/2}}$
d	$e^{-\lvert x\rvert^2/2}$	$(2\pi)^{d/2}e^{-\lvert\xi\rvert^2/2}$

Table 14.1: Fourier transforms from Example 14.8. In the third example, $c_d = \Gamma\left(\frac{d+1}{2}\right)\Big/\,\pi^{(d+1)/2}$, so that $c_1 = 1/\pi$. The fourth example says the Fourier transform of a Gaussian is a Gaussian.

with the differentiation through the integral justified by using difference quotients and dominated convergence (Exercise). Hence

$$\begin{aligned} g'(\xi) &= i\int_{\mathbb{R}} \left(e^{-x^2/2}\right)' e^{-i\xi x}\,dx \\ &= -i\int_{\mathbb{R}} e^{-x^2/2}\left(e^{-i\xi x}\right)'\,dx \qquad \text{by parts} \\ &= -\xi\int_{\mathbb{R}} e^{-x^2/2}e^{-i\xi x}\,dx \\ &= -\xi g(\xi). \end{aligned}$$

Solving the differential equation yields $g(\xi) = \sqrt{2\pi}e^{-\xi^2/2}$.

For $d > 1$, note the product structure $e^{-|x|^2/2} = \prod_{j=1}^d e^{-x_j^2/2}$ and apply Lemma 14.4.

3. For $d = 1$, $\int_{\mathbb{R}} e^{-|x|}e^{-i\xi x}\,dx = \int_0^\infty e^{-(1+i\xi)x}\,dx + \int_{-\infty}^0 e^{(1-i\xi)x}\,dx = 1/(1+i\xi) + 1/(1-i\xi)$, which simplifies to the desired result.

To handle $d > 1$, we need a calculus lemma that expresses a decaying exponential as a superposition of Gaussians.

Lemma 14.9. *For $b > 0$,*

$$e^{-b} = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{e^{-a/2}}{\sqrt{a}} e^{-b^2/2a}\, da.$$

Proof.

$$\begin{aligned}
& e^b \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{e^{-a/2}}{\sqrt{a}} e^{-b^2/2a}\, da \\
&= \frac{2\sqrt{b}}{\sqrt{2\pi}} \int_0^\infty e^{-b(c-1/c)^2/2}\, dc && \text{by letting } a = bc^2 \\
&= \frac{2\sqrt{b}}{\sqrt{2\pi}} \int_0^\infty e^{-b(c-1/c)^2/2} c^{-2}\, dc && \text{by } c \mapsto 1/c \\
&= \frac{\sqrt{b}}{\sqrt{2\pi}} \int_0^\infty e^{-b(c-1/c)^2/2} (1 + c^{-2})\, dc && \text{by averaging the last two formulas} \\
&= \frac{\sqrt{b}}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-bu^2/2}\, du && \text{where } u = c - 1/c \\
&= 1.
\end{aligned}$$

□

Now we compute the Fourier transform of $e^{-|x|}$ as

$$\begin{aligned}
& \int_{\mathbb{R}^d} e^{-|x|} e^{-i\xi x}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{e^{-a/2}}{\sqrt{a}} \int_{\mathbb{R}^d} e^{-|x|^2/2} e^{-i(\xi\sqrt{a})x}\, dx\, a^{d/2}\, da && \text{by Lemma 14.9 and } x \mapsto \sqrt{a}x \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty a^{(d-1)/2} e^{-a/2} (2\pi)^{d/2} e^{-|\xi\sqrt{a}|^2/2}\, da && \text{by the Gaussian in Table 14.1} \\
&= (2\pi)^{(d-1)/2} \big((1 + |\xi|^2)/2\big)^{-(d+1)/2} \int_0^\infty u^{(d-1)/2} e^{-u}\, du
\end{aligned}$$

where $u = a(1+|\xi|^2)/2$. The last integral is $\Gamma((d+1)/2)$, so that the transform equals $(2\pi)^d c_d (1 + |\xi|^2)^{-(d+1)/2}$ as claimed in the Table.

Smoothness and decay

Theorem 14.10 (Differentiation and Fourier transforms)**.**

(a) If $f \in C_c^1(\mathbb{R}^d)$ (or more generally, $f \in W^{1,1}(\mathbb{R}^d)$) then

$$\widehat{(\partial_j f)}(\xi) = i\xi_j \widehat{f}(\xi),$$

where $\partial_j = \partial/\partial x_j$ for $j = 1, \ldots, d$. Thus:

$\widehat{\ }$ *takes differentiation to multiplication by $i\xi_j$.*

(b) If $(1 + |x|)f(x) \in L^1(\mathbb{R}^d)$ then $\widehat{f}$ is continuously differentiable, with

$$\widehat{(-ix_j f)}(\xi) = (\partial_j \widehat{f})(\xi),$$

where $\partial_j = \partial/\partial \xi_j$ for $j = 1, \ldots, d$. Thus:

$\widehat{\ }$ *takes multiplication by $-ix_j$ to differentiation.*

Proof. For (a)

$$\begin{aligned} \int_{\mathbb{R}^d} (\partial_j f)(x) e^{-i\xi x}\, dx &= \int_{\mathbb{R}^d} f(x)(i\xi_j) e^{-i\xi x}\, dx \qquad \text{by parts} \\ &= i\xi_j \widehat{f}(\xi). \end{aligned}$$

For (b) we compute a difference quotient, with $\delta \in \mathbb{R}$ and e_j = unit vector in the j-th direction:

$$\begin{aligned} \frac{\widehat{f}(\xi + \delta e_j) - \widehat{f}(\xi)}{\delta} &= \int_{\mathbb{R}^d} f(x) e^{-i\xi x} \frac{e^{-i\delta x_j} - 1}{\delta}\, dx \\ &\to \int_{\mathbb{R}^d} f(x) e^{-i\xi x}(-ix_j)\, dx = \widehat{(-ix_j f)}(\xi) \end{aligned}$$

as $\delta \to 0$, by dominated convergence with dominating function $f(x)|x| \in L^1(\mathbb{R}^d)$. Hence $\widehat{f}(\xi)$ has partial derivative $\widehat{(-ix_j f)}(\xi)$, which is continuous by Theorem 14.2. □

Theorem 14.11 (Smoothness of f and decay of $\widehat{f}$)**.**

(a) If $f \in L^1(\mathbb{R}^d)$ then $\widehat{f}(\xi) = o(1)$ as $|\xi| \to \infty$, and

$$|\widehat{f}(\xi)| \le \|f\|_{L^1(\mathbb{R}^d)} = O(1).$$

(b) If $f \in C_c^1(\mathbb{R}^d)$ then $\widehat{f}(\xi) = o(1/|\xi|)$ as $|\xi| \to \infty$, and

$$|\widehat{f}(\xi)| \le \frac{d \max_j \|\partial_j f\|_{L^1(\mathbb{R}^d)}}{|\xi|} = O(1/|\xi|).$$

Proof. (a) Use Riemann–Lebesgue (Corollary 14.7) and Theorem 14.2.

(b) For each ξ there exists j such that $|\xi_j| \geq |\xi|/d$ (since $|\xi_1| + \cdots + |\xi_d| \geq |\xi|$). Then

$$\begin{aligned}
|\widehat{f}(\xi)| = \left| \frac{\widehat{(\partial_j f)}(\xi)}{i\xi_j} \right| &\leq \frac{|\widehat{(\partial_j f)}(\xi)|}{|\xi|/d} \\
&\leq \frac{d \max_j |\widehat{(\partial_j f)}(\xi)|}{|\xi|} \\
&= o(1/|\xi|) && \text{by Riemann–Lebesgue} \\
&\leq \frac{d \max_j \|\partial_j f\|_{L^1(\mathbb{R}^d)}}{|\xi|} && \text{by Theorem 14.2} \\
&= O(1/|\xi|).
\end{aligned}$$

Or one could argue more directly using the gradient vector:

$$\begin{aligned}
|\widehat{f}(\xi)| = \frac{|\widehat{(\nabla f)}(\xi)|}{|i\xi|} = o(1/|\xi|) && \text{by Riemann–Lebesgue} \\
\leq \frac{\|\nabla f\|_{L^1(\mathbb{R}^d)}}{|\xi|} && \text{by Theorem 14.2} \\
= O(1/|\xi|).
\end{aligned}$$

□

Convolution

Definition 14.12. Given $f, g \in L^1(\mathbb{R}^d)$, define their *convolution*

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y) g(y)\, dy, \qquad x \in \mathbb{R}^d.$$

Theorem 14.13 (Convolution and Fourier transforms)**.** *If $f, g \in L^1(\mathbb{R}^d)$ then $f * g \in L^1(\mathbb{R}^d)$ with*

$$\|f * g\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}$$

and

$$\widehat{(f * g)}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi), \qquad \xi \in \mathbb{R}^d.$$

Thus the Fourier transform takes convolution to multiplication.

Proof. Like Theorem 1.11. □

Example 14.14. Let $f = \mathbb{1}_{[-1/2,1/2]}$, so that $(f * f)(x) = (1 - |x|)\mathbb{1}_{[-1,1]}(x)$ by direct calculation. We find $\widehat{f}(\xi) = \operatorname{sinc}(\xi/2)$ like example 1 of Table 14.1, and $\widehat{(f * f)}(\xi) = \operatorname{sinc}^2(\xi/2)$ by example 2 of Table 14.1.

Hence $\widehat{(f * f)} = (\widehat{f})^2$, as Theorem 14.13 predicts.

As this example illustrates, **convolution is a smoothing operation**, and hence improves the decay of the transform: $\operatorname{sinc}(\xi/2)$ decays like $1/\xi$ while $\operatorname{sinc}^2(\xi/2)$ decays like $1/\xi^2$.

Convolution facts (similar to Chapter 2)

1. Convolution is commutative: $f * g = g * f$. It is also associative, and linear with respect to f and g.

2. If $f \in L^p(\mathbb{R}^d), 1 \le p \le \infty$, and $g \in L^1(\mathbb{R}^d)$, then $f * g \in L^p(\mathbb{R}^d)$ with

$$\|f * g\|_{L^p(\mathbb{R}^d)} \le \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}.$$

Further, if $f \in C_0(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$ then $f * g \in C_0(\mathbb{R}^d)$.
Proof. For the first claim, use Young's Theorem A.3. For the second, if $f \in C_0(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$ then $f * g$ is continuous because $(f * g)(x + z) \to (f * g)(x)$ as $z \to 0$ by uniform continuity of f (exercise). And $(f * g)(x) \to 0$ as $|x| \to \infty$ by dominated convergence, since $f(x - y) \to 0$ as $|x| \to \infty$.

3. Convolution is continuous on $L^p(\mathbb{R}^d)$: if $f_m \to f$ in $L^p(\mathbb{R}^d), 1 \le p \le \infty$, and $g \in L^1(\mathbb{R}^d)$, then $f_m * g \to f * g$ in $L^p(\mathbb{R}^d)$.
Proof. Use linearity and Fact 2.

4. If $f \in L^1(\mathbb{R}^d)$ and $P(x) = \int_{\mathbb{R}^d} Q(\xi) e^{i\xi x}\, d\xi$ for some $Q \in L^1(\mathbb{R}^d)$, then

$$(P * f)(x) = \int_{\mathbb{R}^d} Q(\xi) \widehat{f}(\xi) e^{i\xi x}\, d\xi. \tag{14.2}$$

Proof.

$$\begin{aligned}(P * f)(x) &= \int_{\mathbb{R}^d} Q(\xi) \int_{\mathbb{R}^d} e^{i\xi(x-y)} f(y)\, dy d\xi \qquad \text{by Fubini} \\ &= \int_{\mathbb{R}^d} Q(\xi) e^{i\xi x} \widehat{f}(\xi)\, d\xi.\end{aligned}$$

Chapter 15

Fourier integrals: summability in norm

Goal

Develop summability kernels in $L^p(\mathbb{R}^d)$

Reference

[Katznelson] Section VI.1

Definition 15.1. A **summability kernel** on $\mathbb{R}^d$ is a family $\{k_\omega\}$ of integrable functions such that

$$\int_{\mathbb{R}^d} k_\omega(x)\,dx = 1 \qquad (\text{NORMALIZATION}) \qquad \text{(SR1)}$$

$$\sup_\omega \int_{\mathbb{R}^d} |k_\omega(x)|\,dx < \infty \qquad (L^1 \text{ BOUND}) \qquad \text{(SR2)}$$

$$\lim_{\omega\to\infty} \int_{\{x:|x|>\delta\}} |k_\omega(x)|\,dx = 0 \qquad (L^1 \text{ CONCENTRATION}) \qquad \text{(SR3)}$$

for each $\delta > 0$.

Some kernels further satisfy

$$\lim_{\omega\to\infty} \sup_{|x|>\delta} |k_\omega(x)| = 0 \qquad (L^\infty \text{ CONCENTRATION}) \qquad \text{(SR4)}$$

for each $\delta > 0$.

(*Notation.* Here $k_\omega(x)$ does *not* mean the translation $k(x-\omega)$.)

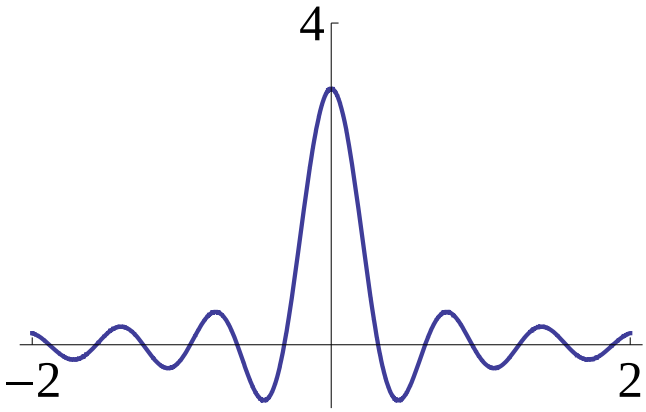

Figure 15.1: Dirichlet kernel with $\omega = 10$

Example 15.2. Suppose $k \in L^1(\mathbb{R}^d)$ is continuous with $\int_{\mathbb{R}^d} k(x)\,dx = 1$. Put

$$k_\omega(x) = \omega^d k(\omega x)$$

for $\omega > 0$. Then $\{k_\omega\}$ is a summability kernel.
Proof. Show (SR1) and (SR2) by changing variable with $y = \omega x, dy = \omega^d dx$. For (SR3),

$$\begin{aligned}\int_{\{x:|x|>\delta\}} |k_\omega(x)|\,dx &= \int_{\{y:|y|>\omega\delta\}} |k(y)|\,dy \\ &\to 0\end{aligned}$$

as $\omega \to \infty$, by dominated convergence.

Example 15.3. For $d = 1$, let

$$D(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathbb{1}_{[-1,1]}(\xi) e^{i\xi x}\,d\xi \tag{15.1}$$

$$= \frac{\sin x}{\pi x} = \frac{1}{\pi}\operatorname{sinc} x. \tag{15.2}$$

The **Dirichlet kernel** is

$$D_\omega(x) = \omega D(\omega x) = \frac{1}{2\pi}\int_{-\omega}^{\omega} e^{i\xi x}\,d\xi \tag{15.3}$$

$$= \frac{\sin(\omega x)}{\pi x}. \tag{15.4}$$

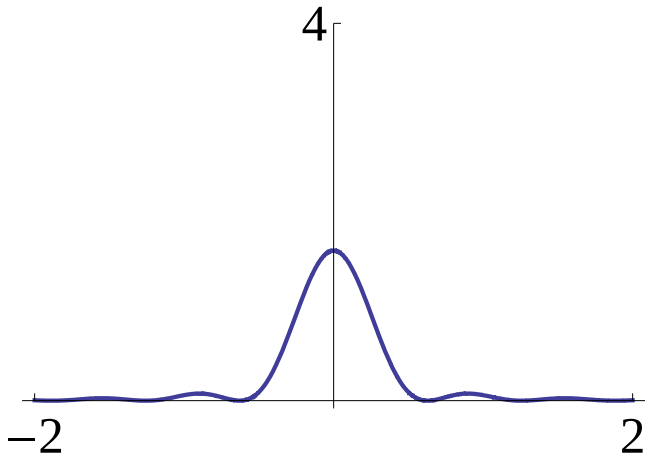

Figure 15.2: Fejér kernel with $\omega = 10$

See Figure 15.1. D is not integrable since $|D(x)| \sim |x|^{-1}$ at infinity.
$\therefore \{D_\omega\}$ is not a summability kernel.

In higher dimensions the Dirichlet function is $\prod_{j=1}^d D(x_j)$, with associated kernel $D_\omega(x) = \prod_{j=1}^d D_\omega(x_j)$.

Example 15.4. For $d = 1$, let

$$F(x) = \frac{1}{2\pi} \int_{\mathbb{R}} (1 - |\xi|) \mathbb{1}_{[-1,1]}(\xi) e^{i\xi x} \, d\xi \tag{15.5}$$

$$= \frac{1}{2\pi} \left(\frac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x} \right)^2 \qquad \text{by Table 14.1.} \tag{15.6}$$

The **Fejér kernel** is

$$F_\omega(x) = \omega F(\omega x) = \frac{1}{2\pi} \int_{-\omega}^{\omega} (1 - |\xi|/\omega) e^{i\xi x} \, d\xi \tag{15.7}$$

$$= \frac{\omega}{2\pi} \left(\frac{\sin\left(\frac{1}{2}\omega x\right)}{\frac{1}{2}\omega x} \right)^2. \tag{15.8}$$

See Figure 15.2. F is integrable since $F(x) \sim x^{-2}$ at infinity. And

$$\begin{aligned}\int_{\mathbb{R}} F(x)\,dx &= \frac{2}{\pi}\lim_{\rho\to\infty}\int_{-\rho}^{\rho}\frac{\sin^2(x/2)}{x^2}\,dx \\ &= \frac{2}{\pi}\lim_{\rho\to\infty}\int_{-\rho}^{\rho}\frac{2\sin(x/2)\cos(x/2)\cdot(1/2)}{x}\,dx \qquad \text{by parts} \\ &= \frac{1}{\pi}\lim_{\rho\to\infty}\int_{-\rho}^{\rho}\frac{\sin x}{x}\,dx \\ &= 1.\end{aligned}$$

$\therefore \{F_\omega\}$ is a summability kernel.

In higher dimensions the Fejér function is $\prod_{j=1}^d F(x_j)$, with associated kernel $F_\omega(x) = \prod_{j=1}^d F_\omega(x_j)$.

The Fejér kernel is an arithmetic mean of Dirichlet kernels; for example, $F(x) = \int_0^1 D_\omega(x)\,d\omega$ in 1 dimension, by integrating (15.3).

Example 15.5.

$$P(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|}e^{i\xi x}\,d\xi \tag{15.9}$$

$$= \frac{c_d}{(1+|x|^2)^{(d+1)/2}} \qquad \text{by Table 14.1.} \tag{15.10}$$

The **Poisson kernel** is

$$P_\omega(x) = \omega^d P(\omega x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|/\omega}e^{i\xi x}\,d\xi \tag{15.11}$$

$$= c_d\frac{\omega^{-1}}{\left(|x|^2+\omega^{-2}\right)^{(d+1)/2}}. \tag{15.12}$$

See Figure 15.3. P is integrable since $P(x) \sim |x|^{-(d+1)}$ at infinity. And $\int_{\mathbb{R}^d} P(x)\,dx = \widehat{P}(0) = 1$ because $\widehat{P}(\xi) = e^{-|\xi|}$ by Example 16.3 below; alternatively, one can integrate (15.10) directly (see [Stein and Weiss, p. 9] for $d > 1$).
$\therefore \{P_\omega\}$ is a summability kernel.

Example 15.6.

$$G(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|^2/2}e^{i\xi x}\,d\xi \tag{15.13}$$

$$= (2\pi)^{-d/2}e^{-|x|^2/2} \qquad \text{by Table 14.1.} \tag{15.14}$$

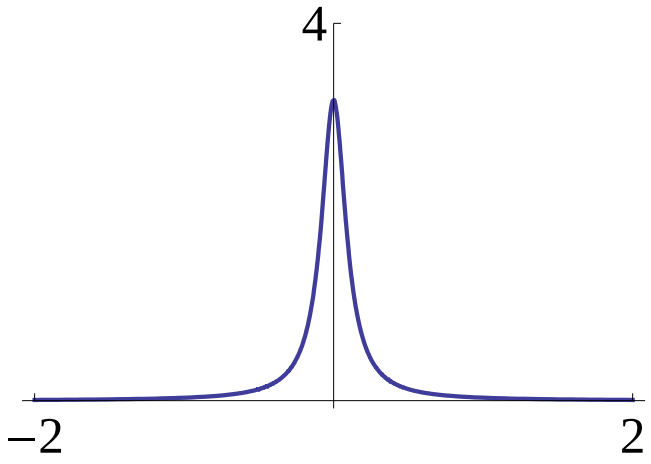

Figure 15.3: Poisson kernel with $\omega = 10$

The **Gauss kernel** is

$$G_\omega(x) = \omega^d G(\omega x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-|\xi/\omega|^2/2} e^{i\xi x}\, d\xi \tag{15.15}$$

$$= \frac{\omega^d}{(2\pi)^{d/2}} e^{-|\omega x|^2/2}. \tag{15.16}$$

See Figure 15.4. G is clearly integrable, and $\int_{\mathbb{R}^d} G(x)\, dx = 1$ from (15.14).
$\therefore \{G_\omega\}$ is a summability kernel.

Connection to Fourier integrals

For $f \in L^1(\mathbb{R}^d)$:

$$(D_\omega * f)(x) = \frac{1}{(2\pi)^d} \int_{[-\omega,\omega]^d} \widehat{f}(\xi) e^{i\xi x}\, d\xi \tag{15.17}$$

$$(F_\omega * f)(x) = \frac{1}{(2\pi)^d} \int_{[-\omega,\omega]^d} \Big(\prod_{j=1}^{d} (1 - |\xi_j|/\omega) \Big)\, \widehat{f}(\xi) e^{i\xi x}\, d\xi \tag{15.18}$$

$$(P_\omega * f)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-|\xi|/\omega} \widehat{f}(\xi) e^{i\xi x}\, d\xi \tag{15.19}$$

$$(G_\omega * f)(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-|\xi/\omega|^2/2} \widehat{f}(\xi) e^{i\xi x}\, d\xi \tag{15.20}$$

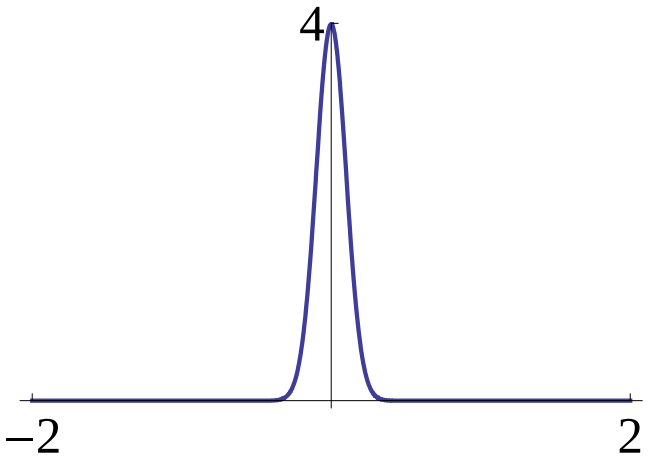

Figure 15.4: Gauss kernel with $\omega = 10$

Proof. Use Convolution Fact (14.2) and definitions (15.1), (15.5), (15.9), (15.13), respectively.

Caution. The left sides of the above formulas make sense for $f \in L^p(\mathbb{R}^d)$, but the right side does not: so far we have defined the Fourier transform only for $f \in L^1(\mathbb{R}^d)$.

Summability in norm

Theorem 15.7 (Summability in $L^p(\mathbb{R}^d)$ and $C_0(\mathbb{R}^d)$)**.** *Assume* $\{k_\omega\}$ *is a summability kernel.*

(a) If $f \in L^p(\mathbb{R}^d), 1 \le p < \infty$, *then* $k_\omega * f \to f$ *in* $L^p(\mathbb{R}^d)$ *as* $\omega \to \infty$.

(b) If $f \in C_0(\mathbb{R}^d)$ *then* $k_\omega * f \to f$ *in* $C_0(\mathbb{R}^d)$ *as* $\omega \to \infty$.

Recall that $C_0(\mathbb{R}^d)$ uses the L^∞ norm.

Proof. Argue as for Theorem 2.6. Use that if $f \in C_0(\mathbb{R}^d)$ then f is uniformly continuous. □

Consequences

• Fejér summability for $f \in L^1(\mathbb{R}^d)$:

$$\frac{1}{(2\pi)^d} \int_{[-\omega,\omega]^d} \Big(\prod_{j=1}^{d} (1 - |\xi_j|/\omega) \Big) \widehat{f}(\xi) e^{i\xi x} \, d\xi \to f(x) \qquad \text{in } L^1(\mathbb{R}^d). \quad (15.21)$$

Similarly for Poisson and Gauss summability.

Proof. Use Theorem 15.7 and formulas (15.18)–(15.20).

• Uniqueness theorem:

$$\text{if } f, g \in L^1(\mathbb{R}^d) \text{ with } \widehat{f} = \widehat{g} \text{ then } f = g. \quad (15.22)$$

That is, the Fourier transform $\widehat{\ }: L^1(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d)$ is injective.

Proof. Use Fejér summability (15.21) on f and g.

Connection to PDEs

Fix $f \in L^1(\mathbb{R}^d)$.

1. The Poisson kernel solves Laplace's equation in a half-space:

$$\begin{aligned} v(x, x_{d+1}) &= (P_{1/x_{d+1}} * f)(x) \\ &= c_d \int_{\mathbb{R}^d} \frac{x_{d+1}}{\big(|x-y|^2 + x_{d+1}^2\big)^{(d+1)/2}} f(y) \, dy \end{aligned}$$

solves

$$(\partial_1^2 + \cdots + \partial_d^2 + \partial_{d+1}^2) v = 0$$

on $\mathbb{R}^d \times (0, \infty)$, with boundary value $v(x, 0) = f(x)$ in the sense of Theorem 15.7.

That is, v is the harmonic extension of f from $\mathbb{R}^d$ to the halfspace $\mathbb{R}^d \times (0, \infty)$.

Proof. Take $\omega = 1/x_{d+1}$ in (15.19) and differentiate through the integral, using

$$\begin{aligned} \sum_{j=1}^{d+1} \frac{\partial^2}{\partial x_j^2} (e^{-|\xi| x_{d+1}} e^{i\xi x}) &= \big((i\xi_1)^2 + \cdots + (i\xi_d)^2 + (-|\xi|)^2 \big) e^{-|\xi| x_{d+1}} e^{i\xi x} \\ &= 0. \end{aligned}$$

For the boundary value, note $\omega = 1/x_{d+1} \to \infty$ as $x_{d+1} \to 0$.

2. The Gauss kernel solves the diffusion (heat) equation:

$$\begin{aligned} w(t,x) &= (G_{1/\sqrt{2t}} * f)(x) \\ &= \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-|x-y|^2/4t} f(y)\, dy \end{aligned}$$

solves

$$w_t = \Delta w$$

for $(t,x) \in (0,\infty) \times \mathbb{R}^d$, with initial value $w(0,x) = f(x)$ in the sense of Theorem 15.7. (Here $\Delta = \partial_1^2 + \cdots + \partial_d^2$.)

Proof. Take $\omega = 1/\sqrt{2t}$ in (15.20) and differentiate through the integral, using

$$\begin{aligned} \left(\frac{\partial}{\partial t} - \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} \right) (e^{-|\xi|^2 t} e^{i\xi x}) &= \big(-|\xi|^2 - (i\xi_1)^2 - \cdots - (i\xi_d)^2 \big) e^{-|\xi|^2 t} e^{i\xi x} \\ &= 0. \end{aligned}$$

For the boundary value, note $\omega = 1/\sqrt{2t} \to \infty$ as $t \to 0$.

Chapter 16

Fourier transforms in $L^1(\mathbb{R}^d)$, and Fourier inversion

Goal

Fourier inversion when $\widehat{f}$ is integrable

Reference

[Katznelson, Section VI.1]

Definition 16.1. Define

$$\begin{aligned}\check{g}(x) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} g(\xi)e^{i\xi x}\,d\xi \\ &= \frac{1}{(2\pi)^d}\widehat{g}(-x).\end{aligned}$$

We call $\check{}$ the *inverse Fourier transform*, in view of the next theorem.

Theorem 16.2. *(Fourier inversion)*

(a) If $f, \widehat{f} \in L^1(\mathbb{R}^d)$ then f is continuous and

$$f(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \widehat{f}(\xi)e^{i\xi x}\,d\xi, \qquad x \in \mathbb{R}^d.$$

(b) If $g, \check{g} \in L^1(\mathbb{R}^d)$ then g is continuous and

$$g(\xi) = \int_{\mathbb{R}^d} \check{g}(x)e^{-i\xi x}\,dx, \qquad \xi \in \mathbb{R}^d.$$

dimension	$f(x)$	$\widehat{f}(\xi)$
d	$F(x) = \frac{1}{(2\pi)^d} \prod_{j=1}^d \left(\frac{\sin(x_j/2)}{x_j/2}\right)^2$	$\widehat{F}(\xi) = \mathbb{1}_{[-1,1]^d}(\xi) \prod_{j=1}^d (1 - \lvert\xi_j\rvert)$
d	$P(x) = \frac{c_d}{(1+\lvert x\rvert^2)^{(d+1)/2}}$	$\widehat{P}(\xi) = e^{-\lvert\xi\rvert}$
d	$G(x) = (2\pi)^{-d/2} e^{-\lvert x\rvert^2/2}$	$\widehat{G}(\xi) = e^{-\lvert\xi\rvert^2/2}$

Table 16.1: Fourier transforms of the Fejér, Poisson and Gauss functions, from Example 16.3.

The theorem says $\boxed{(\hat{f})^{\check{}} = f}$ and $\boxed{(\check{g})^{\hat{}} = g}$.

Proof. (a) The L^1 convergence in Fejér summability (15.21) implies pointwise convergence a.e. for some subsequence of ω-values:

$$\begin{aligned} f(x) &= \lim_{\omega\to\infty} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathbb{1}_{[-\omega,\omega]^d}(\xi) \Big(\prod_{j=1}^d (1 - |\xi_j|/\omega) \Big) \widehat{f}(\xi) e^{i\xi x}\, d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{i\xi x}\, d\xi \end{aligned}$$

by dominated convergence, using that $\widehat{f} \in L^1(\mathbb{R}^d)$.

(b) Apply part (a) to g, change $\xi \mapsto -\xi$, and then swap x and ξ. □

Example 16.3. The Fourier transforms of the Fejér, Poisson and Gauss functions can be computed by Fourier Inversion Theorem 16.2(b), because definitions (15.5), (15.9) and (15.13) express those kernels as inverse Fourier transforms. For example, if we choose $g(\xi) = e^{-|\xi|^2/2}$ then definition (15.13) says $G(x) = \check{g}(x)$, so that $\widehat{G} = g$ by Theorem 16.2(b).

Table 16.1 displays the results.

Chapter 17

Fourier transforms in $L^2(\mathbb{R}^d)$

Goal

Extend the Fourier transform to an isometric bijection of $L^2(\mathbb{R}^d)$ to itself

Reference

[Katznelson] Section VI.3

Notation

Inner product on $L^2(\mathbb{R}^d)$ is $\langle f, g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx$.

Theorem 17.1 (Fourier transform on $L^2(\mathbb{R}^d)$)**.** *The Fourier transform* $\widehat{\ }: L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ *is a bijective isometry (up to a constant factor) with*

$$\begin{aligned} \|f\|_{L^2(\mathbb{R}^d)} &= (2\pi)^{-d/2}\|\widehat{f}\|_{L^2(\mathbb{R}^d)} && \text{(Plancherel)} \\ \langle f, g\rangle &= (2\pi)^{-d}\langle \widehat{f}, \widehat{g}\rangle && \text{(Parseval)} \\ (\hat{f})^{\check{}} = f, &\quad (\check{g})^{\hat{}} = g && \text{(Inversion)} \end{aligned}$$

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

The proof will show $\widehat{\ }: L^1 \cap L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is bounded with respect to the L^2 norm. Then by density of $L^1 \cap L^2$ in L^2, we conclude the Fourier transform extends to a bounded operator from L^2 to itself.

Proof. For $f \in L^1 \cap L^2(\mathbb{R}^d)$,

$$\begin{aligned}
\|f\|^2_{L^2(\mathbb{R}^d)} &= \lim_{\omega\to\infty} \int_{\mathbb{R}^d} f(x)\overline{(G_\omega * f)(x)}\,dx \\
&\qquad\qquad \text{since } G_\omega * f \to f \text{ in } L^2(\mathbb{R}^d) \text{ by Theorem 15.7} \\
&= \lim_{\omega\to\infty} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x)e^{-i\xi x}\overline{\widehat{f}(\xi)}e^{-|\xi/\omega|^2/2}\,d\xi dx \qquad \text{by (15.20)} \\
&= \lim_{\omega\to\infty} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 e^{-|\xi/\omega|^2/2}\,d\xi \quad \text{by Fubini, using } \widehat{f} \in L^\infty(\mathbb{R}^d), \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2\,d\xi \qquad\qquad \text{by monotone convergence} \\
&= \frac{1}{(2\pi)^d} \|\widehat{f}\|^2_{L^2(\mathbb{R}^d)}. \qquad\qquad (17.1)
\end{aligned}$$

By density of $L^1 \cap L^2$ in L^2, the Fourier transform $\widehat{\ }$ extends to a bounded operator from $L^2(\mathbb{R}^d)$ to itself. Plancherel follows from (17.1) by density. Thus the Fourier transform is an isometry, up to a constant factor.

Parseval follows from Plancherel by polarization, or by repeating the argument for Plancherel with $\langle f, f\rangle$ changed to $\langle f, g\rangle$ (and using dominated instead of monotone convergence).

For Inversion, note $\check{\ } : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is bounded by Definition 16.1, since the Fourier transform is bounded. If f is smooth with compact support then $\widehat{f}$ is bounded and decays rapidly at infinity, by repeated use of Theorem 14.11. Hence $\widehat{f} \in L^1(\mathbb{R}^d)$, with $(\hat{f})^{\check{}} = f$ by Inversion Theorem 16.2. So the Fourier transform followed by the inverse transform gives the identity on the dense set $L^1 \cap L^2(\mathbb{R}^d)$, and hence on all of $L^2(\mathbb{R}^d)$ by continuity. Similarly $(\check{g})\hat{\ } = g$ for all $g \in L^2(\mathbb{R}^d)$.

Finally, the Fourier transform is injective by Plancherel, and surjective by Inversion. □

Example 17.2. In 1 dimension, the Dirichlet function

$$D(x) = \frac{\sin x}{\pi x}$$

belongs to $L^2(\mathbb{R})$ and has

$$\widehat{D}(\xi) = \mathbb{1}_{[-1,1]}(\xi).$$

Proof. $D = (\mathbb{1}_{[-1,1]})^{\check{}}$ by definition in (15.1), and so $\widehat{D} = \mathbb{1}_{[-1,1]}$ by Theorem 17.1 Inversion.

dimension	$f(x)$	$\widehat{f}(\xi)$
d	$D(x) = \frac{1}{\pi^d} \prod_{j=1}^{d} \frac{\sin x_j}{x_j}$	$\widehat{D}(\xi) = \mathbb{1}_{[-1,1]^d}(\xi)$

Table 17.1: Fourier transform of the Dirichlet function, from Example 17.2.

Remark 17.3. If $f \in L^2(\mathbb{R}^d)$ then $f\mathbb{1}_{B(n)} \in L^1 \cap L^2(\mathbb{R}^d)$ and $f\mathbb{1}_{B(n)} \to f$ in $L^2(\mathbb{R}^d)$. Hence

$$\begin{aligned} \widehat{f}(\xi) &= \lim_{n\to\infty} \widehat{(f\mathbb{1}_{B(n)})}(\xi) && \text{in } L^2(\mathbb{R}^d), \text{ by Theorem 17.1,} \\ &= \lim_{n\to\infty} \int_{B(n)} f(x) e^{-i\xi x}\, dx. \end{aligned}$$

How can this limit exist, when f need not be integrable? The answer must be that oscillations of $e^{-i\xi x}$ yield cancelations that allow $f(x)e^{-i\xi x}$ to be integrated improperly, as above, for almost every ξ.

Theorem 17.4 (Hausdorff–Young for Fourier transform)**.** *The Fourier transform*

$$\widehat{\ }\colon L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$$

is bounded for $1 \le p \le 2$*, where* $\frac{1}{p} + \frac{1}{p'} = 1$.

Proof. Apply the Riesz–Thorin Interpolation Theorem C.6, using boundedness of

$$\begin{aligned} \widehat{\ }\colon L^1(\mathbb{R}^d) &\to L^\infty(\mathbb{R}^d) && \text{in Theorem 14.2, and} \\ \widehat{\ }\colon L^2(\mathbb{R}^d) &\to L^2(\mathbb{R}^d) && \text{in Theorem 17.1.} \end{aligned}$$

Note the Fourier transform is well defined on $L^1 + L^2(\mathbb{R}^d)$, since the L^1 and L^2 Fourier transforms agree on $L^1 \cap L^2(\mathbb{R}^d)$. □

Remark 17.5. The first five Basic Properties in Theorem 14.2 still hold for the Fourier transform on $L^p(\mathbb{R}^d), 1 \le p \le 2$, and so do Corollary 14.3 (radial functions) and Lemma 14.4 (product functions) and (15.17)–(15.20) (connection to Fourier integrals).

Proof. Given $f \in L^p(\mathbb{R}^d)$, take $f_m \in L^1 \cap L^p(\mathbb{R}^d)$ with $f_m \to f$ in $L^p(\mathbb{R}^d)$. Then $\widehat{f_m} \to \widehat{f}$ in $L^{p'}(\mathbb{R}^d)$ by the Hausdorff–Young Theorem 17.4. Here $\widehat{f_m}$ is the usual Fourier transform of $f_m \in L^1(\mathbb{R}^d)$, so that Theorem 14.2, Corollary 14.3, Lemma 14.4 and (15.17)–(15.20) all apply to f_m. Now let $m \to \infty$ in those results.

Corollary 17.6 (Convolution and Fourier transforms)**.** *If $f \in L^1(\mathbb{R}^d), g \in L^p(\mathbb{R}^d), 1 \le p \le 2$, then $f * g \in L^p(\mathbb{R}^d)$ and*

$$\widehat{(f * g)} = \widehat{f}\,\widehat{g}.$$

Proof. Take $g_m \in L^1 \cap L^p(\mathbb{R}^d)$ with $g_m \to g$ in $L^p(\mathbb{R}^d)$. Then $\widehat{(f * g_m)} = \widehat{f}\,\widehat{g_m}$ by Theorem 17.1. Let $m \to \infty$ and use the Hausdorff–Young Theorem 17.4, noting $\widehat{f}$ is bounded. □

Consequence

Analogue of Weierstrass trigonometric approximation: functions with compactly supported Fourier transform are dense in $L^p(\mathbb{R}^d), 1 \le p \le 2$.

Proof. $F_\omega * f \to f$ in $L^p(\mathbb{R}^d)$ by Theorem 15.7, and $\widehat{(F_\omega * f)} = \widehat{F_\omega}\widehat{f}$ has compact support (because $\widehat{F_\omega}$ has compact support by Table 16.1).

Chapter 18

Fourier integrals: summability pointwise

Goal

Prove sufficient conditions for summability at a single point, and a.e.

Reference

[Grafakos] Sections 2.1b, 3.3b

If $f \in C_0(\mathbb{R}^d)$ then $k_\omega * f \to f$ uniformly by Theorem 15.7(b), and hence convergence holds at every x. But what if f is merely continuous at a *point*?

Theorem 18.1 (Summability at a point)**.** *Assume $\{k_\omega\}$ is a summability kernel. Suppose either $f \in L^1(\mathbb{R}^d)$ and $\{k_\omega\}$ satisfies the L^∞ concentration hypothesis* (SR4)*, or else $f \in L^\infty(\mathbb{R}^d)$.*

If f is continuous at $x_0 \in \mathbb{R}^d$ then $(k_\omega * f)(x_0) \to f(x_0)$ as $\omega \to \infty$.

Proof. Adapt the corresponding result on the torus, Theorem 3.1(a). □

The Poisson and Gauss kernels satisfy (SR4), and so does the Fejér kernel in 1 dimension. More generally, if $k(x) = o(1/|x|^d)$ as $|x| \to \infty$ then $k_\omega(x) = \omega^d k(\omega x)$ satisfies (SR4) (Exercise).

Next we aim at summability a.e., by using maximal functions like we did for Fourier series in Chapter 7.

Definition 18.2. Define the

$$\begin{aligned}
\textit{Dirichlet} \text{ maximal function} \quad & (D^*f)(x) = \sup_{\omega} |(D_\omega * f)(x)| \\
\textit{Fejér} \text{ maximal function} \quad & (F^*f)(x) = \sup_{\omega} |(F_\omega * f)(x)| \\
\textit{Poisson} \text{ maximal function} \quad & (P^*f)(x) = \sup_{\omega} |(P_\omega * f)(x)| \\
\textit{Gauss} \text{ maximal function} \quad & (G^*f)(x) = \sup_{\omega} |(G_\omega * f)(x)| \\
\textit{Lebesgue} \text{ maximal function} \quad & (L^*f)(x) = \sup_{\omega} |(L_\omega * f)(x)|
\end{aligned}$$

where

$$L(x) = \frac{1}{|B(1)|} \mathbb{1}_{B(1)}(x)$$

is the normalized indicator function of the unit ball.

Lemma 18.3.

$$L^*f \le L^*|f| \le Mf$$

where M is the Hardy–Littlewood maximal operator from Chapter 6.

Proof. First,

$$L_{1/\omega}(y) = (1/\omega)^d L(y/\omega) = \frac{1}{|B(\omega)|} \mathbb{1}_{B(\omega)}(y). \tag{18.1}$$

Hence

$$\begin{aligned}
|(L_{1/\omega} * f)(x)| &\le \frac{1}{|B(\omega)|} \int_{B(\omega)} |f(x-y)|\, dy \\
&\le (Mf)(x).
\end{aligned}$$

□

Lemma 18.4 (Majorization)**.** *If $k \in L^1(\mathbb{R}^d)$ is nonnegative and radially symmetric decreasing, then*

$$|(k * f)(x)| \le \|k\|_{L^1(\mathbb{R}^d)} (L^*f)(x) \qquad \textit{for all } x \in \mathbb{R}^d, \quad f \in L^1(\mathbb{R}^d).$$

Proof. Write $k(x) = \rho(|x|)$ where $\rho : [0, \infty) \to \mathbb{R}$ is nonnegative and decreasing. Assume ρ is absolutely continuous, for simplicity. We first establish a layer-cake decomposition of k, like we did on the torus in Lemma 7.2:

$$\begin{aligned} k(y) = \rho(|y|) &= -\int_{|y|}^{\infty} \rho'(\omega)\, d\omega \qquad \text{since } \rho(\infty) = 0 \text{ by integrability of } k \\ &= -\int_0^{\infty} |B(\omega)| L_{1/\omega}(y) \rho'(\omega)\, d\omega, \end{aligned}$$

because by (18.1),

$$L_{1/\omega}(y) = \begin{cases} 1/|B(\omega)| & \text{if } \omega > |y|, \\ 0 & \text{if } \omega \le |y|. \end{cases}$$

Hence

$$(k * f)(x) = \int_0^{\infty} |B(\omega)| (L_{1/\omega} * f)(x) \big(-\rho'(\omega) \big)\, d\omega$$

and so

$$\begin{aligned} |(k * f)(x)| &\le \int_0^{\infty} |B(\omega)| \big(-\rho'(\omega) \big)\, d\omega \cdot (L^* f)(x) \qquad \text{since } \rho' \le 0 \\ &= \int_0^{\infty} \int_0^{\omega} |\partial B(1)| r^{d-1}\, dr \, \big(-\rho'(\omega) \big)\, d\omega \cdot (L^* f)(x) \\ &\qquad \text{by spherical coordinates for } |B(\omega)| = \int_{B(\omega)} dy \\ &= \int_0^{\infty} |\partial B(1)| \omega^{d-1} \rho(\omega)\, d\omega \cdot (L^* f)(x) \\ &\qquad \text{by parts with respect to } \omega \text{ (why does the } \omega = \infty \text{ term vanish?)} \\ &= \int_{\mathbb{R}^d} k(y)\, dy \cdot (L^* f)(x) \end{aligned}$$

by using spherical coordinates again. □

Theorem 18.5 (Lebesgue dominates Poisson and Gauss in all dimensions, and Fejér in 1 dimension)**.**

$$\begin{aligned} F^* f &\le \frac{4}{\pi} L^* |f| \qquad \textit{(when } d = 1\textit{)} \\ P^* f &\le L^* f \\ G^* f &\le L^* f \end{aligned}$$

for all $f \in L^p(\mathbb{R}^d), 1 \le p \le \infty$.

Proof. $P^*f \le L^*f$ by the Majorization Lemma 18.4, since P_ω is nonnegative and radially symmetric decreasing, with $\|P_\omega\|_{L^1(\mathbb{R}^d)} = 1$. Similarly $G^*f \le L^*f$.

When $d = 1$,

$$F_\omega(x) = \frac{\omega}{2\pi}\left(\frac{\sin\left(\frac{\omega}{2}x\right)}{\frac{\omega}{2}x}\right)^2 \qquad \text{by (15.8)}$$
$$\le k(x) \stackrel{\text{def}}{=} \frac{\omega}{2\pi}\begin{cases}1, & |x| \le 2/\omega,\\ 1/\left(\frac{\omega}{2}x\right)^2, & |x| > 2/\omega.\end{cases}$$

Note k is nonnegative, even and decreasing, with $\|k\|_{L^1(\mathbb{R})} = 4/\pi$. Hence $|F_\omega * f| \le k * |f| \le (4/\pi)L^*|f|$ by Majorization Lemma 18.4. □

Remark 18.6. The Fejér kernel is *not* majorized by a radially symmetric decreasing integrable function, when $d \ge 2$. For example, taking $\omega = 2$ gives

$$F_2(x) = \prod_{j=1}^{d} \frac{1}{\pi}\left(\frac{\sin x_j}{x_j}\right)^2,$$

which decays like x_1^{-2} along the x_1-axis. Thus the best possible radial bound would be $O(|x|^{-2})$, which is not integrable at infinity in dimensions $d \ge 2$.

Corollary 18.7. *F^*, P^*, G^* and L^* are weak $(1,1)$ and strong (p,p) on $L^p(\mathbb{R}^d)$, for $1 < p \le \infty$.*

Proof. Combine Theorem 18.5 and Lemma 18.3 with the weak and strong bounds on the Hardy–Littlewood maximal operator in Chapter 6.

For the Fejér kernel in dimensions $d \ge 2$, see [Grafakos, Theorem 3.3.3]. □

Theorem 18.8 (Summability a.e.)**.** *If $f \in L^p(\mathbb{R}^d), 1 \le p \le \infty$, then*

$$F_\omega * f \to f \text{ a.e. as } \omega \to \infty,$$
$$P_\omega * f \to f \text{ a.e. as } \omega \to \infty,$$
$$G_\omega * f \to f \text{ a.e. as } \omega \to \infty,$$
$$L_\omega * f \to f \text{ a.e. as } \omega \to \infty.$$

(The last statement is the Lebesgue differentiation theorem.)

Proof. Assume $1 \le p < \infty$. F^* is weak (p,p) by Corollary 18.7. Hence the Theorem in Chapter 6 says

$$\mathcal{C} = \{f \in L^p(\mathbb{R}^d) : \lim_{\omega\to\infty} F_\omega * f = f \text{ a.e.}\}$$

is closed in $L^p(\mathbb{R}^d)$. Obviously $\mathcal{C}$ contains every $f \in C_c(\mathbb{R}^d)$, because $F_\omega * f \to f$ uniformly by Theorem 15.7. Thus $\mathcal{C}$ is dense in $L^p(\mathbb{R}^d)$ (using here that $p < \infty$). Because $\mathcal{C}$ is closed, it must equal $L^p(\mathbb{R}^d)$, which proves the result.

When $p = \infty$, consider $f \in L^\infty(\mathbb{R}^d)$. For $m \in \mathbb{N}$, put $g = \mathbb{1}_{B(m)} f$ and $h = f - g$. Then $g \in L^1(\mathbb{R}^d)$, and so $F_\omega * g \to g$ a.e., by the part of the theorem already proved. Hence $F_\omega * g \to f$ a.e. on $B(m)$. Next $h \in L^\infty(\mathbb{R}^d)$ is continuous on $B(m)$, with $h = 0$ there, and so $F_\omega * h \to h = 0$ on $B(m)$ by Theorem 18.1. Since $f = g + h$ we deduce $F_\omega * f \to f$ a.e. on $B(m)$. Letting $m \to \infty$ proves the result.

Argue similarly for the other kernels. □

Chapter 19

Fourier integrals: norm convergence

Goal

Show norm convergence for $L^p(\mathbb{R}^d)$ follows from boundedness of the Hilbert transform on $\mathbb{R}$

Reference

I do not know a fully satisfactory reference for this material. Suggestions are welcome!

Definition 19.1. Write

$$S_\omega(f) = D_\omega * f$$

where

$$D_\omega(x) = \prod_{j=1}^{d} \omega D(\omega x_j) = \prod_{j=1}^{d} \frac{\sin(\omega x_j)}{\pi x_j}$$

is the Dirichlet kernel on $\mathbb{R}^d$ and $D(z) = (\sin z)/\pi z$ is the Dirichlet function in 1 dimension.

S_ω is the "partial sum" operator for the Fourier integral, because if $f \in L^p(\mathbb{R}^d), 1 \le p \le 2$, then $S_\omega(f) = (\mathbb{1}_{[-\omega,\omega]^d} \hat{f})\check{}$ by (15.17) and Remark 17.5. In particular,

$$S_\omega : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$$

is bounded, by boundedness of the Fourier transform and its inverse on L^2. Further, $\mathbb{1}_{[-\omega,\omega]^d}\widehat{f} \to \widehat{f}$ and so $S_\omega(f) \to f$ in $L^2(\mathbb{R}^d)$, as $\omega \to \infty$.

$S_\omega(f)$ is well defined whenever $f \in L^p(\mathbb{R}^d), 1 \le p < \infty$, because $D_\omega \in L^q(\mathbb{R}^d)$ for each $q > 1$ and so $D_\omega * f \in L^r(\mathbb{R}^d)$ for each $r \in (p, \infty]$, by the Generalized Young's Theorem in Chapter 13.

We will prove below that $S_\omega(f) \in L^p(\mathbb{R}^d)$ when $f \in L^p(\mathbb{R}^d), 1 < p < \infty$. But $S_\omega(f)$ need not belong to $L^1(\mathbb{R}^d)$ when $f \in L^1(\mathbb{R}^d)$ (Exercise).

Our goal in this Chapter is to improve the L^p summability for Fourier integrals ($F_\omega * f \to f$ in Theorem 15.7) to L^p convergence ($D_\omega * f = S_\omega(f) \to f$ in Theorem 19.4 below). As remarked above, we have the result already for $p = 2$.

First we reduce norm convergence to norm boundedness.

Theorem 19.2. *Let $1 < p < \infty$ and suppose $\sup_\omega \|S_\omega\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} < \infty$. Then Fourier integrals converge in $L^p(\mathbb{R}^d)$: $\lim_{\omega\to\infty}\|S_\omega(f) - f\|_{L^p(\mathbb{R}^d)} = 0$ for each $f \in L^p(\mathbb{R}^d)$.*

Proof. Let $\mathcal{A} = \{g \in L^1 \cap L^p(\mathbb{R}^d) : \widehat{g} \text{ has compact support}\}$. We claim $\mathcal{A}$ is dense in $L^p(\mathbb{R}^d)$. Indeed, if $f \in L^1 \cap L^p(\mathbb{R}^d)$ then $F_\omega * f \in L^1 \cap L^p(\mathbb{R}^d)$ and $\widehat{(F_\omega * f)} = \widehat{F_\omega}\widehat{f}$ has compact support by Table 16.1. Thus $F_\omega * f \in \mathcal{A}$. Since $F_\omega * f \to f$ in $L^p(\mathbb{R}^d)$ by Theorem 15.7, and $L^1 \cap L^p$ is dense in L^p, we see $\mathcal{A}$ is dense in $L^p(\mathbb{R}^d)$.

We further show $S_\omega(g) = g$, when $g \in \mathcal{A}$, provided ω is large enough that $[-\omega, \omega]^d$ contains the support of $\widehat{g}$. To see this fact, note $S_\omega(g) \in L^2(\mathbb{R}^d)$ because $D_\omega \in L^2(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$; thus

$$\begin{aligned}\widehat{S_\omega(g)} &= \widehat{D_\omega}\widehat{g} \\ &= \mathbb{1}_{[-\omega,\omega]^d}\widehat{g} && \text{by Table 17.1} \\ &= \widehat{g}.\end{aligned}$$

Applying Fourier inversion in L^2 gives $S_\omega(g) = g$.

We conclude

$$\mathcal{A} \subset \{f \in L^p(\mathbb{R}^d) : \lim_{\omega\to\infty} S_\omega(f) = f \text{ in } L^p(\mathbb{R}^d)\} \stackrel{\text{def}}{=} \mathcal{C},$$

so that $\mathcal{C}$ is dense in $L^p(\mathbb{R}^d)$. Because $\mathcal{C}$ is closed by Proposition 9.2 (using the assumption that $\sup_\omega \|S_\omega\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} < \infty$), we conclude $\mathcal{C} = L^p(\mathbb{R}^d)$, which proves the theorem. □

Next we reduce to norm boundedness in 1 dimension. For the sake of generality we allow different ω-values in each coordinate direction. (Thus our "square partial sums" for convergence of Fourier integrals can be relaxed to "rectangular partial sums"; proof omitted.)

Given a vector $\vec{\omega} = (\omega_1, \ldots, \omega_d)$ of positive numbers, define

$$D_{\vec{\omega}}(x) = \prod_{j=1}^{d} \omega_j D(\omega_j x_j).$$

The Fourier multiplier

$$\widehat{D_{\vec{\omega}}} = \mathbb{1}_{[-\omega_1,\omega_1]\times\cdots\times[-\omega_d,\omega_d]}$$

is the indicator function of a rectangular box.

Write

$$C_{p,d} = \sup_{\vec{\omega}} \|S_{\vec{\omega}}\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)}$$

for the norm bound on the partial sum operators. We have not yet shown that this constant is finite.

Theorem 19.3 (Reduction to 1 dimension)**.** $C_{p,d} \le (C_{p,1})^d$.

Proof. First observe that for $g \in L^p(\mathbb{R})$ and $\omega > 0$,

$$\begin{aligned}
\int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \omega D(\omega(x-y)) g(y)\, dy \Big|^p dx &= \|D_\omega * g\|^p_{L^p(\mathbb{R})} \\
&\le C^p_{p,1} \|g\|^p_{L^p(\mathbb{R})} \qquad \text{by definition of } C_{p,1} \\
&= C^p_{p,1} \int_{\mathbb{R}} |g(y)|^p\, dy. \qquad (19.1)
\end{aligned}$$

Hence for $f \in L^p(\mathbb{R}^2)$ and $\vec{\omega} = (\omega_1, \omega_2)$,

$$\begin{aligned}
&\int_{\mathbb{R}^2} |(D_{\vec{\omega}} * f)(x_1, x_2)|^p \, dx_1 dx_2 \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \omega_1 D(\omega_1(x_1 - y_1)) \int_{\mathbb{R}} \omega_2 D(\omega_2(x_2 - y_2)) f(y_1, y_2) \, dy_2 dy_1 \Big|^p \, dx_1 dx_2 \\
&\le C_{p,1}^p \int_{\mathbb{R}} \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \omega_2 D(\omega_2(x_2 - y_2)) f(y_1, y_2) \, dy_2 \Big|^p \, dy_1 dx_2 \\
&\qquad\qquad \text{by (19.1) with } g(y_1) = \int_{\mathbb{R}} \omega_2 D(\omega_2(x_2 - y_2)) f(y_1, y_2) \, dy_2 \\
&\le C_{p,1}^{2p} \int_{\mathbb{R}} \int_{\mathbb{R}} |f(y_1, y_2)|^p \, dy_2 dy_1 \\
&\qquad\qquad \text{by (19.1) with } g(y_2) = f(y_1, y_2).
\end{aligned}$$

Taking p-th roots gives $\|S_{\vec{\omega}}\|_{L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2)} \le C_{p,1}^2$, which proves the theorem when $d = 2$.

Argue similarly for $d \ge 3$. □

Aside. The "ball" multiplier $\mathbb{1}_{B(1)}(\xi)$ does *not* yield a partial sum operator with uniform norm bounds, when $p \neq 2$; see [Grafakos, Section 10.1]. Therefore Fourier integrals and series in higher dimensions should be evaluated with "rectangular" partial sums, and not "spherical" sums, when working in L^p for $p \neq 2$.

Boundedness in $L^p(\mathbb{R})$

1. We shall prove (in Chapters 20 and 21) the existence of a bounded linear operator

$$H : L^p(\mathbb{R}) \to L^p(\mathbb{R}), \qquad 1 < p < \infty,$$

called the **Hilbert transform** on $\mathbb{R}$, with the property that

$$\widehat{(Hf)}(\xi) = -i \operatorname{sign}(\xi) \widehat{f}(\xi)$$

when $f \in L^p \cap L^2(\mathbb{R})$. (Thus H is a Fourier multiplier operator.)

2. Then the *Riesz projection* $P : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ defined by

$$Pf = \frac{1}{2}(f + iHf)$$

is also bounded, when $1 < p < \infty$.

Observe P projects onto the positive frequencies:

$$\widehat{(Pf)}(\xi) = \mathbb{1}_{(0,\infty)}(\xi)\widehat{f}(\xi), \qquad f \in L^2(\mathbb{R}),$$

since $i(-i\,\mathrm{sign}(\xi)) = \mathrm{sign}(\xi)$.

3. The following formula expresses the Fourier partial sum operator in terms of the Riesz projection and some modulations: for $\omega > 0$,

$$e^{-i\omega x}P(e^{i\omega x}f) - e^{i\omega x}P(e^{-i\omega x}f) = S_\omega(f), \qquad f \in L^2(\mathbb{R}). \tag{19.2}$$

Proof.

$$\begin{aligned}
[e^{i\omega x}f]\widehat{\ }(\xi) &= \widehat{f}(\xi-\omega)\\
[P(e^{i\omega x}f)]\widehat{\ }(\xi) &= \mathbb{1}_{(0,\infty)}(\xi)\widehat{f}(\xi-\omega)\\
[e^{-i\omega x}P(e^{i\omega x}f)]\widehat{\ }(\xi) &= \mathbb{1}_{(0,\infty)}(\omega+\xi)\widehat{f}(\xi)\\
&= \mathbb{1}_{(-\omega,\infty)}(\xi)\widehat{f}(\xi)\\
[e^{i\omega x}P(e^{-i\omega x}f)]\widehat{\ }(\xi) &= \mathbb{1}_{(\omega,\infty)}(\xi)\widehat{f}(\xi)
\end{aligned}$$

Subtracting the last two formulas gives $\mathbb{1}_{(-\omega,\omega]}\widehat{f}$, which equals $\widehat{S_\omega(f)}$. Fourier inversion now completes the proof.

4. From (19.2) applied to the dense class of $f \in L^p \cap L^2(\mathbb{R})$, and from boundedness of the Riesz projection, it follows that

$$C_{p,1} = \sup_\omega \|S_\omega\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})} \le 2\|P\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})} < \infty$$

when $1 < p < \infty$. Hence from Theorems 19.2 and 19.3 we conclude:

Theorem 19.4 (Fourier integrals converge in $L^p(\mathbb{R}^d)$)**.** *Let $1 < p < \infty$. Then*

$$\lim_{\omega\to\infty} \|S_\omega(f) - f\|_{L^p(\mathbb{R}^d)} = 0 \qquad \textit{for each } f \in L^p(\mathbb{R}^d).$$

It remains to prove L^p boundedness of the Hilbert transform on $\mathbb{R}$.

Chapter 20

Hilbert and Riesz transforms on $L^2(\mathbb{R}^d)$

Goal

Develop spatial and frequency representations of Hilbert and Riesz transforms

Reference

[Duoandikoetxea] Section 4.3
[Grafakos] Section 4.1

Definition 20.1. The Riesz transforms on $\mathbb{R}^d$ are

$$\begin{aligned} R_j : L^2(\mathbb{R}^d) &\to L^2(\mathbb{R}^d) \\ f &\mapsto (-i(\xi_j/|\xi|)\widehat{f}\,)^\vee \end{aligned}$$

for $j = 1, \ldots, d$.

In dimension $d = 1$, the Riesz transform equals the Hilbert transform on $\mathbb{R}$, defined by

$$\begin{aligned} H : L^2(\mathbb{R}) &\to L^2(\mathbb{R}) \\ f &\mapsto (-i\,\mathrm{sign}(\xi)\widehat{f}\,)^\vee \end{aligned}$$

because $\mathrm{sign}(\xi) = \xi/|\xi|$.

R_j is bounded since the Fourier multiplier $-i\xi_j/|\xi|$ is a bounded function (in fact, bounded by 1). Clearly

$$\begin{aligned} \|R_j\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} &\le 1 && \text{by Plancherel,} \\ \sum_{j=1}^{d} R_j^2 &= -I && \text{since } \sum_{j=1}^{d}(-i\xi_j/|\xi|)^2 = -1, \\ R_j^* &= -R_j && \text{by Parseval.} \end{aligned}$$

Proposition 20.2 (Spatial representation of Hilbert transform)**.** *If $f \in L^2(\mathbb{R})$ is C^1-smooth on an interval then*

$$(Hf)(x) = p.v. \int_{\mathbb{R}} f(x-y)\frac{1}{\pi y}\,dy \tag{20.1}$$

for almost every x in the interval.

The proposition says formally that

$$Hf = f * \frac{1}{\pi x}$$

or

$$\big(\text{p.v. } \frac{1}{\pi x}\big)\widehat{\ } = -i\,\text{sign}(\xi).$$

Later we will justify these formulas in terms of distributions.

The right side of (20.1) is a *singular integral*, since the convolution kernel $1/\pi y$ is not integrable.

Proof. [This proof is similar to Proposition 10.3 on $\mathbb{T}$, and so was skimmed only lightly in class.] For $\omega > 0$,

$$\begin{aligned} \frac{1}{2\pi}\int_{[-\omega,\omega]} (-i)\,\text{sign}(\xi) e^{i\xi y}\,d\xi &= \frac{i}{2\pi}\int_{-\omega}^{0} e^{i\xi y}\,d\xi - \frac{i}{2\pi}\int_{0}^{\omega} e^{i\xi y}\,d\xi \\ &= \frac{1-\cos(\omega y)}{\pi y}. \end{aligned} \tag{20.2}$$

If $f \in L^1 \cap L^2(\mathbb{R})$ then

$$\begin{aligned}
(\mathbb{1}_{[-\omega,\omega]}\widehat{Hf})\check{}(x)\,d\xi &= \frac{1}{2\pi}\int_{[-\omega,\omega]} (-i)\,\mathrm{sign}(\xi)\widehat{f}(\xi)e^{i\xi x}\,d\xi \\
&= \int_{\mathbb{R}} f(y)\frac{1}{2\pi}\int_{[-\omega,\omega]} (-i)\,\mathrm{sign}(\xi)e^{i\xi(x-y)}\,d\xi dy \qquad \text{by Fubini} \\
&= \int_{\mathbb{R}} f(x-y)\frac{1-\cos(\omega y)}{\pi y}\,dy \qquad \text{by } y \mapsto x-y \text{ and (20.2)} \\
&= \int_{|y|<1} [f(x-y)-f(x)]\frac{1-\cos(\omega y)}{\pi y}\,dy \\
&+ \int_{|y|>1} f(x-y)\frac{1-\cos(\omega y)}{\pi y}\,dy
\end{aligned}$$

by oddness of $\big(1-\cos(\omega y)\big)/\pi y$. The second integral converges to

$$\int_{|y|>1} f(x-y)\frac{1}{\pi y}\,dy \tag{20.3}$$

as $\omega \to \infty$, by the Riemann–Lebesgue Corollary 14.7. The first integral similarly converges to

$$\int_{|y|<1} [f(x-y)-f(x)]\frac{1}{\pi y}\,dy, \tag{20.4}$$

assuming f is C^1-smooth on a neighborhood of x (which ensures integrability of $y \mapsto [f(x-y)-f(x)]/\pi y$ on $|y|<1$).

Meanwhile, $\mathbb{1}_{[-\omega,\omega]}\widehat{Hf}$ converges to $\widehat{Hf}$ in $L^2(\mathbb{R})$ as $\omega \to \infty$, so that $(\mathbb{1}_{[-\omega,\omega]}\widehat{Hf})\check{}$ converges to Hf. Convergence holds a.e. for some subsequence of ω-values. Formula (20.1) therefore follows from (20.3) and (20.4), since $\int_{\varepsilon<|y|<1}(1/\pi y)\,dy = 0$.

Finally, one deduces (20.1) in full generality by approximating f off a neighborhood of x using functions in $L^1 \cap L^2$. (Obviously f belongs to $L^1 \cap L^2$ already on each neighborhood of x.) □

Proposition 20.3 (Spatial representation of Riesz transform)**.** *If $f \in L^2(\mathbb{R}^d)$ is C^1-smooth on an open set $U \subset \mathbb{R}^d$ then*

$$(R_j f)(x) = p.v. \int_{\mathbb{R}^d} f(x-y)\frac{c_d\, y_j}{|y|^{d+1}}\,dy$$

for almost every $x \in U$, for each $j = 1, \ldots, d$.

Here $c_d = \Gamma\big((d+1)/2\big)/\pi^{(d+1)/2} > 0$. For example, $c_1 = 1/\pi$.

The proposition says formally that

$$R_j f = f * \frac{c_d\, y_j}{|y|^{d+1}}$$

or

$$\big(\text{p.v. } \frac{c_d\, y_j}{|y|^{d+1}}\big)\widehat{\ } = -i\frac{\xi_j}{|\xi|}.$$

Proof. To motivate the following proof, observe

$$\frac{1}{|\xi|} = \int_0^\infty e^{-|\xi| z}\, dz \tag{20.5}$$

and that $e^{-|\xi|z}$ is the Fourier transform of the Poisson kernel $P_{1/z}$. Our proof will use a truncated version of this identity:

$$\frac{e^{-|\xi|\delta} - e^{-|\xi|/\delta}}{|\xi|} = \int_\delta^{1/\delta} e^{-|\xi| z}\, dz. \tag{20.6}$$

In class we proceeded formally, skipping the rest of this proof and using (20.5) instead of (20.6) in the proof of Lemma 20.4 below.

For $f \in L^2(\mathbb{R}^d)$,

$$\begin{aligned}\widehat{(R_j f)}(\xi) &= -i\frac{\xi_j}{|\xi|}\widehat{f}(\xi) \\ &= \lim_{\delta\to 0}(-i\xi_j)\frac{e^{-|\xi|\delta} - e^{-|\xi|/\delta}}{|\xi|}\widehat{f}(\xi)\end{aligned}$$

with convergence in $L^2(\mathbb{R}^d)$ (by dominated convergence). Applying L^2 Fourier inversion yields

$$(R_j f)(x) = \lim_{\delta\to 0}\left(-i\xi_j\frac{e^{-|\xi|\delta} - e^{-|\xi|/\delta}}{|\xi|}\hat{f}\right)\check{}\,(x)$$

in $L^2(\mathbb{R}^d)$, and hence pointwise a.e. for some subsequence of δ values. Thus the theorem is proved when $f \in L^1 \cap L^2(\mathbb{R}^d)$, by Lemma 20.4 below.

Finally, one deduces the theorem for $f \in L^2(\mathbb{R}^d)$ by approximating f off a neighborhood of x using functions in $L^1 \cap L^2$. (Obviously f belongs to $L^1 \cap L^2$ already on each neighborhood of x.) □

Lemma 20.4. *If $f \in L^1 \cap L^2(\mathbb{R}^d)$ is C^1-smooth on an open set $U \subset \mathbb{R}^d$, then*

$$\begin{aligned}&\lim_{\delta\to 0}\left(-i\xi_j\frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|}\hat{f}\right)\check{}\,(x)\\&=\int_{|y|<1}[f(x-y)-f(x)]\frac{c_d\,y_j}{|y|^{d+1}}\,dy+\int_{|y|>1}f(x-y)\frac{c_d\,y_j}{|y|^{d+1}}\,dy\end{aligned}\tag{20.7}$$

for almost every $x \in U$. Further, the first integral in (20.7) *equals*

$$\lim_{\varepsilon\to 0}\int_{\varepsilon<|y|<1}f(x-y)\frac{c_d\,y_j}{|y|^{d+1}}\,dy.$$

Proof. First, $\xi_j/|\xi|$ is bounded by 1, and the exponentials $e^{-|\xi|\delta}$ and $e^{-|\xi|/\delta}$ are square integrable, and so is $\widehat{f}$. Thus their product is integrable, so that by the L^1 Fourier Inversion Theorem 16.2 (and the definition of $\widehat{f}$ for $f \in L^1(\mathbb{R}^d)$),

$$\begin{aligned}&\left(-i\xi_j\frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|}\hat{f}\right)\check{}\,(x)\\&=-\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}i\xi_j\frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|}\int_{\mathbb{R}^d}f(y)e^{-i\xi y}\,dy\,e^{i\xi x}\,d\xi\\&=-\int_{\mathbb{R}^d}f(x-y)\,\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}i\xi_j\frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|}e^{i\xi y}\,d\xi dy\end{aligned}$$

after changing $y \mapsto x - y$. To evaluate the inner integral, we express it using

Poisson kernels:

$$\begin{aligned}
&\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} i\xi_j \frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|} e^{i\xi y}\,d\xi \\
&= \frac{\partial}{\partial y_j}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|} e^{i\xi y}\,d\xi \\
&= \int_\delta^{1/\delta} \frac{\partial}{\partial y_j}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|z} e^{i\xi y}\,d\xi dz && \text{by identity (20.6)} \\
&= \int_\delta^{1/\delta} \frac{\partial}{\partial y_j} P_{1/z}(y)\,dz && \text{by (15.11)} \\
&= \int_\delta^{1/\delta} c_d\, z\, \frac{\partial}{\partial y_j}\,\frac{1}{(|y|^2+z^2)^{(d+1)/2}}\,dz && \text{by (15.12)} \\
&= \int_\delta^{1/\delta} c_d\, y_j\, \frac{\partial}{\partial z}\,\frac{1}{(|y|^2+z^2)^{(d+1)/2}}\,dz && \text{(why?!)} \\
&= \frac{c_d y_j}{(|y|^2+z^2)^{(d+1)/2}}\bigg|_{z=\delta}^{z=1/\delta}.
\end{aligned}$$

By substituting this expression into the above, we find

$$\begin{aligned}
&\left(-i\xi_j \frac{e^{-|\xi|\delta}-e^{-|\xi|/\delta}}{|\xi|}\hat{f}\right)\check{}\,(x) \\
&= -\int_{\mathbb{R}^d} f(x-y)\,\frac{c_d\, y_j}{(|y|^2+z^2)^{(d+1)/2}}\bigg|_{z=\delta}^{z=1/\delta} dy \\
&= -\int_{|y|<1} [f(x-y)-f(x)]\,\frac{c_d\, y_j}{(|y|^2+z^2)^{(d+1)/2}}\bigg|_{z=\delta}^{z=1/\delta} dy && (20.8) \\
&\quad -\int_{|y|>1} f(x-y)\,\frac{c_d\, y_j}{(|y|^2+z^2)^{(d+1)/2}}\bigg|_{z=\delta}^{z=1/\delta} dy && (20.9)
\end{aligned}$$

where we used the oddness of y_j to insert $f(x)$ in (20.8).

Now fix a point $x \in U$. As $\delta \to 0$, expression (20.8) converges to

$$\int_{|y|<1} [f(x-y)-f(x)]\frac{c_d\, y_j}{|y|^{d+1}}\,dy$$

by dominated convergence (noting the C^1-smoothness ensures the integrand is $O(|y|)\cdot O(1/|y|^d) = O(1/|y|^{d-1})$ near the origin, which is integrable). And

as $\delta \to 0$, expression (20.9) converges to

$$\int_{|y|>1} f(x-y) \frac{c_d\, y_j}{|y|^{d+1}}\, dy$$

by dominated convergence (noting $f \in L^2(\mathbb{R}^d)$ and $y_j/|y|^{d+1} = O(1/|y|^d)$ is square integrable for $|y| > 1$). (Exercise: explain why the terms with $z = 1/\delta$ in (20.8) and (20.9) vanish as $\delta \to 0$, using dominated convergence.)

For the final claim in the lemma, write $\int_{|y|<1} = \lim_{\varepsilon \to 0} \int_{\varepsilon<|y|<1}$ and use the oddness of y_j to remove the term with $f(x)$. □

Connections to PDEs

1. The Riesz transforms map the *normal* derivative of a harmonic function to its *tangential* derivatives.

Formal Proof. Given a function f, let

$$u(x, x_{d+1}) = (P_{1/x_{d+1}} * f)(x), \qquad x \in \mathbb{R}^d, \quad x_{d+1} > 0,$$

so that u is harmonic on the upper halfspace $\mathbb{R}^d \times (0, \infty)$ with boundary value $u = f$ when $x_{d+1} = 0$ (see Chapter 15). Put

$$v(x) = \frac{\partial}{\partial x_{d+1}} u(x, x_{d+1}) \bigg|_{x_{d+1}=0} = \text{normal derivative of } u \text{ at the boundary.}$$

Then

$$R_j v = \frac{\partial f}{\partial x_j}, \qquad j = 1, \ldots, d,$$

because

$$\begin{aligned}
\widehat{(R_j v)}(\xi) &= -i\frac{\xi_j}{|\xi|}\widehat{v}(\xi) \\
&= -i\frac{\xi_j}{|\xi|}\,\frac{\partial}{\partial x_{d+1}}\widehat{u}(\xi, x_{d+1})\Big|_{x_{d+1}=0} \\
&= -i\frac{\xi_j}{|\xi|}\,\frac{\partial}{\partial x_{d+1}}\big(e^{-|\xi|x_{d+1}}\widehat{f}(\xi)\big)\Big|_{x_{d+1}=0} \\
&= -i\frac{\xi_j}{|\xi|}(-|\xi|)\widehat{f}(\xi) \\
&= i\xi_j\widehat{f}(\xi) \\
&= \left(\frac{\partial f}{\partial x_j}\right)\widehat{\ }(\xi).
\end{aligned}$$

Thus we have shown the jth Riesz transform maps the normal derivative of u to its jth tangential derivative, on the boundary.

2. Mixed Riesz transforms map the Laplacian to mixed partial derivatives.

Formal Proof.

$$\left(\frac{\partial^2 f}{\partial x_j^2}\right)\widehat{\ }(\xi) = (i\xi_j)^2\widehat{f}(\xi) = -\xi_j^2\widehat{f}(\xi)$$

and so summing over j gives

$$(\Delta f)\widehat{\ }(\xi) = -|\xi|^2\widehat{f}(\xi).$$

Hence

$$\begin{aligned}
(R_j R_k \Delta f)\widehat{\ }(\xi) &= \frac{(-i\xi_j)}{|\xi|}\frac{(-i\xi_k)}{|\xi|}(-|\xi|^2)\widehat{f}(\xi) \\
&= -(i\xi_j)(i\xi_k)\widehat{f}(\xi) \\
&= -\left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)\widehat{\ }(\xi)
\end{aligned}$$

so that

$$R_j R_k \Delta f = -\frac{\partial^2 f}{\partial x_j \partial x_k}.$$

That is, mixed Riesz transforms map the Laplacian to mixed partial derivatives.

The above formal derivation is rigorous if, for example, f is C^2-smooth with compact support.

Consequently, the norm of a mixed second derivative is controlled by the norms of the pure second derivatives in the Laplacian, with

$$\left\| \frac{\partial^2 f}{\partial x_j \partial x_k} \right\|_{L^2(\mathbb{R}^d)} \leq \|\Delta f\|_{L^2(\mathbb{R}^d)}$$

since each Riesz transform has norm 1 on $L^2(\mathbb{R}^d)$. Similar estimates hold on $L^p(\mathbb{R}^d), 1 < p < \infty$, by the L^p boundedness of the Riesz transform proved in the next chapter.

Chapter 21

Hilbert and Riesz transforms on $L^p(\mathbb{R}^d)$

Goal

Prove weak $(1,1)$ for Riesz transform, and deduce strong (p,p) by interpolation and duality

Reference

[Duoandikoetxea] Section 5.1

Theorem 21.1 (weak $(1,1)$ on $L^1 \cap L^2(\mathbb{R}^d)$)**.** *There exists $A > 0$ such that*

$$|\{x \in \mathbb{R}^d : |(R_j f)(x)| > \omega\}| \leq \frac{A}{\omega} \|f\|_{L^1(\mathbb{R}^d)}$$

for all $\omega > 0, j = 1, \ldots, d$ and $f \in L^1 \cap L^2(\mathbb{R}^d)$.

Proof. Apply the Calderón–Zygmund Theorem 11.4 to get $f = g + b$. Note $g \in L^1 \cap L^\infty(\mathbb{R}^d)$ and so $g \in L^2(\mathbb{R}^d)$, hence $R_j g \in L^2(\mathbb{R}^d)$ by Chapter 20. And $b = f - g \in L^2(\mathbb{R}^d)$ so that $R_j b \in L^2(\mathbb{R}^d)$.

Now proceed like in the proof of Theorem 12.1, just changing $\mathbb{T}$ to $\mathbb{R}^d$ and the interval $I(l)$ to the cube $Q(l)$. To finish the proof, we want to show

$$\sum_l \int_{\mathbb{R}^d \setminus 2\sqrt{d}Q(l)} |(R_j b_l)(x)|\, dx \leq (\text{const.}) \|f\|_{L^1(\mathbb{R}^d)}. \tag{21.1}$$

By Proposition 20.3 applied on the open set $U = \mathbb{R}^d \backslash 2\sqrt{d}Q(l)$ (where $b_l = 0$), we have

$$\begin{aligned}
&\int_{\mathbb{R}^d\setminus 2\sqrt{d}Q(l)} |R_j b_l(x)|\,dx \\
&= \int_{\mathbb{R}^d\setminus 2\sqrt{d}Q(l)} \Big| \int_{Q(l)} b_l(y) \frac{c_d(x_j - y_j)}{|x-y|^{d+1}}\,dy \Big|\,dx \\
&\qquad \text{noting } x-y \text{ is bounded away from 0, since } y \in Q(l) \text{ and } x \notin 2\sqrt{d}Q(l), \\
&= \int_{\mathbb{R}^d\setminus 2\sqrt{d}Q(l)} \Big| \int_{Q(l)} b_l(y) \big[\rho_j(x-y) - \rho_j(x-c(l))\big]\,dy \Big|\,dx
\end{aligned}$$

where

$$\rho_j(x) = c_d \frac{x_j}{|x|^{d+1}}$$

is the jth Riesz kernel and $c(l)$ is the center of $Q(l)$; here we used that $\int_{Q(l)} b_l(y)\,dy = 0$. Hence

$$\begin{aligned}
&\int_{\mathbb{R}^d\setminus 2\sqrt{d}Q(l)} |R_j b_l(x)|\,dx \\
&\le \int_{Q(l)} |b_l(y)| \int_{\mathbb{R}^d\setminus 2\sqrt{d}Q(l)} |\rho_j(x-y) - \rho_j(x-c(l))|\,dx dy \\
&\le (\text{const.}) \int_{Q(l)} |b_l(y)|\,dy \qquad (21.2)
\end{aligned}$$

by Lemma 21.2 below; the hypotheses of that lemma are satisfied here because

$$|(\nabla \rho_j)(x)| \le \frac{(\text{const.})}{|x|^{d+1}}$$

and if $x \in \mathbb{R}^d \setminus 2\sqrt{d}Q(l)$ and $y \in Q(l)$ then

$$\begin{aligned}
|x - c(l)| &\ge \frac{1}{2} \operatorname{side}\big(2\sqrt{d}Q(l)\big) \\
&\ge 2|y - c(l)|.
\end{aligned}$$

Now (21.1) follows by summing (21.2) over l and recalling that $\|b\|_{L^1(\mathbb{R}^d)} \le 2\|f\|_{L^1(\mathbb{R}^d)}$ by the Calderón–Zygmund Theorem 11.4. □

Lemma 21.2 (Hörmander condition)**.** *If* $\rho \in C^1(\mathbb{R}^d \setminus \{0\})$ *with*

$$|(\nabla\rho)(x)| \le \frac{(const.)}{|x|^{d+1}}, \qquad x \in \mathbb{R}^d,$$

then

$$\sup_{y,z\in\mathbb{R}^d} \int_{\{x:|x-z|\ge 2|y-z|\}} |\rho(x-y)-\rho(x-z)|\,dx < \infty.$$

Proof. We can take $z = 0$, by a translation. By the Fundamental Theorem,

$$\begin{aligned}\rho(x-y)-\rho(x) &= \int_0^1 \frac{\partial}{\partial s}\rho(x-sy)\,ds \\ &= -\int_0^1 y\cdot(\nabla\rho)(x-sy)\,ds.\end{aligned}$$

Hence

$$\begin{aligned}&\int_{\{x:|x|\ge 2|y|\}} |\rho(x-y)-\rho(x)|\,dx \\ &\le |y| \int_0^1 \int_{|x|\ge 2|y|} |(\nabla\rho)(x-sy)|\,dxds \\ &\le (\text{const.})|y| \int_{|x|\ge 2|y|} \frac{1}{(|x|/2)^{d+1}}\,dx \\ &\qquad\qquad \text{by using the hypothesis, since } |x-sy| \ge |x|-|y| \ge |x|/2, \\ &= (\text{const.})|y| \int_{2|y|}^\infty \frac{1}{r^{d+1}} r^{d-1}\,dr \\ &= (\text{const.})\end{aligned}$$

□

Now we deduce strong (p,p) estimates.

Corollary 21.3. *The Riesz transforms are strong* (p,p) *for* $1 < p < \infty$.

Proof. R_j is strong $(2,2)$ and linear, by definition in Chapter 20, and R_j is weak $(1,1)$ on $L^1 \cap L^2(\mathbb{R}^d)$ (and hence on all simple functions with support of finite measure) by Theorem 21.1. So R_j is strong (p,p) for $1 < p < 2$ by Remark C.4 after Marcinkiewicz Interpolation (in Appendix C). That is, $R_j : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ is bounded and linear for $1 < p < 2$.

For $2 < p < \infty$ we use duality and anti-selfadjointness $R_j^* = -R_j$ on $L^2(\mathbb{R}^d)$ to reduce to the case $1 < p < 2$, just like in the proof of Corollary 12.2. □

Alternatively, for singular integral kernels of the form

$$\frac{\mathcal{O}(x/|x|)}{|x|^d}$$

where $\mathcal{O}$ is an odd function on the unit sphere, one can instead use the *method of rotations* [Grafakos, Section 4.2c]. The idea is to express convolution with this kernel as an average of Hilbert transforms taken in all possible directions in $\mathbb{R}^d$.

The Riesz kernel $c_d(x_j/|x|)/|x|^d$ fits this form, since $\mathcal{O}(y) = y_j$ is odd.

The strong (p, p) bound on the Riesz transform can be generalized to a whole class of convolution-type singular integral operators [Duoandikoetxea, Section 5.1].

Part III

Fourier series and integrals

Chapter 22

Compactly supported Fourier transforms, and the sampling theorem

Goal

Show band limited functions are holomorphic
Prove the Kotelnikov–Shannon–Whittaker sampling theorem

Reference

[Katznelson] Section VI.7

Definition 22.1. We say $f = \check{g}$ is **band limited** if $g \in L^1(\mathbb{R}^d)$ has compact support.

Theorem 22.2 (Band limited functions are holomorphic)**.** *Assume $g \in L^1(\mathbb{R}^d)$ is supported in a ball $B(R)$, and define*

$$f(z) = \check{g}(z) = \frac{1}{(2\pi)^d} \int_{B(R)} g(\xi) e^{i\xi z} \, d\xi$$

for $z = x + iy \in \mathbb{C}^d, x, y \in \mathbb{R}^d$. (Here $\xi z = \xi_1 z_1 + \cdots + \xi_d z_d$.)
Then f is holomorphic, and $|f(z)| = O(e^{R|y|})$.
If in addition $g \in L^2(\mathbb{R}^d)$ then $|f(z)| = O(e^{R|y|}/\sqrt{|y|})$.

Thus once more, decay of the Fourier transform (here, compact support) implies smoothness of the function (here, holomorphicity). The theorem also bounds the rate of growth of the function in the complex directions. (The function must vanish at infinity in the real directions, by the Riemann–Lebesgue corollary, since g is integrable.)

For example, the Dirichlet kernel $D(x) = \sin(x)/\pi x = (\mathbb{1}_{[-1,1]})\check{}(x)$ is band limited with $R = 1$, in 1 dimension. Taking $z = 0 + iy$, we calculate

$$D(iy) = \frac{e^y - e^{-y}}{2\pi y} = O(e^{|y|}/|y|),$$

which is better (by a factor of $\sqrt{|y|}$) than is guaranteed by the theorem.

Proof. f is well defined because $\xi \mapsto e^{i\xi z}$ is bounded on $B(R)$, for each z. And f is holomorphic because $e^{i\xi z}$ is holomorphic and f can be differentiated through the integral with respect to the complex variable z. (Exercise. Justify these claims in detail.)

Clearly

$$\begin{aligned} |f(z)| &\leq \frac{1}{(2\pi)^d} \int_{B(R)} |g(\xi)| e^{-\xi y}\, d\xi \qquad\qquad \text{since } e^{i\xi z} = e^{i\xi x} e^{-\xi y} \\ &\leq \frac{1}{(2\pi)^d} \|g\|_{L^1(\mathbb{R}^d)} e^{R|y|}. \end{aligned}$$

If in addition $g \in L^2(\mathbb{R}^d)$, then

$$|f(z)| \leq \frac{1}{(2\pi)^d} \|g\|_{L^2(\mathbb{R}^d)} \Big(\int_{B(R)} e^{-2\xi y}\, d\xi \Big)^{1/2}$$

and

$$\begin{aligned}\int_{B(R)} e^{-2\xi y}\,d\xi &= \int_{B(R)} e^{-2\xi|y|e_1}\,d\xi \\ &\qquad \text{by } \xi \mapsto \xi A \text{ for some orthogonal matrix } A \text{ with } Ay = |y|e_1 \\ &= \int_{B(R)} e^{-2\xi_1|y|}\,d\xi \\ &\le \int_{[-R,R]^d} e^{-2\xi_1|y|}\,d\xi \\ &= (2R)^{d-1}\frac{e^{2R|y|}-e^{-2R|y|}}{2|y|} \\ &\le \frac{(2R)^{d-1}}{2}\frac{e^{2R|y|}}{|y|}.\end{aligned}$$

Hence $|f(z)| \le (\text{const.})e^{R|y|}/\sqrt{|y|}$. □

Holomorphic functions are known to be determined by their values on lower dimensional sets in $\mathbb{C}^d$. For a band limited function, that "sampling set" can be a lattice in $\mathbb{R}^d$.

Theorem 22.3 (Sampling theorem for band limited functions)**.** *Assume $f \in L^2(\mathbb{R}^d)$ is band limited, with $\widehat{f}$ supported in the cube $[-\omega,\omega]^d$ for some $\omega > 0$.*

Then

$$f(x) = \sum_{n\in\mathbb{Z}^d} f\big(\frac{\pi}{\omega}n\big)\prod_{j=1}^{d}\operatorname{sinc}(\omega x_j - \pi n_j)$$

with the series converging in $L^2(\mathbb{R}^d)$, and also uniformly (in $L^\infty(\mathbb{R}^d)$).

Remark 22.4.

1. The sampling rate ω/π is proportional to the bandwidth ω, that is, to the highest frequency contained in the signal f. Intuitively, the sampling rate must be high when the frequencies are high, because many samples are needed to determine a highly oscillatory function.
2. $\operatorname{sinc}(\omega x_j - \pi n_j)$ is centered at the sampling location $(\pi/\omega)n_j$ and rescaled to have bandwidth ω. It vanishes at all the other sampling locations $(\pi/\omega)m_j$, since

$$\operatorname{sinc}\big(\omega(\pi/\omega)m_j - \pi n_j\big) = \operatorname{sinc}\big(\pi(m_j - n_j)\big) = 0.$$

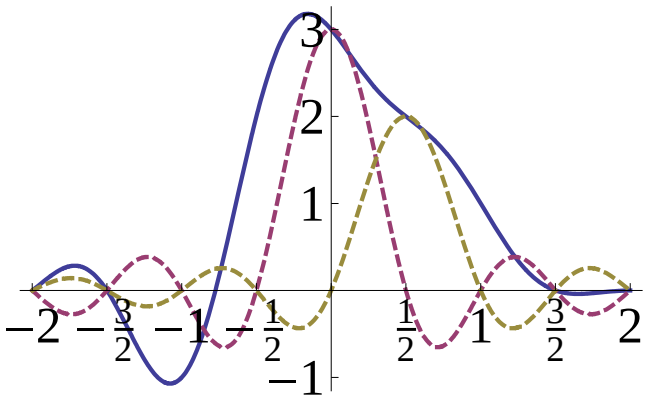

Figure 22.1: Example of sampling formula in Theorem 22.3, with $\omega = 2\pi$ and sampling rate $\omega/\pi = 2$. The dashed curves are two of the sinc functions making up the signal. See Remark 22.4.

3. A graphical example of the sampling formula is shown in Figure 22.1, for $\omega = 2\pi$ and

$$f(x) = -\operatorname{sinc}\big(2\pi(x+1)\big) + 2\operatorname{sinc}\big(2\pi(x+1/2)\big) + 3\operatorname{sinc}\big(2\pi x\big) \\ + 2\operatorname{sinc}\big(2\pi(x-1/2)\big) + 1\operatorname{sinc}\big(2\pi(x-1)\big).$$

The figure shows f with a solid curve, and $3\operatorname{sinc}\big(2\pi x\big)$ and $2\operatorname{sinc}\big(2\pi(x-1/2)\big)$ with dashed curves.

Proof of Sampling Theorem. We can assume $\omega = \pi$, by replacing x with $(\pi/\omega)x$ (Exercise).

Next, $\widehat{f}$ is square integrable and compactly supported, and so is integrable. Hence by L^1 Fourier inversion, f is continuous (after redefining it on some set of measure zero) with

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{i\xi x}\, d\xi, \qquad x \in \mathbb{R}^d. \tag{22.1}$$

Thus the pointwise sampled values $f\big((\pi/\omega)n\big)$ in the theorem are well defined.

We will prove

$$\widehat{f}(\xi) = \sum_{n\in\mathbb{Z}^d} f(-n) e^{i\xi n}, \qquad \xi \in [-\pi,\pi]^d, \tag{22.2}$$

with convergence in $L^2([-\pi,\pi]^d)$. Indeed, if we regard $\widehat{f}$ as a square integrable function on the cube $\mathbb{T}^d = [-\pi,\pi]^d$, then its Fourier coefficients are

$$\begin{aligned}
&\frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \widehat{f}(\xi)e^{-i\xi n}\,d\xi \\
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \widehat{f}(\xi)e^{-i\xi n}\,d\xi \qquad\qquad \text{since } \widehat{f} \text{ is supported in } [-\pi,\pi]^d \\
&= f(-n)
\end{aligned}$$

by the inversion formula (22.1). Thus (22.2) simply expresses the Fourier series of $\widehat{f}$ on the cube.

After changing $n \mapsto -n$ in (22.2), we have

$$\widehat{f}(\xi) = \sum_{n\in\mathbb{Z}^d} f(n)e^{-i\xi n}\mathbb{1}_{[-\pi,\pi]^d}(\xi), \qquad \xi\in\mathbb{R}^d,$$

with convergence in $L^2(\mathbb{R}^d)$ and in $L^1(\mathbb{R}^d)$. Applying L^2 inversion gives

$$\begin{aligned}
f(x) &= \sum_{n\in\mathbb{Z}^d} f(n)\big(e^{-i\xi n}\mathbb{1}_{[-\pi,\pi]^d}\big)^{\vee}(x) \\
&= \sum_{n\in\mathbb{Z}^d} f(n)\prod_{j=1}^{d}\frac{\sin(\pi(x_j-n_j))}{\pi(x_j-n_j)}
\end{aligned}$$

with convergence in $L^2(\mathbb{R}^d)$. Applying L^1 inversion gives convergence in L^∞. □

Paley–Wiener space

For a deeper perspective on Sampling Theorem 22.3, consider the *Paley–Wiener space*

$$PW(\omega) = \{f\in L^2(\mathbb{R}^d) : \widehat{f} \text{ is supported in } [-\omega,\omega]^d\ \}.$$

Clearly $PW(\omega)$ is a subspace of $L^2(\mathbb{R}^d)$, and it is a closed subspace (since if $f = \lim_m f_m$ in $L^2(\mathbb{R}^d)$ and $\widehat{f_m}$ is supported in $[-\omega,\omega]^d$, then $\widehat{f} = \lim_m \widehat{f_m}$ is also supported in $[-\omega,\omega]^d$).

Hence $PW(\omega)$ is a Hilbert space with the L^2 inner product. It is isometric, under the Fourier transform, to $L^2([-\omega,\omega]^d)$ with inner product $(2\pi)^{-d}\langle\cdot,\cdot\rangle_{L^2}$. That space has orthonormal Fourier basis

$$\big\{(\pi/\omega)^{d/2}\mathbb{1}_{[-\omega,\omega]^d}(\xi)e^{-i\xi(\pi/\omega)n}\big\}_{n\in\mathbb{Z}^d},$$

where the indicator function simply reminds us that we are working on the cube. Taking the inverse Fourier transform gives an orthonormal basis of sinc functions for the Paley–Wiener space:

$$\{g_n\}_{n\in\mathbb{Z}^d} = \big\{(\omega/\pi)^{d/2}\prod_{j=1}^{d}\operatorname{sinc}(\omega x_j - \pi n_j)\big\}_{n\in\mathbb{Z}^d}.$$

Using this orthonormal basis, we expand

$$f = \sum_{n\in\mathbb{Z}^d}\langle f, g_n\rangle_{L^2}\, g_n, \qquad \text{for all } f \in PW(\omega), \tag{22.3}$$

where the coefficient is

$$\begin{aligned}\langle f, g_n\rangle_{L^2} &= \frac{1}{(2\pi)^d}\langle \widehat{f}, (\pi/\omega)^{d/2}\mathbb{1}_{[-\omega,\omega]^d}e^{-i\xi(\pi/\omega)n}\rangle_{L^2} \qquad \text{by Parseval}\\ &= (\pi/\omega)^{d/2} f\big((\pi/\omega)n\big)\end{aligned}$$

by Fourier inversion. Thus the orthonormal expansion (22.3) simply restates the Sampling Theorem 22.3.

Our calculations have, of course, essentially repeated the proof of the Sampling Theorem.

Chapter 23

Periodization and Poisson summation

Goal

Periodize functions on $\mathbb{R}^d$ to functions on $\mathbb{T}^d$
Show the Fourier series of periodization gives the Poisson summation formula

References

[Folland] Section 8.3
[Katznelson] Section VI.1

Definition 23.1. Given $f \in L^1(\mathbb{R}^d)$, its **periodization** is the function

$$\mathrm{Pe}(f)(x) = (2\pi)^d \sum_{n\in\mathbb{Z}^d} f(x + 2\pi n), \qquad x \in \mathbb{R}^d.$$

Example 23.2. In 1 dimension, if $f = \mathbb{1}_{[-\pi,2\pi)}$, then $\mathrm{Pe}(f) = 2\pi(2\mathbb{1}_{[-\pi,0)} + \mathbb{1}_{[0,\pi)})$ for $x \in [-\pi, \pi)$, with $\mathrm{Pe}(f)$ extending 2π-periodically to $\mathbb{R}$.

Lemma 23.3. *If $f \in L^1(\mathbb{R}^d)$ then the series for $\mathrm{Pe}(f)(x)$ converges absolutely for almost every x, and $\mathrm{Pe}(f)$ is $2\pi\mathbb{Z}^d$-periodic. Further, $\mathrm{Pe} : L^1(\mathbb{R}^d) \to L^1(\mathbb{T}^d)$ is bounded, with*

$$\|\mathrm{Pe}(f)\|_{L^1(\mathbb{T}^d)} \le \|f\|_{L^1(\mathbb{R}^d)}.$$

The periodization has Fourier coefficients

$$\widehat{\mathrm{Pe}(f)}(j) = \widehat{f}(j), \qquad j \in \mathbb{Z}^d.$$

That is, the jth Fourier coefficient of $\mathrm{Pe}(f)$ equals the Fourier transform of f at j.

Proof. See Problem 19 in Assignment 3. □

Lemma 23.4 (Periodization of a convolution)**.** *If $f, g \in L^1(\mathbb{R}^d)$ then*

$$\mathrm{Pe}(f * g) = \mathrm{Pe}(f) * \mathrm{Pe}(g).$$

Proof. We have

$$\begin{aligned}
\big(\mathrm{Pe}(f*g)\big)\widehat{\ }(j) &= (f*g)\widehat{\ }(j) && \text{by Lemma 23.3} \\
&= \widehat{f}(j)\,\widehat{g}(j) \\
&= \widehat{\mathrm{Pe}(f)}(j)\,\widehat{\mathrm{Pe}(g)}(j) && \text{by Lemma 23.3 again} \\
&= \big(\mathrm{Pe}(f) * \mathrm{Pe}(g)\big)\widehat{\ }(j)
\end{aligned}$$

and so $\mathrm{Pe}(f*g) = \mathrm{Pe}(f)*\mathrm{Pe}(g)$ by the uniqueness theorem for Fourier series.

For a more direct proof, suppose f and g are bounded with compact support, so that the sums in the following argument are all finite rather than infinite. (Thus sums and integrals can be interchanged, below.)

For each $x \in \mathbb{R}^d$,

$$\begin{aligned}
&\mathrm{Pe}(f * g)(x) \\
&= (2\pi)^d \sum_{n\in\mathbb{Z}^d} (f*g)(x+2\pi n) \\
&= (2\pi)^d \sum_{n\in\mathbb{Z}^d} \int_{\mathbb{R}^d} f(x+2\pi n - y) g(y)\, dy \\
&= \int_{\mathbb{R}^d} \mathrm{Pe}(f)(x-y) g(y)\, dy && \text{by definition of } \mathrm{Pe}(f) \\
&= \sum_{m\in\mathbb{Z}^d} \int_{\mathbb{T}^d} \mathrm{Pe}(f)(x-y-2\pi m) g(y+2\pi m)\, dy && \text{since } \mathbb{R}^d = \bigcup_m (\mathbb{T}^d + 2\pi m) \\
&= \sum_{m\in\mathbb{Z}^d} \int_{\mathbb{T}^d} \mathrm{Pe}(f)(x-y) g(y+2\pi m)\, dy && \text{using periodicity of } \mathrm{Pe}(f) \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \mathrm{Pe}(f)(x-y)\, \mathrm{Pe}(g)(y)\, dy \\
&= \big(\mathrm{Pe}(f) * \mathrm{Pe}(g)\big)(x),
\end{aligned}$$

remembering that our definition of convolution on $\mathbb{T}^d$ has a prefactor of $(2\pi)^{-d}$.

Finally, pass to the general case by a limiting argument, using that if $f_m \to f$ in $L^1(\mathbb{R}^d)$ then $\mathrm{Pe}(f_m) \to \mathrm{Pe}(f)$ in $L^1(\mathbb{T}^d)$ by Lemma 23.3. □

Theorem 23.5 (Poisson summation formula)**.** *Suppose $f \in L^1(\mathbb{R}^d)$ is continuous and decays in space and frequency according to:*

$$|f(x)| \le \frac{C}{(1+|x|)^{d+\varepsilon}}, \qquad x \in \mathbb{R}^d, \tag{23.1}$$

$$|\widehat{f}(\xi)| \le \frac{C}{(1+|\xi|)^{d+\varepsilon}}, \qquad \xi \in \mathbb{R}^d, \tag{23.2}$$

for some constants $C, \varepsilon > 0$.

Then the periodization $\mathrm{Pe}(f)$ *equals its Fourier series at every point:*

$$(2\pi)^d \sum_{n\in\mathbb{Z}^d} f(x+2\pi n) = \sum_{j\in\mathbb{Z}^d} \widehat{f}(j) e^{ijx}, \qquad x \in \mathbb{R}^d.$$

In particular, taking $x = 0$ gives

$$(2\pi)^d \sum_{n\in\mathbb{Z}^d} f(2\pi n) = \sum_{j\in\mathbb{Z}^d} \widehat{f}(j).$$

This Poisson summation formula relates a lattice sum of values of the function to a lattice sum of values of its Fourier transform.

Proof. $\mathrm{Pe}(f)$ has Fourier coefficients in $\ell^1(\mathbb{Z}^d)$, since

$$\begin{aligned}
\sum_{j\in\mathbb{Z}^d} \left|\widehat{\mathrm{Pe}(f)}(j)\right| &= \sum_{j\in\mathbb{Z}^d} |\widehat{f}(j)| && \text{by Lemma 23.3} \\
&\le \sum_{j\in\mathbb{Z}^d} \frac{C}{(1+|j|)^{d+\varepsilon}} && \text{by (23.2)} \\
&\le \int_{\mathbb{R}^d} \frac{(\text{const.})}{(1+|\xi|)^{d+\varepsilon}}\, d\xi \\
&< \infty
\end{aligned}$$

by spherical coordinates.

Hence the Fourier series of $\mathrm{Pe}(f)$ converges absolutely and uniformly to a continuous function. That continuous function has the same Fourier coefficients as $\mathrm{Pe}(f)$, and so it equals $\mathrm{Pe}(f)$ a.e. (just like in 1 dimension; see Chapter 4).

To complete the proof we will show $\mathrm{Pe}(f)$ is continuous, for then $\mathrm{Pe}(f)$ equals its Fourier series everywhere (and not just almost everywhere).

Notice that $\mathrm{Pe}(f)(x) = (2\pi)^d \sum_{n\in\mathbb{Z}^d} f(x+2\pi n)$ is a series of continuous functions. The series converges absolutely and uniformly on each ball in $\mathbb{R}^d$ (by using (23.1); exercise), and so $\mathrm{Pe}(f)$ is continuous. □

Example 23.6 (Periodizing the Poisson kernel)**.** The Poisson kernel P_r on $\mathbb{T}$ equals the periodization of the Poisson kernel P_ω on $\mathbb{R}$:

$$\frac{1-r^2}{1-2r\cos x + r^2} = 2\pi \sum_{n\in\mathbb{Z}} \frac{1}{\pi} \frac{\omega^{-1}}{(x+2\pi n)^2 + \omega^{-2}}, \qquad x\in\mathbb{R}, \tag{23.3}$$

provided $r = e^{-1/\omega}$. Hence we obtain a series expansion for the square of the cosecant:

$$\frac{\pi^2}{\sin^2 \pi x} = \sum_{n\in\mathbb{Z}} \frac{1}{(x+n)^2}, \qquad x\in\mathbb{R}\setminus\mathbb{Z}.$$

Proof. First, to partially motivate these results we note $\mathrm{Pe}(P_\omega * f) = \mathrm{Pe}(P_\omega) * \mathrm{Pe}(f)$ by Lemma 23.4, so that it is plausible P_ω periodizes to P_r for some r.

To prove (23.3), observe that P_ω satisfies decay hypotheses (23.1) and (23.2) because

$$\begin{aligned} P_\omega(x) &= \frac{1}{\pi}\frac{\omega^{-1}}{x^2+\omega^{-2}}, && x \in \mathbb{R}, \\ \widehat{P_\omega}(\xi) &= e^{-|\xi|/\omega}, && \xi \in \mathbb{R}, \end{aligned}$$

by (15.12) and Table 16.1. Hence the Poisson Summation Formula says that

$$\begin{aligned} \mathrm{Pe}(P_\omega)(x) &= \sum_{j\in\mathbb{Z}} \widehat{P_\omega}(j) e^{ijx} \\ &= \sum_{j\in\mathbb{Z}} e^{-|j|/\omega} e^{ijx} \\ &= \sum_{j\in\mathbb{Z}} r^{|j|} e^{ijx} && \text{since } r = e^{-1/\omega} \\ &= P_r(x) \end{aligned}$$

by (2.8), which proves (23.3).

Changing x to $2\pi x$ in (23.3) gives

$$\sum_{n\in\mathbb{Z}} \frac{1}{(x+n)^2 + (2\pi\omega)^{-2}} = 2\pi^2\omega \frac{1-r^2}{1-2r\cos(2\pi x)+r^2}.$$

Since

$$r = e^{-1/\omega} = 1 - \frac{1}{\omega} + O\big(\frac{1}{\omega^2}\big),$$

letting $\omega \to \infty$ implies that

$$\sum_{n\in\mathbb{Z}} \frac{1}{(x+n)^2} = \frac{4\pi^2}{2-2\cos(2\pi x)} = \frac{\pi^2}{(\sin \pi x)^2},$$

where we used monotone convergence on the left side.

Example 23.7 (Periodizing the Gauss kernel)**.** The Gauss kernel $G_s(t) = \sum_{j\in\mathbb{Z}} e^{-j^2 s} e^{ijt}$ on $\mathbb{T}$ equals the periodization of the Gauss kernel G_ω on $\mathbb{R}$:

$$\sum_{j\in\mathbb{Z}} e^{-j^2 s} e^{ijx} = 2\pi \sum_{n\in\mathbb{Z}} \frac{\omega}{\sqrt{2\pi}} e^{-\omega^2(x+2\pi n)^2/2}, \qquad x \in \mathbb{R}, \tag{23.4}$$

provided $s > 0$ and $\omega = 1/\sqrt{2s}$. Hence

$$\sum_{n\in\mathbb{Z}} e^{-n^2\pi s} = s^{-1/2} \sum_{n\in\mathbb{Z}} e^{-n^2\pi/s}, \qquad s > 0.$$

In terms of the **theta function** $\vartheta(s) = \sum_{n\in\mathbb{Z}} e^{-n^2\pi s}$, the last formula expresses the functional equation

$$\vartheta(s) = s^{-1/2}\vartheta(s^{-1}).$$

Proof. Decay hypotheses (23.1) and (23.2) hold for G_ω because

$$\begin{aligned} G_\omega(x) &= \frac{\omega}{\sqrt{2\pi}} e^{-(\omega x)^2/2}, && x \in \mathbb{R}, \\ \widehat{G_\omega}(\xi) &= e^{-(\xi/\omega)^2/2}, && \xi \in \mathbb{R}, \end{aligned}$$

by (15.16) and Table 16.1. Hence the Poisson Summation Formula says that

$$\begin{aligned} \mathrm{Pe}(G_\omega)(x) &= \sum_{j\in\mathbb{Z}} \widehat{G_\omega}(j) e^{ijx} \\ &= \sum_{j\in\mathbb{Z}} e^{-(j/\omega)^2/2} e^{ijx} \\ &= \sum_{j\in\mathbb{Z}} e^{-j^2 s} e^{ijx} && \text{since } \omega = 1/\sqrt{2s} \\ &= G_s(x), \end{aligned}$$

which proves (23.4).

Taking $x = 0$ in (23.4) and changing s to πs yields the functional equation for the theta function.

Chapter 24

Uncertainty principles

Goal

Establish qualitative and quantitative uncertainty principles

References

[Goh and Micchelli] Section 2
[Jaming] Section 1

Uncertainty principles say that f and $\widehat{f}$ cannot both be too localized. Consequently, if $\widehat{f}$ is well localized then f is not, and so we are "uncertain" of the value of f.

Proposition 24.1 (Qualitative uncertainty principles)**.**
(a) If $f \in L^2(\mathbb{T})$ is continuous, f has infinitely many zeros in $\mathbb{T}$, and $\widehat{f}$ is finitely supported, then $f \equiv 0$.
(b) If $f \in L^2(\mathbb{R}^d)$ is continuous, f vanishes on some open set, and $\widehat{f}$ is compactly supported, then $f \equiv 0$.

Proof.
(a) f is a trigonometric polynomial since it has only finitely many nonzero Fourier coefficients. Thus part (a) says:

> a trigonometric polynomial that vanishes infinitely often in $\mathbb{T}$ must vanish identically.

To prove this claim, write $f(t) = \sum_{n=-N}^{N} a_n e^{int}$. Then $f(t) = p(e^{it})/e^{iNt}$ where p is the polynomial

$$p(z) = \sum_{n=0}^{2N} a_{n-N} z^n, \qquad z \in \mathbb{C}.$$

Since f has infinitely many zeros $t \in \mathbb{T}$, we see p has infinitely many zeros e^{it} on the unit circle. The Fundamental Theorem of Algebra implies $p \equiv 0$.

(b) f is band limited, and hence is holomorphic on $\mathbb{C}^d$ by Theorem 22.2. In particular, f is real analytic on $\mathbb{R}^d$.

Choose $x_0 \in \mathbb{R}^d$ such that $f \equiv 0$ on a neighborhood of x_0; then the Taylor series of f centered at x_0 is identically zero. That Taylor series equals f on $\mathbb{R}^d$, and so $f \equiv 0$. □

Theorem 24.2 (Benedicks' qualitative uncertainty principle)**.** *If $f \in L^2(\mathbb{R}^d)$ is continuous and f and $\widehat{f}$ are supported on sets of finite measure, then $f \equiv 0$.*

In contrast to Proposition 24.1, here the support of $\widehat{f}$ need not be compact.

Proof. We prove only the 1 dimensional case.

Let $A = \{x \in \mathbb{R} : f(x) \neq 0\}$ and $B = \{\xi \in \mathbb{R} : \widehat{f}(\xi) \neq 0\}$. By dilating f we can suppose $|A| < 2\pi$. Then

$$\begin{aligned}
&\big|\{x \in \mathbb{T} : f(x + 2\pi n) \neq 0 \text{ for some } n \in \mathbb{Z}\}\big| \\
&= \big|\{x \in \mathbb{T} : \sum_{n\in\mathbb{Z}} \mathbb{1}_A(x + 2\pi n) \geq 1\}\big| \\
&\leq \int_{\mathbb{T}} \sum_{n\in\mathbb{Z}} \mathbb{1}_A(x + 2\pi n)\, dx \\
&= \int_{\mathbb{R}} \mathbb{1}_A(x)\, dx \\
&= |A| \\
&< |\mathbb{T}| = 2\pi.
\end{aligned}$$

Therefore the complementary set

$$E = \{x \in \mathbb{T} : f(x + 2\pi n) = 0 \text{ for all } n \in \mathbb{Z}\}$$

has positive measure.

Next,

$$\begin{aligned}\int_{[0,1)} \sum_{j\in\mathbb{Z}} \mathbb{1}_B(\xi+j)\,d\xi &= \int_{\mathbb{R}} \mathbb{1}_B(\xi)\,d\xi \\ &= |B| \\ &< \infty,\end{aligned}$$

so that $\sum_{j\in\mathbb{Z}} \mathbb{1}_B(\xi+j)$ is finite for almost every $\xi \in [0,1)$, say for all $\xi \in F \subset [0,1)$ where F has full measure, $|[0,1)\setminus F| = 0$. Hence when $\xi \in F$, the set $\{j \in \mathbb{Z} : \widehat{f}(\xi+j) \neq 0\}$ is finite.

Fix $\xi \in F$ and consider the periodization

$$\mathrm{Pe}(fe^{-i\xi x})(x) = 2\pi \sum_{n\in\mathbb{Z}} f(x+2\pi n)e^{-i\xi(x+2\pi n)},$$

which is well defined since $f \in L^1(\mathbb{R})$. The jth Fourier coefficient of the periodization equals

$$(fe^{-i\xi x})\widehat{\ }(j) = \widehat{f}(\xi+j),$$

which equals zero for but finitely many j, since $\xi \in F$. Thus $\mathrm{Pe}(fe^{-i\xi x})$ equals some trigonometric polynomial $Q(x)$ a.e. But $\mathrm{Pe}(fe^{-i\xi x})(x) = 0$ for all $x \in E$, and so Q vanishes a.e. on E. In particular, Q vanishes at infinitely many points in $\mathbb{T}$ (using here that E has positive measure). Hence $Q \equiv 0$ by Proposition 24.1(a). The Fourier coefficient $\widehat{f}(\xi+j)$ of Q therefore vanishes for all j.

Since $\widehat{f}(\xi+j) = 0$ for all $j \in \mathbb{Z}$ and almost every $\xi \in [0,1)$, we deduce $\widehat{f}(\xi) = 0$ a.e., and so $f \equiv 0$. □

Theorem 24.3 (Nazarov's quantitative uncertainty principle)**.** *A constant $C_d > 0$ exists such that*

$$\|f\|^2_{L^2(\mathbb{R}^d)} = \|\widehat{f}\|^2_{L^2(\mathbb{R}^d)} \le C_d^{|A||B|+1}\Big(\int_{\mathbb{R}^d\setminus A} |f(x)|^2\,dx + \int_{\mathbb{R}^d\setminus B} |\widehat{f}(\xi)|^2\,d\xi\Big)$$

for all sets $A, B \subset \mathbb{R}^d$ of finite measure and all $f \in L^2(\mathbb{R}^d)$.

We omit the proof.

Nazarov's theorem implies Benedicks' theorem, because if f is supported in A and $\widehat{f}$ is supported in B, then the right side is zero and so $f \equiv 0$.

Next we develop an abstract commutator inequality that leads to the Heisenberg Uncertainty Principle.

Let H be a Hilbert space. Suppose T is a linear operator from a subspace $\mathcal{D}(T)$ into H. Write T^* for its adjoint, defined on a subspace $\mathcal{D}(T^*)$, meaning T^* is linear and

$$\langle Tf, g\rangle = \langle f, T^*g\rangle \qquad \text{whenever} \quad f \in \mathcal{D}(T), \quad g \in \mathcal{D}(T^*).$$

Define

$$\begin{aligned}\Delta_f(T) &= \min_{\alpha\in\mathbb{C}} \|Tf - \alpha f\| \\ &= \text{norm of component of } Tf \text{ perpendicular to } f.\end{aligned}$$

The minimum is attained for $\alpha = \langle Tf, f\rangle/\|f\|^2$.

Theorem 24.4 (Commutator estimate)**.** *Let T and U be linear operators like above. Then*

$$\big|\langle [T,U]f, f\rangle\big| \le \Delta_f(T^*)\Delta_f(U) + \Delta_f(T)\Delta_f(U^*)$$

for all $f \in \mathcal{D}(TU) \cap \mathcal{D}(UT) \cap \mathcal{D}(T^*) \cap \mathcal{D}(U^*)$.

Here $[T,U] = TU - UT$ is the *commutator* of T and U.

Proof.

$$\begin{aligned}\big|\langle [T,U]f, f\rangle\big| &= \big|\langle TUf, f\rangle - \langle UTf, f\rangle\big| \\ &= \big|\langle Uf, T^*f\rangle - \langle Tf, U^*f\rangle\big| \\ &\le \|Uf\|\|T^*f\| + \|Tf\|\|U^*f\|. \qquad (24.1)\end{aligned}$$

Let $\alpha, \beta \in \mathbb{C}$. Note that

$$[T - \alpha I, U - \beta I] = [T, U].$$

Hence by replacing T with $T - \alpha I$ and U with $U - \beta I$ in (24.1) we find

$$\big|\langle [T,U]f, f\rangle\big| \le \|Uf - \beta f\|\|T^*f - \overline{\alpha}f\| + \|Tf - \alpha f\|\|U^*f - \overline{\beta}f\|.$$

Minimizing over α and β proves the theorem, noting for the adjoints that

$$\alpha = \frac{\langle Tf, f\rangle}{\|f\|^2} \quad \Longleftrightarrow \quad \overline{\alpha} = \frac{\langle T^*f, f\rangle}{\|f\|^2}.$$

□

Example 24.5 (Heisenberg Uncertainty Principle)**.** Take $H = L^2(\mathbb{R})$,

$$\begin{aligned}(Tf)(x) &= xf(x) \qquad \text{with} \quad \mathcal{D}(T) = \{f \in L^2(\mathbb{R}) : xf(x) \in L^2(\mathbb{R})\},\\ (Uf)(x) &= -if'(x) \qquad \text{with} \quad \mathcal{D}(U) = \{f \in L^2(\mathbb{R}) : f' \in L^2(\mathbb{R})\}.\end{aligned}$$

Here T is the *position* operator and U is the *momentum* operator.

Observe $T^* = T, U^* = U$ and

$$\begin{aligned}[T,U]f &= TUf - UTf\\ &= x \cdot \big(-i\frac{d}{dx}f(x)\big) + i\frac{d}{dx}\big(xf(x)\big)\\ &= if(x).\end{aligned}$$

The Commutator Theorem 24.4 implies

$$\begin{aligned}\|f\|^2_{L^2(\mathbb{R})} &\le 2\Delta_f(T)\Delta_f(U)\\ &\le 2\|xf - \alpha f\|_{L^2(\mathbb{R})}\,\|-if' - \beta f\|_{L^2(\mathbb{R})}\\ &= 2\|(x-\alpha)f\|_{L^2(\mathbb{R})}\,\frac{1}{\sqrt{2\pi}}\|(\xi-\beta)\widehat{f}\|_{L^2(\mathbb{R})}\end{aligned}$$

by Plancherel. Squaring yields the **Heisenberg Uncertainty Principle**:

$$\frac{1}{4}\|f\|^4_{L^2(\mathbb{R})} \le \int_{\mathbb{R}} |x-\alpha|^2|f(x)|^2\,dx \cdot \frac{1}{2\pi}\int_{\mathbb{R}} |\xi-\beta|^2|\widehat{f}(\xi)|^2\,d\xi \tag{24.2}$$

for all $\alpha, \beta \in \mathbb{C}$.

We interpret (24.2) as restricting how localized f and $\widehat{f}$ can be, around the locations α and β.

In quantum mechanics, we normalize $\|f\|_{L^2(\mathbb{R})} = 1$ and interpret $|f(x)|^2$ as the probability density for the position x of some particle, and regard $|\widehat{f}|^2/2\pi$ as the probability density for the momentum ξ. Thus the Heisenberg Uncertainty Principle implies that the variance (or uncertainty) in position multiplied by the variance in momentum is at least $1/4$.

Roughly, the Principle says that the more precisely one knows the position of a quantum particle, the less precisely one knows its momentum, and vice versa.

Remark 24.6.

1. Equality holds in the Heisenberg Principle (24.2) if and only if $f(x) = Ce^{i\beta x}e^{-\gamma(x-\alpha)^2}$ is a β-modulated Gaussian at α (with $C \in \mathbb{C}, \gamma > 0$).

2. A more direct proof of (24.2) can be given by integrating by parts in

$$\|f\|_{L^2(\mathbb{R})}^2 = \int_{\mathbb{R}} f(x)\overline{f(x)}(x-\alpha)'\,dx$$

and then applying Cauchy–Schwarz.

3. The Heisenberg Uncertainty Principle extends naturally to higher dimensions.

4. On $\mathbb{T}$, the analogous uncertainty principle says

$$\frac{1}{4}m^2\Big|\frac{1}{2\pi}\int_{\mathbb{T}} e^{imt}|f(t)|^2\,dt\Big|^2 \le \frac{1}{2\pi}\int_{\mathbb{T}} |e^{imt}-\alpha|^2|f(t)|^2\,dt \cdot \sum_{n\in\mathbb{Z}} |n-\beta|^2|\widehat{f}(n)|^2$$

for all $\alpha, \beta \in \mathbb{C}, m \in \mathbb{Z}$ (exercise).

One considers here a quantum particle at position e^{it} on the unit circle, with momentum $n \in \mathbb{Z}$. When $\alpha = 0$ we deduce a lower bound on the localization of momentum, in terms of Fourier coefficients of the position density $|f|^2$:

$$\frac{1}{4}\|f\|_{L^2(\mathbb{T})}^{-2} \sup_{m\in\mathbb{Z}} m^2\big|\widehat{(|f|^2)}(m)\big|^2 \le \sum_{n\in\mathbb{Z}} |n-\beta|^2|\widehat{f}(n)|^2.$$

Part IV

Problems

Assignment 1

Problem 1. Do the following problems, but do not hand them in:
[Katznelson] Ex. 1.1.2, 1.1.4.

Problem 2. ([Katznelson] Ex. 1.1.5: downsampling)
Let $f \in L^1(\mathbb{T}), m \in \mathbb{N}$, and define

$$f_{(m)}(t) = f(mt).$$

(a) Prove that $\widehat{f_{(m)}}(n) = \widehat{f}(n/m)$ if $m|n$ and $\widehat{f_{(m)}}(n) = 0$ otherwise. Use only the definition of the Fourier coefficients, and elementary manipulations.
(b) Then give a quick, formal (nonrigorous) proof using the Fourier series of f.

Problem 3. ([Katznelson] Ex. 1.2.8: Fejér's Lemma)
Let $f \in L^p(\mathbb{T})$ and $g \in L^q(\mathbb{T})$, where $1 < p \leq \infty, 1 \leq q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Prove that

$$\lim_{m\to\infty} \frac{1}{2\pi} \int_{\mathbb{T}} f(mt)\overline{g(t)}\, dt = \widehat{f}(0)\overline{\widehat{g}(0)}.$$

Hint. Use that trigonometric polynomials are dense in $L^q(\mathbb{T})$.

Problem 4. (Weak convergence and oscillation)
Let H be a Hilbert space. We say u_n converges weakly to u, written $u_n \rightharpoonup u$ weakly, if $\langle u_n, v\rangle \to \langle u, v\rangle$ as $n \to \infty$, for each $v \in H$. Clearly if $u_n \to u$ in norm (meaning $\|u_n - u\| \to 0$) then $u_n \rightharpoonup u$ weakly.
(a) Show that $e^{imt} \rightharpoonup 0$ weakly in $L^2(\mathbb{T})$, as $m \to \infty$.
(b) Let $f \in L^2(\mathbb{T})$. Show

$$f_{(m)} \rightharpoonup \widehat{f}(0) = (\text{mean value of } f)$$

weakly in $L^2(\mathbb{T})$, as $m \to \infty$.
Remark. Thus rapid oscillation yields weak convergence to the mean.

Problem 5. (Smoothness of f implies rate of decay of $\widehat{f}$)

(a) Show that if f has bounded variation, then $\widehat{f}(n) = O(|n|^{-1})$.

(b) Show that if f is absolutely continuous and f' has bounded variation, then $\widehat{f}(n) = O(|n|^{-2})$.

Remark. These results cover most of the functions encountered in elementary courses. For example, functions that are smooth expect for finitely many jumps (such as the sawtooth $f(t) = t, t \in (-\pi, \pi]$) have bounded variation. And functions that are smooth except for finitely many corners (such as the triangular wave $f(t) = |t|, t \in (-\pi, \pi]$) have first derivative with bounded variation. That is why one encounters so many functions with Fourier coefficients decaying like $1/n$ or $1/n^2$.

Problem 6. ([Katznelson] Ex. 1.3.2: rate of uniform summability)

Assume f is Hölder continuous, with $f \in C^\alpha(\mathbb{T})$ for some $0 < \alpha < 1$. Prove there exists $C > 0$ (depending on the Hölder constant of f) such that

$$\|\sigma_N(f) - f\|_{L^\infty} \leq \frac{C}{1-\alpha} \frac{1}{N^\alpha}, \qquad N \in \mathbb{N}.$$

Remark. Thus the "smoother" f is, the faster $\sigma_N(f)$ converges to f as $N \to \infty$.

Problem 7. ([Katznelson] Ex. 1.5.4)

Let f be absolutely continuous on $\mathbb{T}$ with $f' \in L^2(\mathbb{T})$. In other words, $f \in W^{1,2}(\mathbb{T})$.

(a) Prove that

$$\|\widehat{f}\|_{\ell^1(\mathbb{Z})} \leq \|f\|_{L^1(\mathbb{T})} + \left(2 \sum_{n=1}^{\infty} \frac{1}{n^2}\right)^{1/2} \|f'\|_{L^2}.$$

Hint. First evaluate $\|f'\|_{L^2}^2$.

(b) Deduce that $f \in A(\mathbb{T})$.

Remark. Hence the Fourier series of f converges uniformly by Chapter 4, so that $S_n(f) \to f$ in $L^\infty(\mathbb{T})$. In particular, if f is smooth except for finite many corners (such as the triangular wave $f(t) = |t|$ for $t \in (-\pi, \pi]$), then the Fourier series converges uniformly to f.

Problem 8. (A lacunary series)

Assume $0 < \alpha < 1$.

(a) Suppose that f is continuous on $\mathbb{T}$ and that

$$\sum_{2^n \le |j| < 2^{n+1}} |\widehat{f}(j)| \le C 2^{-n\alpha}$$

for each $n \ge 0$. Prove $f \in A(\mathbb{T})$, and then $f \in C^\alpha(\mathbb{T})$.

(b) Let $f(t) = \sum_{n=0}^{\infty} 2^{-n\alpha} e^{i2^n t}$. Show $f \in C^\alpha(\mathbb{T})$. Deduce that the rate of decay $\widehat{f}(n) = O(|n|^{-\alpha})$ proved in Theorem 1.6 for $C^\alpha(\mathbb{T})$ is sharp. (That is, show $\widehat{f}(n) = O(|n|^{-\beta})$ fails for some $f \in C^\alpha(\mathbb{T})$, when $\beta > \alpha$.)

Problem 9. (Maximal function when $p = 1$)

Define $L \log L(\mathbb{R}^d)$ to be the class of measurable functions for which $\int_{\mathbb{R}^d} |f(x)| \log\big(1 + |f(x)|\big)\, dx < \infty$. Prove that

$$f \in L \log L(\mathbb{R}^d) \quad \Longrightarrow \quad Mf \in L^1_{loc}(\mathbb{R}^d).$$

Remark. Thus if the singularities of f are "logarithmically better than L^1" then the Hardy–Littlewood maximal function belongs to L^1 (at least locally).

Problem 10. Enjoyable reading (nothing to hand in).

Read Chapter 8 "Compass and Tides" from [Körner], which shows how sums of Fourier series having different underlying periods can be used to model the heights of tides.

Sums of periodic functions having different periods are called *almost periodic functions*. Their theory was developed by the Danish mathematician Harald Bohr, brother of physicist Niels Bohr. Harald Bohr won a silver medal at the 1908 Olympics, in soccer.

Assignment 2

Problem 11 (Hilbert transform of indicator function)**.**

(a) Evaluate $(H\mathbb{1}_{[a,b]})(t)$, where $[a,b] \subset (-\pi,\pi)$ is a closed interval. Sketch the graph, for $t \in [-\pi,\pi]$.

(b) Conclude that the Hilbert transform on $\mathbb{T}$ is not strong (∞,∞).

Problem 12 (Fourier synthesis on ℓ^p)**.** Let $1 \le p \le 2$.

Prove that the Fourier synthesis operator T, defined by

$$(T\{c_n\})(t) = \sum_{n\in\mathbb{Z}} c_n e^{int},$$

is bounded from $\ell^p(\mathbb{Z})$ to $L^{p'}(\mathbb{T})$. Estimate the norm of T.

Extra credit. Show the series converges unconditionally, in $L^{p'}(\mathbb{T})$.

Problem 13 (Parseval on L^p)**.** Do part (a) or part (b). You may do both parts if you wish.

(a) Let $1 \le p \le 2$. Take $f \in L^p(\mathbb{T})$ and $g \in L^1(\mathbb{T})$ with $\{\widehat{g}(n)\} \in \ell^p(\mathbb{Z})$. Prove that $g \in L^{p'}(\mathbb{T})$, and establish the Parseval identity

$$\frac{1}{2\pi}\int_{\mathbb{T}} f(t)\overline{g(t)}\,dt = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\overline{\widehat{g}(n)}.$$

(In your solution, explain why the integral and sum are absolutely convergent.)

(b) Let $1 < p < \infty$. Take $f \in L^p(\mathbb{T})$ and $g \in L^{p'}(\mathbb{T})$. Prove the Parseval identity

$$\frac{1}{2\pi}\int_{\mathbb{T}} f(t)\overline{g(t)}\,dt = \lim_{N\to\infty} \sum_{|n|\le N} \widehat{f}(n)\overline{\widehat{g}(n)}.$$

Problem 14 (Fourier analysis into a weighted space)**.** Let $1 < p \le 2$.

(a) Show

$$\left(\sum_{n \neq 0} |\widehat{f}(n)|^p |n|^{p-2} \right)^{1/p} \le C_p \|f\|_{L^p(\mathbb{T})} \qquad \text{for all } f \in L^p(\mathbb{T}).$$

Hint. $Y = \mathbb{Z} \setminus \{0\}$ with $\nu =$ counting measure weighted by n^{-2}.

(b) Show that combining the Hölder and Hausdorff–Young inequalities in the obvious way does *not* prove part (a).

Problem 15 (Poisson extension)**.** Recall P_r denotes the Poisson kernel on $\mathbb{T}$, and write $\mathbb{D}$ for the open unit disk in the complex plane. Suppose $f \in C(\mathbb{T})$ and define

$$v(re^{it}) = \begin{cases} (P_r * f)(t) & \text{for } 0 \le r < 1,\ t \in \mathbb{T}, \\ f(t) & \text{for } r = 1,\ t \in \mathbb{T}, \end{cases}$$

so that v is defined on the closed disk $\overline{\mathbb{D}}$.

(a) Show v is C^∞ smooth and harmonic ($\Delta v = 0$) in $\mathbb{D}$.

(b) Show v is continuous on $\overline{\mathbb{D}}$.

(c) [Optional; no credit] Assume $f \in C^\infty(\mathbb{T})$ and show $v \in C^\infty(\overline{\mathbb{D}})$. (Parts (a) and (b) show v is smooth on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$. Thus the task is to prove each partial derivative of v on $\mathbb{D}$ extends continuously to $\overline{\mathbb{D}}$.)

Aside. $(P_r * f)(t)$ is called the *harmonic extension* to the disk of the boundary function f.

Problem 16 (Boundary values lose half a derivative)**.** Assume u is a smooth, real-valued function on a neighborhood of $\overline{\mathbb{D}}$, and define

$$f(t) = u(e^{it})$$

for the boundary value function of u. Hence $f \in C^\infty(\mathbb{T})$, and so the Poisson extension v belongs to $C^\infty(\overline{\mathbb{D}})$ by Problem 15(c).

(a) Prove

$$\frac{1}{2\pi} \int_{\mathbb{D}} |\nabla v|^2 \, dA = \sum_{n \in \mathbb{Z}} |n| |\widehat{f}(n)|^2.$$

Hint. Use one of Green's formulas, and remember $v = \overline{v}$ since f and v are real-valued.

(b) Prove

$$\int_{\mathbb{D}} |\nabla v|^2 \, dA \leq \int_{\mathbb{D}} |\nabla u|^2 \, dA.$$

Hint. Write $u = v + (u - v)$ and use one of Green's formulas.

Aside. This result is known as "Dirichlet's principle". It asserts that among all functions having the same boundary values, the harmonic function has smallest Dirichlet integral. As your proof reveals, this result holds on arbitrary domains.

(c) Conclude

$$\sum_{n \in \mathbb{Z}} |n| |\widehat{f}(n)|^2 \leq \frac{1}{2\pi} \int_{\mathbb{D}} |\nabla u|^2 \, dA.$$

Discussion. We say f has "half a derivative" in L^2, since $\{|n|^{1/2} \widehat{f}(n)\} \in \ell^2(\mathbb{Z})$. Justification: if f has zero derivatives ($f \in L^2(\mathbb{T})$) then $\{\widehat{f}(n)\} \in \ell^2(\mathbb{Z})$, and if f has one derivative ($f' \in L^2(\mathbb{T})$) then $\{n\widehat{f}(n)\} \in \ell^2(\mathbb{Z})$. Halfway inbetween lies the condition $\{|n|^{1/2} \widehat{f}(n)\} \in \ell^2(\mathbb{Z})$.

Boundary trace inequalities like in part (c) are important for partial differential equations and Sobolev space theory. The inequality says, basically, that if a function u has one derivative ∇u belonging to L^2 on a domain, then u has half a derivative in L^2 on the boundary. Thus the boundary value loses half a derivative, compared to the original function.

Note that in this problem, $f \in C^\infty(\mathbb{T})$ and so certainly $f' \in L^2(\mathbb{T})$, which implies $\{n\widehat{f}(n)\} \in \ell^2(\mathbb{Z})$. You might wonder, then, why you should bother proving the weaker result $\{|n|^{1/2} \widehat{f}(n)\} \in \ell^2(\mathbb{Z})$ in part (c). But actually you prove more in part (c): you obtain a *norm estimate* on $\{|n|^{1/2} \widehat{f}(n)\} \in \ell^2(\mathbb{Z})$ in terms of the L^2 norm of ∇u. (We do not have such a norm estimate on $\{n\widehat{f}(n)\}$.) This norm estimate means that the restriction map

$$\begin{aligned} H^1(\mathbb{D}) &\to H^{1/2}(\partial\mathbb{D}) \\ u &\mapsto f \end{aligned}$$

is bounded from the Sobolev space $H^1(\mathbb{D})$ on the disk with one derivative in L^2 to the Sobolev space $H^{1/2}(\partial\mathbb{D})$ on the boundary circle with half a derivative in L^2.

Aside. The notion of fractional derivatives defined via Fourier coefficients can be extended to fractional derivatives in $\mathbb{R}^d$, by using Fourier transforms.

Problem 17 (Measuring diameters of stars)**.**

Enjoyable reading; nothing to hand in.

Read Chapter 95 "The Diameter of Stars" from [Körner], which shows how the diameters of stars can be estimated using Fourier transforms of radial functions, and convolutions.

Assignment 3

Problem 18 (Adjoint of Fourier transform)**.**
Find the adjoint of the Fourier transform on $L^2(\mathbb{R}^d)$.

Problem 19 (Periodization, and Fourier coefficients and transforms)**.**
Suppose $f \in L^1(\mathbb{R}^d)$.

(a) Prove that the periodization

$$\mathrm{Pe}(f)(x) = (2\pi)^d \sum_{n\in\mathbb{Z}^d} f(x + 2\pi n)$$

of f satisfies

$$\|\mathrm{Pe}(f)\|_{L^1(\mathbb{T}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}.$$

(b) Deduce from your argument that the series for $\mathrm{Pe}(f)(x)$ converges absolutely for almost every x, and that $\mathrm{Pe}(f)$ is $2\pi\mathbb{Z}^d$-periodic.

(c) Show that the jth Fourier coefficient of $\mathrm{Pe}(f)$ equals the Fourier transform of f at j:

$$\widehat{\mathrm{Pe}(f)}(j) = \widehat{f}(j), \qquad j \in \mathbb{Z}^d$$

Problem 20 (Course summary)**.**
Write a one page description of the most important and memorable results and general techniques from this course. Be brief, but thoughtful; explain how these main results fit together.

You need not state the results technically — intuition is more helpful than rigor, at this stage.

Part V

Appendices

Appendix A

Minkowski's integral inequality

Goal

State Minkowski's integral inequality, and apply it to norms of convolutions

Minkowski's inequality on a measure space (X, μ) is simply the triangle inequality for $L^p(X)$, saying that the norm of a sum is bounded by the sum of the norms:

$$\Big\| \sum_j f_j \nu_j \Big\|_{L^p(X)} \leq \sum_j \|f_j\|_{L^p(X)} \nu_j$$

whenever $f_j \in L^p(X)$ and the constants ν_j are nonnegative. Similarly, the norm of an integral is bounded by the integral of the norms:

Theorem A.1. *Suppose (X, μ) and (Y, ν) are σ-finite measure spaces, and that $f(x, y)$ is measurable on the product space $X \times Y$. If $1 \leq p \leq \infty$ then*

$$\Big\| \int_Y f(x, y)\, d\nu(y) \Big\|_{L^p(X)} \leq \int_Y \|f(x, y)\|_{L^p(X)}\, d\nu(y)$$

whenever the right side is finite.

Proof. Take q to be the conjugate exponent, with $\frac{1}{p} + \frac{1}{q} = 1$. Then for all

$g \in L^q(X)$,

$$\begin{aligned}
&\Big| \int_X \Big(\int_Y f(x,y)\, d\nu(y) \Big) g(x)\, d\mu(x) \Big| \\
&\le \int_Y \int_X |f(x,y)||g(x)|\, d\mu(x) d\nu(y) \\
&\le \int_Y \Big(\int_X |f(x,y)|^p\, d\mu(x) \Big)^{1/p} \|g\|_{L^q(X)}\, d\nu(y) \qquad \text{by Hölder} \\
&= \int_Y \|f(x,y)\|_{L^p(X)}\, d\nu(y) \cdot \|g\|_{L^q(X)}.
\end{aligned}$$

Now the theorem follows from the dual characterization of the norm on $L^p(X)$ (see [Folland, Theorem 6.14]). □

Definition A.2. Define the *convolution* of functions f and g on $\mathbb{T}$ by

$$(f * g)(t) = \frac{1}{2\pi} \int_{\mathbb{T}} f(t-\tau) g(\tau)\, d\tau, \qquad t \in \mathbb{T},$$

whenever the integral makes sense.

Define the *convolution* of functions f and g on $\mathbb{R}^d$ by

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

whenever the integral makes sense.

Theorem A.3 (Young's theorem)**.** *Fix $1 \le p \le \infty$. Then*

$$\begin{aligned}
\|f * g\|_{L^p(\mathbb{T})} &\le \|f\|_{L^p(\mathbb{T})} \|g\|_{L^1(\mathbb{T})}, \\
\|f * g\|_{L^p(\mathbb{R}^d)} &\le \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)},
\end{aligned}$$

whenever the right sides are finite. In particular, the convolution $f * g$ is well defined a.e. whenever $f \in L^p$ and $g \in L^1$.

Proof.

$$\begin{aligned}
\|f * g\|_{L^p(\mathbb{R}^d)} &= \Big\| \int_{\mathbb{R}^d} f(\cdot - y) g(y)\, dy \Big\|_{L^p(\mathbb{R}^d)} \\
&\le \int_{\mathbb{R}^d} \|f(\cdot - y)\|_{L^p(\mathbb{R}^d)} |g(y)|\, dy \\
&\qquad\qquad \text{by Minkowski's integral inequality, Theorem A.1,} \\
&= \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}.
\end{aligned}$$

Argue similarly on $\mathbb{T}$. □

Appendix B

L^p norms and the distribution function

Goal

Express L^p-norms in terms of the distribution function

Given a σ-finite measure space (X, μ) and a measurable function f on X, write

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

for the level set of f above level λ. The *distribution function* of f is $\mu(E(\lambda))$.

Lemma B.1. *Let $\alpha > 0$.*

If $-\infty < r < p < \infty$ then

$$\int_0^\infty \lambda^{p-r-1} \int_{E(\lambda/\alpha)} |f(x)|^r \, d\mu(x) d\lambda = \frac{\alpha^{p-r}}{p-r} \int_X |f(x)|^p \, d\mu(x). \tag{B.1}$$

If $-\infty < p < r < \infty$ then

$$\int_0^\infty \lambda^{p-r-1} \int_{E(\lambda/\alpha)^c} |f(x)|^r \, d\mu(x) d\lambda = \frac{\alpha^{p-r}}{r-p} \int_X |f(x)|^p \, d\mu(x). \tag{B.2}$$

In particular, when $r = 0 < p < \infty$ and $\alpha = 1$, formula (B.1) *expresses the L^p-norm in terms of the distribution function:*

$$\int_0^\infty p\lambda^{p-1} \mu(E(\lambda)) \, d\lambda = \int_X |f(x)|^p \, d\mu(x) = \|f\|^p_{L^p(X)}. \tag{B.3}$$

Proof. We can assume $\alpha = 1$ without loss of generality, by changing variable with $\lambda \mapsto \alpha\lambda$.

Write $E = \{(x, \lambda) \in X \times (0, \infty) : |f(x)| > \lambda\}$, so that $(x, \lambda) \in E \Leftrightarrow x \in E(\lambda)$. Then the left hand side of (B.1) equals

$$\begin{aligned}
&\int_0^\infty \lambda^{p-r-1} \int_X \mathbb{1}_E(x, \lambda) |f(x)|^r \, d\mu(x) d\lambda \\
&= \int_X |f(x)|^r \int_0^\infty \lambda^{p-r-1} \mathbb{1}_E(x, \lambda) \, d\lambda d\mu(x) \qquad \text{by Fubini} \\
&= \int_X |f(x)|^r \int_0^{|f(x)|} \lambda^{p-r-1} \, d\lambda d\mu(x) \qquad \text{since } \lambda < |f(x)| \text{ on } E \\
&= \int_X |f(x)|^r \frac{1}{p-r} |f(x)|^{p-r} \, d\mu(x)
\end{aligned}$$

since $p - r > 0$. Thus we have proved (B.1), and (B.2) is similar. □

Appendix C

Interpolation

Goal

Interpolation of operators on L^p spaces, assuming either weak endpoint bounds (Marcinkiewicz) or strong endpoint bounds (Riesz–Thorin)

References

[Folland] Chapter 6
[Grafakos] Section 1.3

Definition C.1. An operator is *sublinear* if

$$|T(f+g)(y)| \leq |(Tf)(y)| + |(Tg)(y)|$$

$$|T(cf)(y)| = |c||(Tf)(y)|$$

for all f, g in the domain of T, all $c \in \mathbb{C}$, and all y in the underlying set.

Theorem C.2 (Marcinkiewicz Interpolation)**.** *Let $1 \leq p_0 < p_1 \leq \infty$ and suppose (X, μ) and (Y, ν) are measure spaces. Assume*

$$T : L^{p_0} + L^{p_1}(X) \to \{\textit{measurable functions on } Y\}$$

is sublinear. If T is weak (p_0, p_0) and weak (p_1, p_1), then T is strong (p, p) whenever $p_0 < p < p_1$.

Proof. Write A_0, A_1 for the constants in the weak (p_0, p_0) and (p_1, p_1) estimates. Let $\alpha > 0$. Consider $f \in L^p(X), \lambda > 0$. Split f into "large" and "small" parts:

$$g = f\mathbb{1}_{\{x:|f(x)|>\lambda/\alpha\}} \quad \text{and} \quad h = f\mathbb{1}_{\{x:|f(x)|\leq\lambda/\alpha\}}.$$

Notice that

$$\begin{aligned} g &\in L^{p_0}(X) \text{ since } |g|^{p_0} \leq |f|^p(\lambda/\alpha)^{p_0-p}, \\ h &\in L^{p_1}(X) \text{ since } |h|^{p_1} \leq |f|^p(\lambda/\alpha)^{p_1-p}. \end{aligned}$$

Hence $f = g + h \in L^{p_0} + L^{p_1}(X)$. By sublinearity, $|Tf| \leq |Tg| + |Th|$.

Case 1. Assume $p_1 < \infty$. Then

$$\begin{aligned} &\nu\left(\{y \in Y : |Tf(y)| > \lambda\}\right) \\ &\leq \nu\left(\{y \in Y : |Tg(y)| > \lambda/2\}\right) + \nu\left(\{y \in Y : |Th(y)| > \lambda/2\}\right) \quad \text{by sublinearity} \\ &\leq \left(\frac{A_0}{\lambda/2}\|g\|_{L^{p_0}(X)}\right)^{p_0} + \left(\frac{A_1}{\lambda/2}\|h\|_{L^{p_1}(X)}\right)^{p_1} \qquad \text{by the weak estimates on } T \\ &= (2A_0)^{p_0}\lambda^{-p_0}\int_{\{x:|f(x)|>\lambda/\alpha\}} |f(x)|^{p_0}\, d\mu(x) \\ &+ (2A_1)^{p_1}\lambda^{-p_1}\int_{\{x:|f(x)|\leq\lambda/\alpha\}} |f(x)|^{p_1}\, d\mu(x). \end{aligned} \tag{C.1}$$

Therefore

$$\begin{aligned} \|Tf\|_{L^p(Y)}^p &= \int_0^\infty p\lambda^{p-1}\nu\left(\{y \in Y : |Tf(y)| > \lambda\}\right) d\lambda \\ &\leq p(2A_0)^{p_0}\frac{\alpha^{p-p_0}}{p-p_0}\|f\|_{L^p(X)}^p + p(2A_1)^{p_1}\frac{\alpha^{p-p_1}}{p_1-p}\|f\|_{L^p(X)}^p \end{aligned}$$

by (C.1) and formulas (B.1), (B.2). We have proved strong (p, p).

Choosing $\alpha = 2A_1^{p_1/(p_1-p_0)}/A_0^{p_0/(p_1-p_0)}$ gives simple constants:

$$\|Tf\|_{L^p(Y)} \leq 2p^{1/2}\left(\frac{1}{p-p_0} + \frac{1}{p_1-p}\right)^{1/p} A_0^{1-\theta}A_1^\theta\|f\|_{L^p(X)} \tag{C.2}$$

where $0 < \theta < 1$ is determined by expressing $\frac{1}{p}$ as a convex combination of $\frac{1}{p_0}$ and $\frac{1}{p_1}$:

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Note the estimate in (C.2) blows up as p approaches p_0 or p_1.

Case 2. Assume $p_1 = \infty$. Let $\alpha = 2A_1$. We have $\|Th\|_{L^\infty(Y)} \leq A_1\|h\|_{L^\infty(X)}$, because weak (∞,∞) is defined to mean strong (∞,∞), and so

$$\|Th\|_{L^\infty(Y)} \leq A_1\frac{\lambda}{\alpha} = \frac{\lambda}{2}$$

by definitions of h and α. Hence

$$\nu\left(\{y \in Y : |Tf(y)| > \lambda\}\right) \leq \nu\left(\{y \in Y : |Tg(y)| > \lambda/2\}\right)$$

because $|Tf| \leq |Tg| + |Th|$. Now argue like in Case 1 to get strong (p,p). □

Next we weaken the hypotheses of Marcinkiewicz Interpolation.

Definition C.3. Given a measure space (X,μ), write

$$\Sigma(X) = \{\text{simple functions on } X \text{ with support of finite measure}\}.$$

That is, $f \in \Sigma(X)$ provided $f = \sum \alpha_j \mathbb{1}_{F_j}$ where the sum is finite, $\alpha_j \in \mathbb{C} \setminus \{0\}$, and the sets F_j have finite measure and are disjoint.

Remark C.4 (Linear Operators)**.** Suppose

$$T : \Sigma(X) \to \{\text{measurable functions on } Y\}$$

is linear. Then Marcinkiewicz Interpolation still holds: if T is weak (p_0,p_0) and weak (p_1,p_1) on the simple functions in $\Sigma(X)$, then T is strong (p,p) on $L^p(X)$ whenever $p_0 < p < p_1$.

Proof. If f is simple with support of finite measure, then so are $g = f\mathbb{1}_{\{|f|>\lambda/\alpha\}}$ and $h = f\mathbb{1}_{\{|f|\leq\lambda/\alpha\}}$. And $Tf = Tg + Th$ by linearity. Hence the proof of Marcinkiewicz Interpolation gives a strong (p,p) bound for T on $\Sigma(X)$. By density of $\Sigma(X)$ in $L^p(X)$ (using here that $p < p_1$ implies $p < \infty$), we deduce T has a unique extension to a bounded linear operator on $L^p(X)$. (This extension step uses linearity of T.) □

Our next interpolation result needs:

Lemma C.5 (Hadamard's Three Lines)**.** *Assume $H(z)$ is holomorphic on $U = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) < 1\}$ and continuous and bounded on $\overline{U} = \{z \in \mathbb{C} : 0 \leq \mathrm{Re}(z) \leq 1\}$. Let $B_0 = \sup_{\mathrm{Re}(z)=0} |H(z)|$ and $B_1 = \sup_{\mathrm{Re}(z)=1} |H(z)|$.*

Then $|H(z)| \leq B_0^{1-\theta}B_1^\theta$ whenever $\mathrm{Re}(z) = \theta \in [0,1]$.

(Exercise. Let $B_\theta = \sup_{\mathrm{Re}(z)=\theta} |H(z)|$, so that $B_\theta \le B_0^{1-\theta} B_1^\theta$ by the Lemma. Show that $\theta \mapsto \log B_\theta$ is convex.)

Proof. Assume $B_0 > 0$ and $B_1 > 0$. Then

$$G(z) = \frac{H(z)}{B_0^{1-z} B_1^z}$$

is holomorphic on U and bounded on $\overline{U}$, since H is bounded and $|B_0^{1-z} B_1^z| = B_0^{1-\mathrm{Re}(z)} B_1^{\mathrm{Re}(z)} \ge \min(B_0, B_1) > 0$. Let $G_m = G(z) e^{(z^2-1)/m}, m > 0$. Then G_m is holomorphic on U with

$$\begin{aligned} |G_m(z)| &= |G(z)| e^{-(y^2+1)/m} e^{(x^2-1)/m} && \text{where } z = x + iy \\ &\le |G(z)| e^{-(y^2+1)/m} && \text{since } x^2 \le 1 \text{ on } \overline{U}. \end{aligned}$$

Hence $G_m \to 0$ as $|z| \to \infty$ in U. Therefore the Maximum Principle applied to G_m in U says

$$\begin{aligned} \sup_{z \in U} |G_m(z)| &\le \sup_{\partial U \cup \{\infty\}} |G_m| \\ &= \sup_{\partial U} |G_m| \\ &\le \sup_{\partial U} |G| \\ &\le 1, \end{aligned}$$

since $|H| \le B_0$ on $\{\mathrm{Re}(z) = 0\}$ and $|H| \le B_1$ on $\{\mathrm{Re}(z) = 1\}$. Letting $m \to \infty$ gives $|G(z)| \le 1$, which proves the lemma.

If $B_0 = 0$ or $B_1 = 0$, then add ε to H and argue as above. Let $\varepsilon \to 0$. □

Theorem C.6 (Riesz–Thorin Interpolation)**.** *Let $1 \le p_0, p_1, q_0, q_1 \le \infty$, and (X, μ) and (Y, ν) be measure spaces. (If $q_0 = q_1 = \infty$, then further assume ν is semi-finite.) Suppose*

$$T : L^{p_0} + L^{p_1}(X) \to L^{q_0} + L^{q_1}(Y)$$

is linear.

If T is strong (p_0, q_0) and (p_1, q_1), then T is strong (p, q) whenever

$$\Big(\frac{1}{p}, \frac{1}{q}\Big) = (1-\theta)\Big(\frac{1}{p_0}, \frac{1}{q_0}\Big) + \theta\Big(\frac{1}{p_1}, \frac{1}{q_1}\Big)$$

for some $0 < \theta < 1$. Specifically,

$$\|T\|_{L^p(X) \to L^q(Y)} \le \|T\|^{1-\theta}_{L^{p_0}(X) \to L^{q_0}(Y)} \|T\|^{\theta}_{L^{p_1}(X) \to L^{q_1}(Y)}.$$

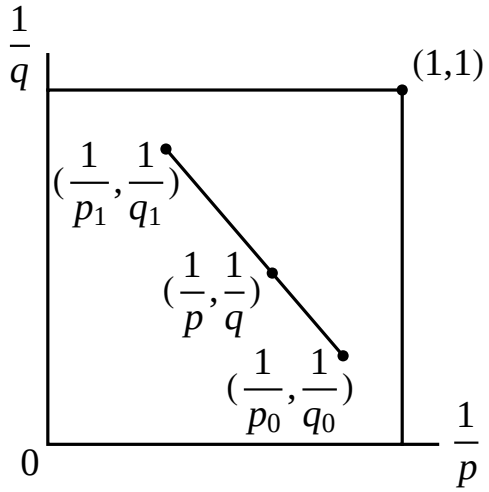

Figure C.1: Parameters in the Riesz–Thorin theorem.

Remark C.7.
1. The relationship between the p and q parameters is shown in Figure C.1. In particular, if $\theta = 0$ then $(p, q) = (p_0, q_0)$, and if $\theta = 1$ then $(p, q) = (p_1, q_1)$.

2. The space

$$L^{p_0} + L^{p_1}(X) = \{f_0 + f_1 : f_0 \in L^{p_0}(X), f_1 \in L^{p_1}(X)\}$$

consists of all sums of functions in L^{p_0} and L^{p_1}. Recall from measure theory that

$$L^p \subset L^{p_0} + L^{p_1},$$

by splitting $f \in L^p$ into large and small parts.

A subtle aspect of the theorem is that when we assume T maps $L^{p_0} + L^{p_1}(X)$ into $L^{q_0} + L^{q_1}(Y)$, we need the value of Tf to be independent of the choice of decomposition $f = f_0 + f_1$.

In applications of the theorem, usually one has T defined and linear on $L^{p_0}(X)$ and $L^{p_1}(X)$, and the two definitions agree on the intersection $L^{p_0} \cap L^{p_1}(X)$. Then one defines T on $f = f_0 + f_1 \in L^{p_0} + L^{p_1}$ by $Tf = Tf_0 + Tf_1$. This definition is independent of the decomposition, as follows. For suppose $f = \widetilde{f}_0 + \widetilde{f}_1$. Then

$$f_0 - \widetilde{f}_0 = \widetilde{f}_1 - f_1 \in L^{p_0} \cap L^{p_1}(X)$$

and so $T\big(f_0 - \widetilde{f}_0\big) = T\big(\widetilde{f}_1 - f_1\big)$, where on the left side we use T defined on $L^{p_0}(X)$ and on the right side we use T on $L^{p_1}(X)$. Linearity of T now yields $Tf_0 + Tf_1 = T\widetilde{f}_0 + T\widetilde{f}_1$ so that the definition of Tf is independent of the decomposition of f.

3. When $T =$ identity, Riesz–Thorin says that

$$L^{p_0} \cap L^{p_1} \subset L^p$$

with

$$\|f\|_{L^p(X)} \le \|f\|^{1-\theta}_{L^{p_0}(X)} \|f\|^{\theta}_{L^{p_1}(X)} \tag{C.3}$$

where $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Here is a direct proof:

$$\begin{aligned}
\|f\|^p_{L^p(X)} &= \int_X |f|^p \, d\mu \\
&= \int_X |f|^{p(1-\theta)} |f|^{p\theta} \, d\mu \\
&\le \Big(\int_X |f|^{p(1-\theta)\cdot p_0/p(1-\theta)} \, d\mu \Big)^{p(1-\theta)/p_0} \Big(\int_X |f|^{p\theta \cdot p_1/p\theta} \, d\mu \Big)^{p\theta/p_1} \quad \text{by Hölder} \\
&= \|f\|^{p(1-\theta)}_{L^{p_0}(X)} \|f\|^{p\theta}_{L^{p_1}(X)}
\end{aligned}$$

Proof of Riesz–Thorin Interpolation. First suppose $p_0 = p_1$, so that $p_0 = p_1 = p$. Then

$$\|Tf\|_{L^q(Y)} \le \|Tf\|^{1-\theta}_{L^{q_0}(Y)} \|Tf\|^{\theta}_{L^{q_1}(Y)}$$

by (C.3) applied to Tf on Y. Now the (p_0, q_0) and (p_1, q_1) bounds can be applied directly to give the (p, q) bound.

Next suppose $p_0 \ne p_1$, so that $p < \infty$.

We will prove an $L^p \to L^q$ bound on Tf for $f \in \Sigma(X)$. Then at the end we will prove the bound for $f \in L^p(X)$.

Let $f \in \Sigma(X)$ and $g \in \Sigma(Y)$, say $f = \sum \alpha_j \mathbb{1}_{F_j}$ and $g = \sum \beta_j \mathbb{1}_{G_j}$. Fix $\theta \in (0,1)$, which fixes p and q. For $z \in \mathbb{C}$, define

$$\begin{aligned}
P(z) &= \frac{p}{p_0}(1-z) + \frac{p}{p_1} z, \\
Q'(z) &= \frac{q'}{q'_0}(1-z) + \frac{q'}{q'_1} z
\end{aligned}$$

where $\frac{1}{q} + \frac{1}{q'} = 1$. (The $'$ in Q' does not denote a derivative, here.) Let

$$\begin{aligned} f_z(x) &= |f(x)|^{P(z)} e^{i \arg f(x)}, && x \in X, \\ g_z(y) &= |g(y)|^{Q'(z)} e^{i \arg g(y)}, && y \in Y. \end{aligned}$$

Note $f_\theta = f$ and $g_\theta = g$, since $P(\theta) = 1$ and $Q'(\theta) = 1$. Clearly

$$g_z = \sum |\beta_j|^{Q'(z)} e^{i \arg \beta_j} \mathbb{1}_{G_j}. \tag{C.4}$$

Therefore $g_z(y)$ is bounded for $y \in Y, z \in \overline{U}$, and it has support (independent of z) with finite measure. Similarly,

$$Tf_z = \sum |\alpha_j|^{P(z)} e^{i \arg \alpha_j} (T \mathbb{1}_{F_j}) \tag{C.5}$$

by linearity, so that

$$\begin{aligned} |Tf_z| &\le \sum |\alpha_j|^{\operatorname{Re} P(z)} |T \mathbb{1}_{F_j}| \\ &\le (\text{const.}) \sum |T \mathbb{1}_{F_j}| \qquad\qquad \text{for } z \in \overline{U}. \end{aligned}$$

The right side belongs to $L^{q_0} \cap L^{q_1}(Y)$ by the strong (p_0, p_0) and (p_1, p_1) bounds, since $\mathbb{1}_{F_j} \in L^{p_0} \cap L^{p_1}(X)$. Hence the function

$$H(z) = \int_Y (Tf_z)(y) g_z(y) \, d\nu(y) \tag{C.6}$$

is well-defined and bounded for $z \in \overline{U}$, by Hölder. And H is holomorphic, as one sees by substituting (C.4) and (C.5) into (C.6) and taking the sums outside the integral. Next,

$$\begin{aligned} \operatorname{Re}(z) = 0 \quad &\Rightarrow \quad \operatorname{Re} P(z) = \frac{p}{p_0}, \qquad \operatorname{Re} Q'(z) = \frac{q'}{q_0'} \\ &\Rightarrow \quad |f_z|^{p_0} = |f|^{p_0 \operatorname{Re} P(z)} = |f|^p \\ &\qquad\quad |g_z|^{q_0'} = |g|^{q_0' \operatorname{Re} Q'(z)} = |g|^{q'} \\ &\Rightarrow \quad \|f_z\|_{L^{p_0}(X)} = \|f\|_{L^p(X)}^{p/p_0} \\ &\qquad\quad \|g_z\|_{L^{q_0'}(Y)} = \|g\|_{L^{q'}(Y)}^{q'/q_0'} \\ &\qquad\qquad\qquad (\text{valid even when } p_0 = \infty \text{ or } q_0 = \infty) \\ &\Rightarrow \quad |H(z)| \le \|Tf_z\|_{L^{q_0}(Y)} \|g_z\|_{L^{q_0'}(Y)} \qquad \text{by Hölder} \\ &\qquad\quad |H(z)| \le \|T\|_{L^{p_0}(X) \to L^{q_0}(Y)} \|f\|_{L^p(X)}^{p/p_0} \|g\|_{L^{q'}(Y)}^{q'/q_1'}. \end{aligned}$$

Similarly,

$$\mathrm{Re}(z) = 1 \quad \Rightarrow \quad |H(z)| \le \|T\|_{L^{p_1}(X)\to L^{q_1}(Y)} \|f\|_{L^p(X)}^{p/p_1} \|g\|_{L^{q'}(Y)}^{q'/q_1'}.$$

Hence by the Hadamard Three Lines Lemma C.5 and a short calculation, if $z = \theta$ then

$$\begin{aligned} |\langle Tf, \overline{g}\rangle| &= |H(\theta)| \\ &\le \|T\|_{L^{p_0}(X)\to L^{q_0}(Y)}^{1-\theta} \|T\|_{L^{p_1}(X)\to L^{q_1}(Y)}^{\theta} \|f\|_{L^p(X)} \|g\|_{L^{q'}(Y)}. \end{aligned}$$

Now the dual characterization of the norm on $L^q(Y)$ implies

$$\|Tf\|_{L^q(Y)} \le \|T\|_{L^{p_0}(X)\to L^{q_0}(Y)}^{1-\theta} \|T\|_{L^{p_1}(X)\to L^{q_1}(Y)}^{\theta} \|f\|_{L^p(X)}, \tag{C.7}$$

which is the desired strong (p,p) bound. (See [Folland, Theorem 6.14] for the dual characterization of the norm, which uses semi-finiteness of ν when $q = \infty$.)

We must extend this bound (C.7) from $f \in \Sigma(X)$ to $f \in L^p(X)$. So fix $f \in L^p(X)$ and let $E = \{x : |f(x)| > 1\}$. Choose a sequence of simple functions $f_n \in \Sigma(X)$ with $|f_n| \le |f|$ and $f_n \to f$ at every point, and with $f_n \to f$ uniformly on $X \backslash E$; such a sequence exists by [Folland, Theorem 2.10]. Define

$$g = f\mathbb{1}_E, \qquad g_n = f_n\mathbb{1}_E,$$

and

$$h = f\mathbb{1}_{X\backslash E}, \qquad h_n = f_n\mathbb{1}_{X\backslash E},$$

so that $f = g + h$, $f_n = g_n + h_n$, and $|g_n| \le |g|, |h_n| \le |h|$. Suppose $p_0 < p_1$, by swapping p_0 and p_1 if necessary. Then $g_n \to g$ in $L^{p_0}(X)$ by dominated convergence, and so $Tg_n \to Tg$ in $L^{q_0}(Y)$. By passing to a subsequence we can further suppose $Tg_n \to Tg$ pointwise a.e.

Also $h_n \to h$ in $L^{p_1}(X)$ by dominated convergence (or, if $p_1 = \infty$, by the uniform convergence $f_n \to f$ on $X \setminus E$). Hence $Th_n \to Th$ in $L^{q_1}(Y)$. By passing to a subsequence we can suppose $Th_n \to Th$ a.e.

Therefore by linearity of T, we have $Tf_n \to Tf$ pointwise a.e. and so

$$\begin{aligned} \|Tf\|_{L^q(Y)} &\le \liminf_n \|Tf_n\|_{L^q(Y)} \qquad \text{by Fatou's lemma} \\ &\le \|T\|_{L^{p_0}(X)\to L^{q_0}(Y)}^{1-\theta} \|T\|_{L^{p_1}(X)\to L^{q_1}(Y)}^{\theta} \liminf_n \|f_n\|_{L^p(X)} \\ &\qquad \text{by (C.7), the strong } (p,p) \text{ bound on the simple functions,} \\ &= \|T\|_{L^{p_0}(X)\to L^{q_0}(Y)}^{1-\theta} \|T\|_{L^{p_1}(X)\to L^{q_1}(Y)}^{\theta} \|f\|_{L^p(X)} \end{aligned}$$

since $f_n \to f$ in $L^p(X)$ by dominated convergence.

We have proved the desired strong (p,p) bound for all $f \in L^p(X)$, and so the proof is complete. $\square$

Bibliography

[Benedetto] J. J. Benedetto. Harmonic Analysis and Applications. CRC-Press, 2001.

[Duoandikoetxea] J. Duoandikoetxea. Fourier Analysis. Transl. D. Cruz-Uribe. Graduate Studies in Mathematics, 29. American Mathematical Society, Providence, RI, 2001.

[Dym and McKean] H. Dym and H. P. McKean. Fourier series and integrals. Probability and Mathematical Statistics, No. 14. Academic Press, New York-London, 1972.

[Edwards and Gaudry] R. E. Edwards and G. I. Gaudry. Littlewood-Paley and Multiplier Theory. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 90. Springer–Verlag, Berlin–New York, 1977.

[Folland] G. B. Folland. Real Analysis. Modern Techniques and their Applications. Second edition. Pure and Applied Mathematics (New York). A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York, 1999.

[Goh and Micchelli] S. S. Goh and C. A. Micchelli. *Uncertainty principles in Hilbert spaces.* J. Fourier Anal. Appl. 8:335–373, 2002.

[Grafakos] L. Grafakos. Classical and Modern Fourier Analysis. Prentice Hall, Upper Saddle River, New Jersey, 2003.

[Gröchenig] K. Gröchenig. Foundations of Time–Frequency Analysis. Birkhäuser, Boston, 2001.

[Jaming] P. Jaming. *Nazarov's uncertainty principles in higher dimension.* J. Approx. Theory 149:30–41, 2007.

[Katznelson] Y. Katznelson. An Introduction to Harmonic Analysis. Second corrected edition. Dover Publications, Inc., New York, 1976.

[Körner] T. W. Körner. Fourier Analysis. Second edition. Cambridge University Press, Cambridge, 1989.

[Rudin] W. Rudin. Fourier Analysis on Groups. Wiley-Interscience, New York, 1990.

[Stein] E. M. Stein. Singular Integrals and Differentiability Properties of Functions. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[Stein and Shakarchi] E. M. Stein and R. Shakarchi. Fourier analysis. An introduction. Princeton Lectures in Analysis, 1. Princeton University Press, Princeton, NJ, 2003.

[Stein and Weiss] E. M. Stein and G. Weiss. Introduction to Fourier Analysis on Euclidean Spaces. Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.

[Strichartz] R. S. Strichartz. A Guide to Distribution Theory and Fourier Transforms. World Scientific Publishing Company, 2003.

[Zygmund] A. Zygmund. Trigonometric Series. Vol. I, II. Third edition. Foreword by R. A. Fefferman. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2002.