Notes on MATH 258: Fourier Analysis

Haoen Li

Latest revision: December 6, 2025

This is a note on the Harmonic Analysis course given by Prof. Ruixiang Zhang (UC Berkeley) in Fall 2025. This course mainly cover the materials in Ch. 1-5 + 8.1-8.4 in the GSM book *Fourier Analysis* by Javier Duoandikoetxea.

Contents

Chapter 1

Introduction

References:

- *Classical & Modern Fourier Analysis* by Loukas Grafakos;
- *Introduction to Fourier Analysis in Euclidean Spaces* by Elias M. Stein and Guido Weiss;
- *Singular Integrals and Differentiability Properties of Functions* by Elias M. Stein;
- *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals* ("dictionary") by Elias M. Stein.

Motivations of Fourier Analysis:

- FT "diagonalizes" differential operators on $\mathbb{R}^d$ or torus.

$$f(x) = \text{"the linear combination of } e^{2\pi \mathrm{i}\xi\cdot x}\text{"} = \int_{\mathbb{R}^d} \widehat{f}(\xi)e^{2\pi\mathrm{i}\xi\cdot x}\mathrm{d}\xi. \tag{1.0.1}$$

  The advantage: $e^{2\pi\mathrm{i}\xi\cdot x}$ are eigenfunctions of the differential operators.

- FT "occasionally diagonalizes" translation operators (applications to geometric problems).
- FT can also be used to study "rotations" (DOES NOT directly diagonalize, but related).
- FT is closely related to convolution operators to functions (after all, the convolution can be viewed as *an integral operator defined using translation*)

$$(g * f)(x) = \int_{\mathbb{R}^d} g(x-\xi)f(\xi)\mathrm{d}\xi. \tag{1.0.2}$$

Applications of Fourier Analysis:

Example 1 (Schrödinger equation)**.** $u(x,t) = e^{\mathrm{i}t\triangle}u(x,0)$*, where* $\triangle$ *is the differential (Laplacian) operator. FT provides some estimation about the solution on the time evolution.*

Example 2 (Fractal geometry)**.** *e.g. Falconer distance conjecture. For compact $E \subset \mathbb{R}^2$, if the Hausdorff dimension of $E > 1$, then*

$$\left|\underbrace{\{|x-y| : x, y \in E\}}_{\subset \mathbb{R}_{\geq 0}}\right| > 0. \tag{1.0.3}$$

The proof of the weaker theorem ($\dim_H E > \frac{5}{4}$) is based on Fourier analysis.

Example 3. *Applications in tomography (Radon transform) or signal processing (e.g. image reconstruction).*

Example 4. *Applications in number theory (use mean value estimation of L^p-functions to study the behaviors of $\mathbb{N}$).*

Main Topics:

- Fundamental problems in harmonic analysis.

  FT:
  $$f(x) = \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi i \xi \cdot x} \mathrm{d}\xi. \tag{1.0.4}$$

  Fourier series (on $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$):
  $$f(x) = \sum_{n \in \mathbb{Z}^d} \widehat{f}(n) e^{2\pi i n \cdot x}. \tag{1.0.5}$$

  Questions: How to solve $\widehat{f}$? In what sense the formula
  $$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \xi \cdot x} \mathrm{d}x \tag{1.0.6}$$

  holds? In what sense do we have the convergence
  $$\int_{|\xi|<R} \widehat{f}(\xi) e^{2\pi i \xi \cdot x} \mathrm{d}\xi \xrightarrow{R \to \infty} f(x) \quad \text{or} \quad \sum_{|n|<N} \widehat{f}(n) e^{2\pi i n \cdot x} \xrightarrow{N \to \infty} f(x)? \tag{1.0.7}$$

- Study of L^p-spaces and operators on them, using Fourier analysis techniques.

Technical Topics:

- *Quantitative* analysis problems. e.g. In 1-dimensional, sometimes the property is less sensitive to the "integral weight" p of L^p spaces, how can we understand this?
- "The art of decomposition and recomposition". e.g. $A = \underbrace{A_1}_{\leq ?\text{ (technique 1)}} + \underbrace{A_2}_{\leq ?\text{ (technique 2)}} \leq B(\surd)$
- Perhaps useful way of thinking when working on other problems (mindset related to Fourier analysis).

Chapter 2

Fourier series and Fourier transforms

2.1 Fourier series on $\mathbb{T}$

Study $f : \mathbb{T} = \mathbb{T}^1 = [0, 1] \to \mathbb{C}$. Fourier:

$$f \sim \sum_{k=0}^{\infty} [a_k \cos(2\pi kx) + b_k \sin(2\pi kx)] . \tag{2.1.1}$$

Fourier noticed the "nice diagonalization":

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) \sim \sum_{k=0}^{\infty} [2\pi k b_k \cos(2\pi kx) - 2\pi k a_k \sin(2\pi kx)] . \tag{2.1.2}$$

Modern notations:

$$f(x) \sim \sum_{n\in\mathbb{Z}} \widehat{f}(n) e^{2\pi inx}, \quad \text{where } \widehat{f}(n) \text{“should be”} \int_0^1 f(x) e^{-2\pi inx} \mathrm{d}x \text{ by “orthogonality”.} \tag{2.1.3}$$

A natural question: Does the RHS converge (in L^p; pointwise; almost surely pointwise)? Does the RHS converge to the LHS?

The first positive result is given by P. G. L. Dirichlet (1829), which proved that "under some conditions" $\lim S_N f(x)$ exists and equals $\frac{f(x+)+f(x-)}{2}$. The idea is to represent the partial sum

$$S_N f(x) := \sum_{n=-N}^{N} \widehat{f}(n) e^{2\pi \mathrm{i} nx} \tag{2.1.4}$$

using a more manageable convolution expression by the *Dirichlet kernel*.

$$S_N(f)(x) = \sum_{|n|\le N} \int_0^1 f(t) e^{-2\pi int} \mathrm{d}t e^{2\pi \mathrm{i} nx} = \int_0^1 f(t) \underbrace{\sum_{|n|\le N} e^{2\pi \mathrm{i} n(x-t)}}_{D_N(x-t)} \mathrm{d}t = (f * D_N)(x). \tag{2.1.5}$$

The function D_N is called the Dirichlet kernel. We can compute it explicitly using geometric series:

$$D_N(t) = \sum_{|n|\le N} e^{2\pi \mathrm{i} nt} = \frac{\sin\Big((N + \frac{1}{2})2\pi t\Big)}{\sin(\pi t)}. \tag{2.1.6}$$

We easily find that

Proposition 1.

$$\int_0^1 D_N(t)\mathrm{d}t = 1, \quad |D_N(t)| \leq \frac{1}{|\sin(\pi t)|} \leq \frac{1}{\sin \pi\delta} \quad \text{for } \delta \leq |t| \leq \frac{1}{2}. \tag{2.1.7}$$

Proof. By noting that for $f(x) \equiv 1$, $\widehat{f}(n) = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases}$. Thus $S_N(f)(x) = 1$ for all N, which implies $\int_0^1 D_N(t)\mathrm{d}t = 1$. □

Remark 1. *We will see that the role of D_N is similar to "localization". Just like many other reproducing kernels, the convolution by D_N tries to reproduce the information of f by localizing the mass of f near x. Thus the convergence of a Fourier series is effectively a local property (i.e. the modification of f outside of a neighborhood of x does not affect the limit).*

Proposition 2 (Estimation of Fourier coefficients)**.** *To make $\widehat{f}$ well-defined, we assume $f \in L^1(\mathbb{T})$.*

- $\left|\widehat{f}(n)\right| \leq \|f\|_{L^1(\mathbb{T})}$ *for all $n \in \mathbb{Z}$.*
- *(Riemann-Lebesgue) $\widehat{f}(n) \to 0$ as $|n| \to \infty$.*

Remark 2. *The idea for Riemann-Lebesgue is to find the mechanism of* cancellation *in the oscillatory integral when the frequency becomes large.*

Proof. (1) trivial. (2) by periodicity, we have $\widehat{f}(n) = \int_0^1 f(x)e^{-2\pi inx}\mathrm{d}x = -\int_0^1 f(x)e^{-2\pi i(x+\frac{1}{2n})}\mathrm{d}x = -\int_0^1 f(x-\frac{1}{2n})e^{-2\pi inx}\mathrm{d}x$. Hence we have

$$\widehat{f}(n) = \frac{1}{2}\int_0^1 [f(x) - f(x-1/2n)]e^{-2\pi inx}\mathrm{d}x \tag{2.1.8}$$

Given $\varepsilon > 0$, for arbitrary $f \in L^1(\mathbb{T})$, we choose a continuous function $g \in C(\mathbb{T})$ such that $\|f-g\|_{L^1(\mathbb{T})} < \varepsilon/2$. Note that $\widehat{g}(n) \to 0$ as $|n| \to \infty$ by the uniform continuity of g, we can choose n large enough such that $\left|\widehat{g}(n)\right| < \varepsilon/2$. Thus we have

$$\left|\widehat{f}(n)\right| \leq \left|\widehat{g}(n)\right| + \left|(\widehat{f-g})(n)\right| < \varepsilon/2 + \|f-g\|_{L^1(\mathbb{T})} < \varepsilon. \tag{2.1.9}$$

□

Theorem 2.1.1 (Dini's criterion)**.** $f \in L^1(\mathbb{T})$. *If $x \in \mathbb{T}$, $\delta > 0$, such that*

$$\int_{|t|<\delta} \frac{|f(x+t) - f(x)|}{|t|}\mathrm{d}t < \infty, \tag{2.1.10}$$

then $\lim_{N\to\infty} S_N f(x) = f(x)$.

Remark 3. *An example of this is the α-Hölder condition for $\alpha > 0$.*

Theorem 2.1.2 (Jordan's criterion)**.** *f has bounded variation in $(x-\delta, x+\delta)$, $x \in \mathbb{T}$, $\delta > 0$, and f is bounded on $\mathbb{T}$, then*

$$\lim_{N\to\infty} S_N f(x) = \frac{f(x+) - f(x-)}{2}. \tag{2.1.11}$$

Proof of Dini's criterion. Note that

$$\begin{aligned} |S_N f(x) - f(x)| &= \left| \int_0^1 f(x+t) D_N(t) \mathrm{d}t - f(x) \right| \\ &= \left| \int_0^1 [f(x+t) - f(x)] D_N(t) \mathrm{d}t \right| \\ &= \left| \int_0^1 \frac{f(x+t) - f(x)}{\sin(\pi t)} \sin((2N+1)\pi t) \mathrm{d}t \right| \to 0 \quad (N \to \infty). \end{aligned} \tag{2.1.12}$$

This is because (1) $\frac{f(x+t)-f(x)}{\sin(\pi t)}$ is $L^1(\mathbb{T})$ (by computing the integral $\int_0^1 \left|\frac{f(x+t)-f(x)}{\sin(\pi t)}\right| \mathrm{d}t < \infty$ near and away from $t = 0$ and $t = 1$) and (2) Riemann-Lebesgue lemma applies. $\square$

Proof of Jordan's criterion. Without loss of generality, we can assume that $x = 0$ and f is monotonic (any BV function can be decomposed into difference of two monotonic functions). Since

$$S_N f(x) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x-t) D_N(t) \mathrm{d}t = \int_0^{\frac{1}{2}} [f(x+t) + f(x-t)] D_N(t) \mathrm{d}t \tag{2.1.13}$$

It suffices to show that

$$\int_0^{\frac{1}{2}} f(t) D_N(t) \mathrm{d}t \to \frac{1}{2} f(0+) \quad (N \to \infty). \tag{2.1.14}$$

For this, we want to estimate

$$\left| \int_0^{\frac{1}{2}} [f(t) - f(0+)] D_N(t) \mathrm{d}t \right| \le \left| \int_0^{\delta} \right| + \left| \int_{\delta}^{\frac{1}{2}} \right| \tag{2.1.15}$$

The second part $\to 0$ as $N \to \infty$, which follows from the integrability $\frac{f(t)-f(0+)}{\sin(\pi t)}$ (since it has no singularity) and Riemann-Lebesgue lemma. For the first part, since $t \mapsto f(t) - f(0+)$ is monotonic, by the **second mean-value theorem of integration**, we have $\exists \nu \in [0, \delta]$ s.t.

$$\begin{aligned} \int_0^{\delta} [f(t) - f(0+)] D_N(t) &= [f(0+) - f(0+)] \int_0^{\nu} D_N(t) \mathrm{d}t + [f(\delta) - f(0+)] \int_{\nu}^{\delta} D_N(t) \mathrm{d}t \\ &= [f(\delta) - f(0+)] \int_{\nu}^{\delta} D_N(t) \mathrm{d}t. \end{aligned} \tag{2.1.16}$$

But we note that

$$\left| \int_{\nu}^{\delta} D_N(t) \mathrm{d}t \right| \le \left| \int_{\nu}^{\delta} \frac{\sin(2N+1)\pi t}{\pi t} \mathrm{d}t \right| + \left| \int_{\nu}^{\delta} \left(\frac{1}{\sin(\pi t)} - \frac{1}{\pi t} \right) \sin(2N+1)\pi t \mathrm{d}t \right| \tag{2.1.17}$$

$$\text{The first term} = \frac{1}{\pi} \left| \int_{(2N+1)\pi\nu}^{(2N+1)\pi\delta} \frac{\sin t}{t} \mathrm{d}t \right| \le \frac{1}{\pi} \left| \int_0^{(2N+1)\pi\nu} \right| + \left| \int_0^{(2N+1)\pi\delta} \right| \le \frac{2}{\pi} \sup_{M>0} \left| \int_0^M \frac{\sin t}{t} \mathrm{d}t \right| \overset{\text{convergence}}{\le} C_1, \tag{2.1.18}$$

$$\text{The second term} \le \int_{\nu}^{\delta} \left| \frac{1}{\sin(\pi t)} - \frac{1}{\pi t} \right| \mathrm{d}t \overset{\text{the integrand has no singularity}}{\le} C_2. \tag{2.1.19}$$

Here $C_1, C_2 < \infty$ are independent of N. Therefore,

$$\left|\int_0^{\delta}[f(t)-f(0+)]D_N(t)\mathrm{d}t\right| \le (C_1+C_2)|f(\delta)-f(0+)| \tag{2.1.20}$$

For any $\varepsilon > 0$, we choose $\delta = \delta(\varepsilon) > 0$ such that $|f(t)-f(0+)| < \varepsilon$ for $t \in (0,\delta)$ (monotonic function admits right-hand limit), then

$$0 \le \limsup_{N\to\infty}\left|\int_0^{\frac{1}{2}}[f(t)-f(0+)]D_N(t)\mathrm{d}t\right| \le (C_1+C_2)\varepsilon. \tag{2.1.21}$$

By the arbitrariness of ε, we conclude that the limit exists and

$$\lim_{N\to\infty}\left|\int_0^{\frac{1}{2}}[f(t)-f(0+)]D_N(t)\mathrm{d}t\right| = 0. \tag{2.1.22}$$

□

2.2 Existence of divergent Fourier series for $f \in C(\mathbb{T})$: an application of Banach-Steinhaus theorem

A natural question is whether the Fourier series converges for all continuous functions. The answer is negative.

Theorem 2.2.1. *There exists a continuous function $f \in C(\mathbb{T})$ such that the Fourier series of f diverges at $x = 0$.*

Proof.

- We define two Banach spaces: $X = C(\mathbb{T})$ with the sup-norm and $Y = \mathbb{C}$ with the norm $|z|$. We consider the familiy of linear operators $T_N : X \to Y$ defined by $T_N(f) = S_N f(0)$.
- We note that $\|T_N\|$ is a bounded linear operator for any $N \in \mathbb{N}$. In fact, we have

$$|T_N f| = |S_N f(0)| \le \|f\|_X \int_0^1 |D_N(t)|\mathrm{d}t \le \|f\|_X \sum_{|n|\le N}\int_0^1 \left|e^{2\pi i n t}\right|\mathrm{d}t = (2N+1)\|f\|_X. \tag{2.2.1}$$

- We claim that $\|T_N\|_{L(X,Y)} \to \infty (N \to \infty)$. In fact we need to find some function that behaves like $\mathrm{sgn}(D_n(t))$. By modifying the sign function on a small neighborhood of each discontinuity, we can find a continuous function $f_\varepsilon \in C(\mathbb{T})$ such that $\|f_\varepsilon\|_X = 1$ and $|T_N f_\varepsilon| \ge L_N - \varepsilon$. Thus we have $\|T_N\|_{L(X,Y)} \ge L_N$, where L_N is the L^1-norm of D_N.
- We estimate L_N as follows:

$$\begin{aligned}
L_N &= \int_0^1 |D_N(t)|\mathrm{d}t = 2\int_0^{\frac{1}{2}}\left|\frac{\sin((2N+1)\pi t)}{\sin(\pi t)}\right|\mathrm{d}t \ge 2\sum_{k=1}^{N}\int_{\frac{k-1}{2N+1}}^{\frac{k}{2N+1}}\left|\frac{\sin((2N+1)\pi t)}{\sin(\pi t)}\right|\mathrm{d}t \\
&= 2\sum_{k=1}^{N}\int_{k-1}^{k}\frac{|\sin(\pi t)|}{\sin(\pi t/(2N+1))/(2N+1)}\mathrm{d}t \ge 2\sum_{k=1}^{N}\int_{k-1}^{k}\frac{|\sin(\pi t)|}{k\pi}\mathrm{d}t = \frac{2}{\pi}\sum_{k=1}^{N}\frac{1}{k}\int_0^1 \sin\pi t\mathrm{d}t \\
&= \frac{4}{\pi^2}\sum_{k=1}^{N}\frac{1}{k} \to \infty \quad (N\to\infty).
\end{aligned} \tag{2.2.2}$$

- Therefore $\sup_{N\in\mathbb{N}} \|T_N\|_{L(X,Y)} = \infty$. By the Banach-Steinhaus theorem, there exists $f \in X$ such that $\sup_{N\in\mathbb{N}} |T_N f| = \infty$, which means that the Fourier series of f diverges at $x = 0$.

□

2.3 L^p-convergence of Fourier series via summability methods

We will first try to study the L^p-convergence of Fourier series for partial sum. By Young's inequality of convolution, we have

$$\|S_N f\|_{L^p(\mathbb{T})} = \|f * D_N\|_{L^p(\mathbb{T})} \leq \|f\|_{L^p(\mathbb{T})}\|D_N\|_{L^1(\mathbb{T})}. \tag{2.3.1}$$

This means that if $f \in L^p(\mathbb{T})$, then $S_N f$ is also L^p. However, simply applying this just gives us pessimistic estimation since the L^1-norm of D_N diverges as $N \to \infty$. In fact, the L^p convergence of $S_N f$ only holds conditionally, i.e. the following theorem holds:

Theorem 2.3.1. *For $1 \leq p < \infty$, $S_N f \to f$ in $L^p(\mathbb{T})$, iff $f \in L^p(\mathbb{T})$ and there exists a constant C_p independent of N and f, such that $\|S_N f\|_{L^p(\mathbb{T})} \leq C_p\|f\|_{L^p(\mathbb{T})}$.*

Proof. The necessity is trivial (follows by the Banach-Steinhaus theorem, noting that if the familiy of operators $\{S_N\}_{N\in\mathbb{N}}$ is bounded on each orbit, then it is uniformly bounded). For the sufficiency, we first note that if g itself is a trigonometric polynomial, then $S_N g = g$ for sufficiently large N. Since the trigonometric polynomials are dense in $L^p(\mathbb{T})$ (by Stone-Weierstrass, or the very simple corollary that we will discuss later), for any $\varepsilon > 0$, we can choose a trigonometric polynomial g such that $\|f - g\|_{L^p(\mathbb{T})} < \varepsilon$. Thus we have

$$\|S_N f - f\|_{L^p(\mathbb{T})} \leq \|S_N(f - g)\|_{L^p(\mathbb{T})} + \|S_N g - g\|_{L^p(\mathbb{T})} + \|g - f\|_{L^p(\mathbb{T})} \leq (C_p + 1)\varepsilon. \tag{2.3.2}$$

Note that ε is arbitrary, we conclude that $S_N f \to f$ in $L^p(\mathbb{T})$. □

We will give some interesting remark about this theorem.

Remark 4. *The inequality in the theorem seems strong, but in fact if $1 < p < \infty$, then it holds. Therefore, if $f \in L^p(1 < p < \infty)$, then the Fourier series is L^p convergent. This is a consequence of a celebrated theorem and we will deal with it in the later chapters.*

When $p = 1$, the inequality CANNOT hold since again we have $\|S_N\|_{L^1(\mathbb{T})\to L^1(\mathbb{T})} = \|D_N\|_{L^1(\mathbb{T})} \to \infty$ as $N \to \infty$.

When $p = 2$, thanks to the orthogonality structure, we can easily see that the inequality holds with $C_2 = 1$ (Parseval's identity).

Remark 5. *For the question whether $S_N f \to f$ almost everywhere, the answer is much more complicated. The statement is* false *for $p = 1$ (counterexample by Kolmogorov) and* true *for $1 < p < \infty$ ($p = 2$ by Calerson and $p > 1$ for Hunt).*

cf. T. Tao 247B, note 4 in May 2020.

In order to avoid the "conditional" convergent of Fourier series of L^p functions (in particular, L^1 functions), we can consider a "better" summablility method. The *Cesàro mean* of the partial is somehow a "smoother" version of the partial sum, and thus is more likely to give better convergence.

Definition 2.3.2 (Cesàro mean)**.** *The Cesàro mean of the partial sum of Fourier series is defined as*

$$\sigma_N f(x) := \frac{1}{N+1}\sum_{k=0}^{N} S_k f(x). \tag{2.3.3}$$

We can also represent $\sigma_N f$ as a convolution operator by the Fejér kernel F_N*:*

$$\sigma_N f(x) = (f * F_N)(x), \quad \textit{where } F_N(t) = \frac{1}{N+1}\sum_{k=0}^{N} D_k(t) = \frac{1}{N+1}\left(\frac{\sin((N+1)\pi t)}{\sin(\pi t)}\right)^2. \tag{2.3.4}$$

The Fejér kernel has the following properties:

Proposition 3.

- $F \geq 0$, $\|F_N\|_{L^1(\mathbb{T})} = \int_0^1 |F_N(t)|\mathrm{d}t = \int_0^1 F_N(t)\mathrm{d}t = 1$. *[This is a very good property compared to the Dirichlet kernel, which means that there is no contribution form the cancellation of oscillatory in the tail of the kernel function.]*
- *Localization: For any $\delta > 0$, $\int_\delta^{\frac{1}{2}} F_N(t)\mathrm{d}t \to 0$ as $N \to \infty$.*
- $$F_N(t) \leq \frac{1}{\sin \pi\delta^2}\frac{1}{N+1} \leq \frac{4}{\pi^2\delta^2(N+1)} \quad \textit{for } \delta \leq |t| \leq \frac{1}{2}. \tag{2.3.5}$$

Theorem 2.3.3 (L^p-convergence of Cesàro mean)**.** *For $1 \leq p < \infty$, $f \in L^p(\mathbb{T})$, or for $p = \infty$ and $f \in C(\mathbb{T}) \subset L^\infty(\mathbb{T})$ (L^∞ spaces are generally to large and we want to consider a smaller/seperable space of continuous functions), then $\sigma_N f \to f$ in $L^p(\mathbb{T})$.*

Proof. Since $\int F_N = 1$, by Minkowski's inequality, we have

$$\begin{aligned} \|\sigma_N f - f\|_p &= \left\| \int_{-\frac{1}{2}}^{\frac{1}{2}} [f(x-t) - f(x)]F_N(t)\mathrm{d}t \right\|_p \leq \int_{-\frac{1}{2}}^{\frac{1}{2}} \|f(\cdot - t) - f(\cdot)\|_p F_N(t)\mathrm{d}t \\ &= \int_{|t|<\delta} \|f(\cdot - t) - f(\cdot)\|_p F_N(t)\mathrm{d}t + 2\|f\|_p \int_{\delta<|t|\leq\frac{1}{2}} F_N(t)\mathrm{d}t \end{aligned} \tag{2.3.6}$$

Note that $\mathbb{T}$ is compact and thus $C(\mathbb{T})$ is dense in $L^p(\mathbb{T})(1 \leq p < \infty)$, thus for any $\varepsilon > 0$, we can choose $g \in C(\mathbb{T})$ such that $\|f - g\|_p < \varepsilon/3$. Since g is uniformly continuous, we can choose $\delta = \delta(\varepsilon) > 0$ such that $\|g(\cdot - t) - g(\cdot)\|_p < \varepsilon/3$ for $|t| < \delta$. Thus we have

$$\|f(\cdot - t) - f(\cdot)\|_p \leq \|g(\cdot - t) - g(\cdot)\|_p + 2\|f - g\|_p \leq \|g(\cdot - t) - g(\cdot)\|_p + 2\varepsilon/3 \leq \varepsilon, \quad \text{for } |t| < \delta. \tag{2.3.7}$$

Same statement obviously also holds for $p = \infty$ and $f \in C(\mathbb{T})$. Therefore, we have

$$\|\sigma_N f - f\|_p \leq \varepsilon + 2\|f\|_p \int_{\delta<|t|\leq\frac{1}{2}} F_N(t)\mathrm{d}t \overset{\text{localization}}{\Longrightarrow} 0 \leq \limsup_{N\to\infty} \|\sigma_N f - f\|_p \leq \varepsilon. \tag{2.3.8}$$

By the arbitrariness of ε, we conclude that $\lim_{N\to\infty} \|\sigma_N f - f\|_p$ exists and equals 0.

□

We will give some very interesting consequence of the nice convergent property of Cesàro mean.

Corollary 1. • *The trigonometric polynomials are dense in $L^p(\mathbb{T})$ for $1 \le p < \infty$ and in $C(\mathbb{T})$. [(1) This is because $\sigma_N f$ itself is a trigonometric polynomial. (2) The Stone-Weierstrass theorem (a general result for the structure of $C(K)$ for K being a compact metric space) implies that the trigonometric polynomials are dense in $C(\mathbb{T})$. But $C(\mathbb{T})$ is again dense in $L^p(\mathbb{T})$ for $1 \le p < \infty$, therefore we also have the density of trigonometric polynomials in $L^p(\mathbb{T})$. However, the proof of Stone-Weierstrass theorem is not so elementary.]*

- *(Uniqueness) If $f \in L^1(\mathbb{T})$ and $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$ a.e. [This is because $\sigma_N f$ is always zero but tries to converge to f in $L^1(\mathbb{T})$, which forces f to be zero.]*

Another good summability method is related to the *Poisson kernel*. This time we view the Fourier series as a formal power series in the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$:

$$u(z) = \sum_{k=0}^{\infty} \widehat{f}(k) z^k + \sum_{k=-\infty}^{-1} \widehat{f}(k) \bar{z}^{|k|}, \quad z = re^{2\pi i\theta} \in \mathbb{D}. \tag{2.3.9}$$

Note that for Fourier coefficients $\{\widehat{f}(k)\}$, by Riemann-Lebesgue lemma it is always a bounded sequence. Thus for $|z| < 1$ the power series converge absolutely and $u(z)$ is well-defined. We can then present $u(z)$ as a convolution operator by the *Poisson kernel* P_r:

$$u(re^{2\pi i\theta}) = \sum_{k=-\infty}^{\infty} \widehat{f}(k) r^{|k|} e^{2\pi i k\theta} = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\varphi) \sum_{k\in\mathbb{Z}} r^{|k|} e^{2\pi i k(\theta-\varphi)} \mathrm{d}\varphi = (f * P_r)(\theta), \tag{2.3.10}$$

Here,

$$P_r(\varphi) := \sum_{k\in\mathbb{Z}} r^{|k|} e^{2\pi i k\varphi} = \frac{1 - r^2}{1 - 2r\cos(2\pi\varphi) + r^2} = \mathrm{Re}\left(\frac{1 + re^{2\pi i\varphi}}{1 - re^{2\pi i\varphi}}\right). \tag{2.3.11}$$

$P_r(\varphi)$ also has the analogous properties as the Fejér kernel:

Proposition 4. *For $0 < r < 1$,*

- $P_r(\varphi) > 0$, $\int_{-\frac{1}{2}}^{\frac{1}{2}} P_r(\varphi)\mathrm{d}\varphi = 1$.
- *Localization: For any $\delta > 0$, $\int_{\delta \le |\varphi| \le \frac{1}{2}} P_r(\varphi)\mathrm{d}\varphi \to 0$ as $r \to 1^-$.*

Therefore, we have the following theorem:

Theorem 2.3.4. *If $f \in L^p$, $1 \le p < \infty$ or f is continuous and $p = \infty$, then $\lim_{r\to 1^-} \left\| u(re^{2\pi i\theta}) - f(\theta) \right\|_p = 0$.*

Remark 6 (Remark of history)**.** *The historical motivation of the Poisson kernel is related to some problems in complex analysis. Another motivation is, since u is (complex) harmonic, it solves the Dirichlet problem in $\mathbb{D}$ with boundary value f.*

2.4 The L^p theory of Fourier transforms

Definition 2.4.1. *Given $f \in L^1(\mathbb{R}^n)$, its FT is $\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi} \mathrm{d}x$.*

Remark 7. *If we change L^1 to L^p, then unlike on the torus, $\widehat{f}(\xi)$ is generally not well-defined (L^1 function does not imply L^p for $p > 1$).*

We will prove later than the Fourier inversion formula

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{2\pi \mathrm{i} x\cdot\xi} \mathrm{d}\xi \tag{2.4.1}$$

holds in suitable sense. (In fact, $\widehat{f}(\xi)$ is generally not integrable, thus the integral does not make sense in the Lebesgue sense.)

We list some useful properties of FT. All of these properties follows directly by expanding the definition of FT.

Proposition 5. *We assume the functions are all in $L^1(\mathbb{R}^n)$.*

- $\widehat{\alpha f + \beta g} = \alpha \widehat{f} + \beta \widehat{g}$ *(linearity).*
- $\left\|\widehat{f}\right\|_\infty \leq \|f\|_1$ *(boundedness).*
- *For $f \in L^1(\mathbb{R}^n)$, $\widehat{f}$ is continuous (by dominated convergence theorem, we take the dominating function to be $|f|$).*
- $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ *(Riemann-Lebesgue lemma, by the same idea as on the torus).*
- *(convolution) If $f, g \in L^1(\mathbb{R}^n)$, then $f * g \in L^1(\mathbb{R}^n)$ and $(\widehat{f * g})(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$. (just expanding the definition and using Fubini's theorem)*
- *(translation) If $f \in L^1(\mathbb{R}^n)$, then $(\widehat{f(\cdot + h)})(\xi) = e^{2\pi \mathrm{i} h\cdot\xi}\widehat{f}(\xi)$ for any $h \in \mathbb{R}^n$.*
- *(rotation) If $\rho \in \mathrm{O}(n)$, then $(\widehat{f(\rho\cdot)})(\xi) = \widehat{f}(\rho\xi)$.*

  Proof. Expanding the definition, we have

$$\begin{aligned} \widehat{f(\rho\cdot)}(\xi) &= \int_{\mathbb{R}^n} f(\rho x) e^{-2\pi \mathrm{i} x\cdot\xi} \mathrm{d}x = \int_{\mathbb{R}^n} f(y) e^{-2\pi \mathrm{i}\xi\cdot\rho^T y} \mathrm{d}y \\ &= \int_{\mathbb{R}^n} f(y) e^{-2\pi \mathrm{i}(\rho\xi)\cdot y} \mathrm{d}y = \widehat{f}(\rho\xi). \end{aligned} \tag{2.4.2}$$

  □

- *(scaling) If $f \in L^1(\mathbb{R}^n)$, then for any $\lambda > 0$, $(\widehat{f(\lambda\cdot)})(\xi) = \frac{1}{\lambda^n}\widehat{f}(\xi/\lambda)$.*

  Proof. By noting that the Jacobian ($|\det(\lambda I_n)|$) of the transformation $x \mapsto \lambda x$ is λ^n. □

- *(differentiation) $(\widehat{\partial_j f})(\xi) = 2\pi \mathrm{i}\xi_j \widehat{f}(\xi)$ if $\partial_j f \in L^1(\mathbb{R}^n)$.*

  Remark 8. *We see that the Fourier transform is somehow a good "diagonalization" of differential operators.*

Proof. Since the $C_c^\infty(\mathbb{R}^n)$ is L^1-dense, we only need to prove it for $f \in C_c^\infty(\mathbb{R}^n)$. The integration over $\mathbb{R}$ of compact supported smooth function satisfies the integration-by-parts formula. By integration by parts, we have

$$\begin{aligned}\widehat{(\partial_j f)}(\xi) &= \int_{\mathbb{R}^n} \partial_j f(x) e^{-2\pi \mathrm{i} x\cdot\xi} \mathrm{d}x = \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} \partial_j f(x) e^{-2\pi \mathrm{i} x_j \xi_j} \mathrm{d}x_j \right) e^{-2\pi \mathrm{i} x'\cdot\xi'} \mathrm{d}x' \\ &= -\int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} f(x)(-2\pi \mathrm{i}\xi_j) e^{-2\pi \mathrm{i} x_j \xi_j} \mathrm{d}x_j \right) e^{-2\pi \mathrm{i} x'\cdot\xi'} \mathrm{d}x' \\ &= 2\pi \mathrm{i}\xi_j \int_{\mathbb{R}^n} f(x) e^{-2\pi \mathrm{i} x\cdot\xi} \mathrm{d}x = 2\pi \mathrm{i}\xi_j \widehat{f}(\xi).\end{aligned} \tag{2.4.3}$$

□

- *$(\widehat{-2\pi \mathrm{i} x_j} f)(\xi) = \partial_j \widehat{f}(\xi)$ if $x_j f \in L^1(\mathbb{R}^n)$.*

  Proof. Since $C_c^\infty(\mathbb{R}^n)$ is L^1-dense in $L^1(\mathbb{R}^n)$, we only need to prove it for $f \in C_c^\infty(\mathbb{R}^n)$. By noting that $\partial_j e^{-2\pi \mathrm{i} x\cdot\xi} = -2\pi \mathrm{i} x_j e^{-2\pi \mathrm{i} x\cdot\xi}$, we have

$$\begin{aligned}(\widehat{-2\pi \mathrm{i} x_j} f)(\xi) &= \int_{\mathbb{R}^n} -2\pi \mathrm{i} x_j f(x) e^{-2\pi \mathrm{i} x\cdot\xi} \mathrm{d}x = \int_{\mathbb{R}^n} f(x) \partial_j e^{-2\pi \mathrm{i} x\cdot\xi} \mathrm{d}x \\ &\overset{f \text{ has compact support}}{=} \partial_j \int_{\mathbb{R}^n} f(x) e^{-2\pi \mathrm{i} x\cdot\xi} \mathrm{d}x = \partial_j \widehat{f}(\xi).\end{aligned} \tag{2.4.4}$$

  □

We are not satisfied with Fourier transform only for L^1 functions. To study the L^p theory for Fourier transforms, we need to introduce the **Schwartz class** and **tempered distributions**.

Definition 2.4.2 (Schwartz class)**.** • *For smooth function $f \in C^\infty(\mathbb{R}^n)$, we define the seminorm $p_{\alpha,\beta} : C^\infty(\mathbb{R}^n) \to [0,\infty]$ by*

$$p_{\alpha,\beta}(f) = \sup_{x\in\mathbb{R}^n} \left| x^\alpha \partial^\beta f(x) \right|, \quad \alpha, \beta \in \mathbb{N}^n \ (\textit{multi-index}). \tag{2.4.5}$$

We say $f \in \mathscr{S}$ or f is a Schwartz function, if $p_{\alpha,\beta}(f) < \infty$ for all $\alpha, \beta \in \mathbb{N}^n$. We call $\mathscr{S}$ the Schwartz class.

Example 5. $C_c^\infty(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$. $\exp\left(-\pi|x|^2\right) \in \mathscr{S}(\mathbb{R}^n)$.

- *We can use the countible familiy of seminorms $\{p_{\alpha,\beta}\}_{\alpha,\beta\in\mathbb{N}^n}$ to define a Frechét topology on $\mathscr{S}$ (meaning that it is metrizable using $d(f,g) = \sum_{\alpha,\beta\in\mathbb{N}^n} c_{\alpha,\beta} \frac{p_{\alpha,\beta}(f-g)}{1+p_{\alpha,\beta}(f-g)}$).*

  In practice, we use the analytic statement: a sequence ϕ_k converges to ϕ in $\mathscr{S}$ if and only if $p_{\alpha,\beta}(\phi_k - \phi) \to 0$ for all $\alpha, \beta \in \mathbb{N}^n$.

- *$\mathscr{S}(\mathbb{R}^n)$ is dense in every $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. [Thus it is natural for us to use $\mathscr{S}(\mathbb{R}^n)$ as a test function space in the study of L^p theory of Fourier transforms.]*

Definition 2.4.3 (tempered distributions)**.** *The dual space $\mathscr{S}^*(\mathbb{R}^n)$ (the space of continuous linear functionals) of $\mathscr{S}(\mathbb{R}^n)$ is called the space of tempered distributions. [This will be the largest space that we attempt to define Fourier transforms.]*

Example 6.
- *For* $1 \le p \le \infty$*, if* $g \in L^p(\mathbb{R}^n)$*, then* g *is in* $\mathscr{S}^*(\mathbb{R}^n)$*, acting on* f *by* $f \mapsto \int_{\mathbb{R}^n} f(x)g(x)\mathrm{d}x$*.*
- *The Dirac delta* δ *defined by* $\delta(f) := \langle \delta, f\rangle(\text{pairing}) = f(0)$ *is in* $\mathscr{S}^*(\mathbb{R}^n)$*.*
- *"The derivative of Dirac delta"* $\langle \frac{\partial}{\partial x_j}\delta, f\rangle := -\frac{\partial f}{\partial x_j}(0)$ *is in* $\mathscr{S}^*(\mathbb{R}^n)$*.*

Remark 9. *Consistent with the integration by parts formula:*

$$\left\langle \frac{\partial g}{\partial x_j}, f \right\rangle = \int_{\mathbb{R}^n} \frac{\partial g}{\partial x_j}(x) f(x)\mathrm{d}x = -\int_{\mathbb{R}^n} g(x)\frac{\partial f}{\partial x_j}(x)\mathrm{d}x, \quad \text{if} \quad f, g \in \mathscr{S}(\mathbb{R}^n). \tag{2.4.6}$$

Proof. Just check that for $T \in \mathscr{S}^*(\mathbb{R}^n)$, whenever $\{\phi_k\} \subset \mathscr{S}(\mathbb{R}^n)$ converges to 0, we have

$$\lim_{k\to\infty} \langle T, \phi_k\rangle = 0. \tag{2.4.7}$$

□

Theorem 2.4.4 (Fourier transform on Schwartz class)**.** *FT is a continuous linear operator from* $\mathscr{S}$ *to* $\mathscr{S}$*, satisfying*

$$\int_{\mathbb{R}^n} \widehat{f} g \mathrm{d}x = \int_{\mathbb{R}^n} f\widehat{g}\mathrm{d}x, \quad f, g \in \mathscr{S}(\mathbb{R}^n). \tag{2.4.8}$$

And also the inversion formula

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{2\pi \mathrm{i} x\cdot\xi}\mathrm{d}\xi, \quad f \in \mathscr{S}(\mathbb{R}^n). \tag{2.4.9}$$

Lemma 1 (Fourier transform of Gaussian)**.** *For* $a > 0$*,* $(\widehat{e^{-\pi|x|^2}})(\xi) = e^{-\pi|\xi|^2}$*.*

Proof. By Fubini's theorem, we have

$$(\widehat{e^{-\pi|x|^2}})(\xi) = \int_{\mathbb{R}^n} e^{-\pi|x|^2} e^{-2\pi \mathrm{i} x\cdot\xi}\mathrm{d}x = \prod_{j=1}^n \left(\int_{-\infty}^{\infty} e^{-\pi x_j^2} e^{-2\pi \mathrm{i} x_j \xi_j}\mathrm{d}x_j\right). \tag{2.4.10}$$

We only need to compute the one-dimensional integral. By completing the square, we have

$$\int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi \mathrm{i} x\xi}\mathrm{d}x = e^{-\pi\xi^2}\int_{-\infty}^{\infty} e^{-\pi(x+\mathrm{i}\xi)^2}\mathrm{d}x. \tag{2.4.11}$$

Since $e^{-\pi z^2}$ is an entire function, by Cauchy's integral theorem, we can shift the contour of integration from $\mathbb{R}$ to $\mathbb{R} + \mathrm{i}\xi$ without changing the value of the integral. Thus we have

$$\int_{-\infty}^{\infty} e^{-\pi(x+\mathrm{i}\xi)^2}\mathrm{d}x = \int_{-\infty}^{\infty} e^{-\pi t^2}\mathrm{d}t = 1. \tag{2.4.12}$$

Therefore, we conclude that $(\widehat{e^{-\pi x^2}})(\xi) = e^{-\pi\xi^2}$ and thus $(\widehat{e^{-\pi|x|^2}})(\xi) = e^{-\pi|\xi|^2}$. □

Proof of Theorem 2.4.4. For continuity, we note by Proposition 5 that

$$\xi^\alpha D^\beta \widehat{f}(\xi) = C_{\alpha,\beta}(\widehat{D^\alpha(x^\beta f)})(\xi). \tag{2.4.13}$$

Hence,
$$\left\|\xi^\alpha D^\beta \widehat{f}\right\|_\infty \le C_{\alpha,\beta}\left\|D^\alpha(x^\beta f)\right\|_1. \tag{2.4.14}$$
Also, by Hölder's inequality, we have
$$\|g\|_1 \lesssim \max\{\left\|x_1^{n+1}g\right\|_\infty, \cdots, \left\|x_n^{n+1}g\right\|_\infty, \|g\|_\infty\}. \tag{2.4.15}$$
Thus,
$$p_{\alpha,\beta}(\widehat{f}) = \left\|\xi^\alpha D^\beta \widehat{f}\right\|_\infty \lesssim \max_{|\gamma|,|\delta|\le n+1} p_{\gamma,\delta}(f). \tag{2.4.16}$$
By these we see that FT is a continuous linear operator from $\mathscr{S}$ to $\mathscr{S}$.

For the second part, by the Fubini's theorem, we have
$$\int_{\mathbb{R}^n}\widehat{f}(\xi)g(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(x)e^{-2\pi \mathrm{i}x\cdot\xi}\mathrm{d}x\right)g(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^n} f(x)\left(\int_{\mathbb{R}^n} g(\xi)e^{-2\pi \mathrm{i}x\cdot\xi}\mathrm{d}\xi\right)\mathrm{d}x = \int_{\mathbb{R}^n} f(x)\widehat{g}(x)\mathrm{d}x. \tag{2.4.17}$$
For the inversion formula, from the second part as well as Proposition 5, we have
$$\int f(x)\widehat{g}(\lambda x)\mathrm{d}x = \int \widehat{f}(x)g(\lambda^{-1}x)\lambda^{-n}\mathrm{d}x. \tag{2.4.18}$$
On the other hand, by the scaling property of FT, we have
$$\int f(x)\widehat{g}(\lambda x)\mathrm{d}x = \lambda^{-n}\int f(\lambda^{-1}x)\widehat{g}(x)\mathrm{d}x. \tag{2.4.19}$$
By DCT, when taking $\lambda \to \infty$, we have
$$f(0)\int_{\mathbb{R}^n}\widehat{g}(x)\mathrm{d}x = g(0)\int_{\mathbb{R}^n}\widehat{f}(x)\mathrm{d}x. \tag{2.4.20}$$
Now we take $g = \exp\left(-\pi|x|^2\right)$, then by the lemma we have $f(0) = \int_{\mathbb{R}^n}\widehat{f}(\xi)\mathrm{d}\xi$. By translation (Proposition 5), we conclude that for any $x \in \mathbb{R}^n$,
$$f(x) = (\tau_x f)(0) = \int_{\mathbb{R}^n}\widehat{(\tau_x f)}(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^n}\widehat{f}(\xi)e^{2\pi\mathrm{i}x\cdot\xi}\mathrm{d}\xi. \tag{2.4.21}$$
□

Corollary 2. *The Fourier transform operator* $\widehat{\ }$ *is* 4*-periodic. In fact,* $\widehat{\widehat{f}} = \widetilde{f}$ *where* $\widetilde{f}(x) = f(-x)$.

Once we have the FT for Schwartz functions (along with the very nice properties we established), we can extend the definition of FT to tempered distributions readily by duality.

Definition 2.4.5 (Fourier transform of tempered distributions)**.** *For* $T \in \mathscr{S}^*(\mathbb{R}^n)$*, we define its FT* $\widehat{T} \in \mathscr{S}^*(\mathbb{R}^n)$ *by its action on* $\phi \in \mathscr{S}(\mathbb{R}^n)$*:*
$$\langle\widehat{T},\phi\rangle := \langle T,\widehat{\phi}\rangle, \quad \forall\phi \in \mathscr{S}(\mathbb{R}^n). \tag{2.4.22}$$
Proof. The linearity of $\widehat{T}$ is obvious. The continuity follows from the continuity of T and FT on Schwartz class. □

Example 7. *We will see that L^p functions ($1 \le p \le \infty$) can all be viewed as tempered distributions. This is because for any sequence of Schwartz functions ϕ_k converging to 0, by Hölder's inequality we have*

$$\langle f, \phi_k \rangle = \int_{\mathbb{R}^n} f(x)\phi_k(x)\mathrm{d}x \le \|f\|_p \|\phi_k\|_{p'}. \tag{2.4.23}$$

We have $\|\phi_k\|_{p'}$ is dominated by functions with form $\|x^a\phi_k\|_\infty$, because

$$\|\phi_k\|_{p'}^{p'} = \int_{\mathbb{R}^n} |\phi_k(x)|^{p'}\mathrm{d}x = \int_{|x|\le 1} |\phi_k(x)|^{p'}\mathrm{d}x + \int_{|x|>1} |\phi_k(x)|^{p'}\mathrm{d}x \le \|\phi_k\|_\infty^{p'} vol(B_n) + C_n\|x^a\phi_k\|_\infty^{p'}, \tag{2.4.24}$$

and thus goes to 0 as $k \to \infty$. Therefore, $\langle f, \phi_k\rangle \to 0$ as $k \to \infty$, which implies that $f \in \mathscr{S}^*(\mathbb{R}^n)$.

Example 8 (A sanity check of the definition of FT for tempered distributions)**.** *If $f \in L^1(\mathbb{R}^n)$, then $f \in \mathscr{S}^*(\mathbb{R}^n)$ and $\widehat{f}$ defined here coincides with the usual FT of f. This is because for any $\phi \in \mathscr{S}(\mathbb{R}^n)$, we can repeat the computation in Theorem 2.4.4 to get*

$$\langle \widehat{f}, \phi\rangle = \langle f, \widehat{\phi}\rangle = \int_{\mathbb{R}^n} f(x)\left(\int_{\mathbb{R}^n} \phi(\xi)e^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{d}\xi\right)\mathrm{d}x = \int_{\mathbb{R}^n} \phi(\xi)\left(\int_{\mathbb{R}^n} f(x)e^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{d}x\right)\mathrm{d}\xi \overset{usual\ defn}{=} \int_{\mathbb{R}^n}\phi(\xi)\widehat{f}(\xi)\mathrm{d}\xi. \tag{2.4.25}$$

Example 9. *For Borel measure μ (a bounded linear functional on $C_c(\mathbb{R}^n)$), we can view μ as a tempered distribution by restricting its action to $\mathscr{S}(\mathbb{R}^n) \subset C_c(\mathbb{R}^n)$. Then its FT is given by*

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{d}\mu(x). \tag{2.4.26}$$

δ is the Dirac measure at the origin, this gives $\widehat{\delta}(\xi) = 1$.

Theorem 2.4.6 (The Fourier transform of tempered distributions is a continuous linear bijection)**.** *The Fourier transform $\widehat{\ }: \mathscr{S}^*(\mathbb{R}^n) \to \mathscr{S}^*(\mathbb{R}^n)$ is a continuous linear bijection, whose inverse is also continuous.*

Proof. If $T_n \to T$ in $\mathscr{S}^*(\mathbb{R}^n)$, then for any $\phi \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\widehat{T_n}(\phi) = T_n(\widehat{\phi}) \to T(\widehat{\phi}) = \widehat{T}(\phi). \tag{2.4.27}$$

Furthermore, since $\widehat{\ }$ is 4-periodic, thus its inverse is equivalent to acting by $\widehat{\ }$ three times, which is also continuous. □

We are at the position to study the L^p theory of Fourier transforms. We begin with L^2 which has a very nice Plancherel theorem.

Theorem 2.4.7 (Plancherel theorem)**.** *The Fourier transform is an isometry on $L^2(\mathbb{R}^n)$, i.e.*

$$\left\|\widehat{f}\right\|_2 = \|f\|_2, \quad \forall f \in L^2(\mathbb{R}^n). \tag{2.4.28}$$

Proof. Given $f, g \in \mathscr{S}$, let $g := \overline{\widehat{h}}$, thus we have $\widehat{g} = \overline{h}$. By the property of FT on Schwartz class (Theorem 2.4.4), we have

$$\int_{\mathbb{R}^n} \widehat{f}(\xi)\overline{\widehat{h}(\xi)}\mathrm{d}\xi = \int_{\mathbb{R}^n} f(x)\overline{h(x)}\mathrm{d}x. \tag{2.4.29}$$

Since $\mathscr{S} \supset C_c^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, we conclude that for any L^2 functions f, h, we have

$$(f, h) = (\widehat{f}, \widehat{h}). \tag{2.4.30}$$

By taking $f = h$, we conclude the Plancherel theorem. □

Remark 10. *Since $f\chi_{B(0,R)}$ and $\widehat{f\chi_{B(0,R)}}$ converge to f and $\widehat{f}$ in L^2 respectively as $R \to \infty$, by Plancherel theorem and the continuity of FT, we can actually define*

$$\widehat{f}(\xi) = \lim_{R\to\infty} \int_{|x|\le R} f(x) e^{-2\pi i x\cdot\xi} \mathrm{d}x, \quad \text{in } L^2(\mathbb{R}^n) \tag{2.4.31}$$

and

$$f(x) = \lim_{R\to\infty} \int_{|\xi|\le R} \widehat{f}(\xi) e^{2\pi i x\cdot\xi} \mathrm{d}\xi, \quad \text{in } L^2(\mathbb{R}^n). \tag{2.4.32}$$

Therefore, for L^2 functions f, $\widehat{f}$ can be understood in the way of eq. (2.4.31), which is a well-defined function (not just a tempered distribution).

Definition 2.4.8 (FT of L^2 function, in the sense of eq. (2.4.31))**.** *For $f \in L^2(\mathbb{R}^n)$, we define its FT $\widehat{f}$ by eq. (2.4.31).*

$$\widehat{f}(\xi) = \lim_{R\to\infty} \int_{|x|\le R} f(x) e^{-2\pi i x\cdot\xi} \mathrm{d}x. \tag{2.4.33}$$

It is well-defined and has Fourier inversion formula by the remark.

Proposition 6 (FT is defined as a function for $L^p, 1 \le p \le 2$ via decomposition)**.** *Once we have the theory for L^2, we can decompose any L^p $(1 < p < 2)$ function to $f_1 \in L^1$ and $f_2 \in L^2$, for example $f_1 = f\chi_{\{x:|f(x)|>1\}}$ and $f_2 = f - f_1$, we have*

$$\int |f_1| = \int_{\{x:|f(x)|>1\}} |f(x)|\mathrm{d}x \le \int_{\{x:|f(x)|>1\}} |f(x)|^p \mathrm{d}x \le \|f\|_p^p < \infty, \tag{2.4.34}$$

$$\int |f_2|^2 = \int_{\{x:|f(x)|\le 1\}} |f(x)|^2\mathrm{d}x \le \int_{\{x:|f(x)|\le 1\}} |f(x)|^p \mathrm{d}x \le \|f\|_p^p < \infty. \tag{2.4.35}$$

Therefore, $\widehat{f} = \widehat{f_1} + \widehat{f_2} \in L^\infty + L^2$, and thus is at least a ***well-defined function****.*

However, by applying an interpolation theorem, we can get a more refined result.

Theorem 2.4.9 (Riesz-Thorin interpolation theorem)**.** *Let $1 \le p_0, p_1, q_0, q_1 \le \infty$, $\theta \in (0,1)$. Define p and q by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{2.4.36}$$

If T is a linear operator from $L^{p_0} + L^{p_1}$ to $L^{q_0} + L^{q_1}$ such that

$$\|Tf\|_{q_0} \le M_0\|f\|_{p_0}, \quad \forall f \in L^{p_0}, \quad \|Tf\|_{q_1} \le M_1\|f\|_{p_1}, \quad \forall f \in L^{p_1}, \tag{2.4.37}$$

then T is a bounded linear operator from L^p to L^q and

$$\|Tf\|_q \le M_0^{1-\theta} M_1^{\theta} \|f\|_p, \quad \forall f \in L^p. \tag{2.4.38}$$

Proof. The proof is based on the Hadamard three-lines theorem in complex analysis. □

We have two immediate corollaries of Riesz-Thorin interpolation theorem: **Hausdorff-Young inequality** and **Young's inequality for convolution**.

Corollary 3 (Hausdorff-Young inequality)**.** *For $1 \leq p \leq 2$, if $f \in L^p(\mathbb{R}^n)$, then $\widehat{f} \in L^{p'}(\mathbb{R}^n)$ and*

$$\left\|\widehat{f}\right\|_{p'} \leq \|f\|_p. \tag{2.4.39}$$

Here, p' is the Hölder conjugate of p defined by $\frac{1}{p} + \frac{1}{p'} = 1$.

Proof. By the L^1-FT theory we have $\left\|\widehat{f}\right\|_\infty \leq \|f\|_1$. By the Plancherel theorem, we have $\left\|\widehat{f}\right\|_2 = \|f\|_2$. Thus we can apply the Riesz-Thorin interpolation theorem with $p_0 = 1, p_1 = q_0 = q_1 = 2$ and θ such that $\frac{1}{p} = 1 - \theta + \frac{\theta}{2}$, i.e. $\theta = 2 - \frac{2}{p}$, to conclude that $\widehat{\ }: L^1 + L^2 \to L^\infty + L^2$ is a bounded linear operator from L^p to $L^{p'}$ with operator norm at most 1. □

Remark 11. *To sum up, we have that $\widehat{f}$ can be understood as a tempered distribution for any $f \in L^p, 1 \leq p \leq \infty$; $\widehat{f}$ is a well-defined function for any $f \in L^p, 1 \leq p \leq 2$. Specifically, for $p = 1$ we can understand it in the usual Lebesgue sense; in the case $p = 2$ we can understand it in the sense of eq. (2.4.31) using Plancherel; for $1 < p < 2$ we can understand it as a sum of an L^2 function and an L^∞ function. Furthermore, by the Hausdorff-Young inequality, we have $\widehat{f} \in L^{p'}$ for $1 \leq p \leq 2$.*

Remark 12. *Riesz-Thorin interpolation theorem is a very useful tool in analysis. For example, we have:*

Corollary 4 (Young's inequality for convolution)**.** *For $1 \leq p, q, r \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, if $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ and*

$$\|f * g\|_r \leq \|f\|_p \|g\|_q. \tag{2.4.40}$$

Proof. For fixed $f \in L^p$, we have

$$\|f * g\|_p = \left\|\int f(\cdot - y) g(y) \mathrm{d}y\right\|_p \overset{\text{Minkowski}}{\leq} \int \|f(\cdot - y)\|_p |g(y)| \mathrm{d}y = \|f\|_p \|g\|_1. \tag{2.4.41}$$

$$|(f * g)(x)| = \left|\int f(x - y) g(y) \mathrm{d}y\right| \leq \|f(x - \cdot) g(\cdot)\|_1 \overset{\text{Hölder}}{\leq} \|f\|_p \|g\|_{p'}. \tag{2.4.42}$$

Then the result follows by Riesz-Thorin. □

Example 10. *If $f \in L^1$, $g \in L^2$, then we have*

$$\|f * g\|_2 \leq \|f\|_1 \|g\|_2. \tag{2.4.43}$$

This means that $g \mapsto f * g$ is a bounded linear operator on L^2. Note that FT is also a bounded linear operator on L^2. Thus $g \mapsto \widehat{(f * g)}$ is also a bounded linear operator on L^2. But note that $f \in L^1$ implies that $\widehat{f}$ is well-defined and is L^∞, thus $g \mapsto \widehat{f} \cdot \widehat{g}$ is also a bounded linear operator on L^2. But by Proposition 5 we know that these two agree on $L^1 \cap L^2$, which is dense in L^2 (for example, L^2 functions can be approximated in norm by linear combinations of characteristic functions with compact support). Therefore, by continuity, we conclude that

$$\widehat{(f * g)} = \widehat{f} \cdot \widehat{g}, \quad \textit{for all } g \in L^2. \tag{2.4.44}$$

That is, we can extend the convolution theorem to the case $f \in L^1, g \in L^2$.

2.5 Summability methods for Fourier inversion

The Fourier inversion formula is also trying to reproduce the original function f from its Fourier transform $\widehat{f}$, analogous to the Fourier series:

$$\lim_{R\to\infty}\int_{B_R}\widehat{f}(\xi)e^{2\pi i\cdot\xi}d\xi \sim f(x). \tag{2.5.1}$$

We already see that this limit holds for L^2 function in L^2 sense. What about other L^p and almost everywhere sense? To study this, we define

$$S_Rf(x) := \int_{|\xi|\le R}\widehat{f}(\xi)e^{2\pi i x\cdot\xi}d\xi. \tag{2.5.2}$$

Then the question is: whether we have $S_Rf(x) \to f(x)$ as $R \to \infty$ in some sense.

Lemma 2. *For $1 \le p \le \infty$, if $\|S_Rf\|_p \le C_p\|f\|_p$ for some $C_p > 0$ and all $R > 0$, then $S_Rf \to f$ in L^p as $R \to \infty$.*

Proof. Analogous to the proof in the Fourier series case. □

Remark 13. *For $n = 1$, this equality holds for $p > 1$ but fails for $p = 1$;*
For $n > 1$, interestingly, this equality fails for all $p \neq 1$ but $p = 2$.

Remark 14. *By rescaling, we have $\|S_R\|_{p\to p} = \|S_1\|_{p\to p}$ for all $R > 0$.*

Definition 2.5.1 (Dirichlet kernel)**.** *We set $n = 1$. The Dirichlet kernel D_R is defined by*

$$D_R(x) = \int_{-R}^{R} e^{2\pi i x\xi}d\xi = \frac{\sin(2\pi Rx)}{\pi x}. \tag{2.5.3}$$

Then we have $S_Rf = f * D_R$.

Remark 15. *Sanity check: $D_R(x)$ is clearly not integrable, but it is in L^q for any $q > 1$. Therefore $D_R * f$ is still well-defined for $f \in L^p, 1 < p < \infty$.*

The method of Cesàro summation consists in taking the average of the partial sums. In our case, we define

$$\sigma_Rf(x) := \frac{1}{R}\int_0^R S_\rho f(x)d\rho. \tag{2.5.4}$$

For $n = 1$, we have

$$\sigma_Rf = f * F_R, \quad F_R(x) = \frac{1}{R}\int_0^R D_\rho(x)d\rho = \frac{1}{R}\int_0^R \frac{\sin(2\pi\rho x)}{\pi x}d\rho = \frac{1-\cos(2\pi Rx)}{2\pi^2Rx^2}. \tag{2.5.5}$$

F_R is called the Fejér kernel. Unlike D_R, F_R is nonnegative and integrable with $\|F_R\|_1 = 1$ by direct computation. Thus, we can expect better convergence results for σ_Rf than S_Rf.

Theorem 2.5.2. *For $1 \le p < \infty$, if $f \in L^p(\mathbb{R})$, then $\sigma_Rf \to f$ in L^p as $R \to \infty$.*

In the next chapter we will prove more general results from which we deduce convergence in L^p and almost everywhere for this and other summability methods.

Example 11. • ***Abel-Poisson summation***: *Define*

$$u(x,t) = \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{2\pi \mathrm{i} x\cdot\xi} e^{-2\pi|t|\xi} \mathrm{d}\xi. \tag{2.5.6}$$

We have

$$u(x,t) = f * P_t(x), \quad P_t(x) = \frac{\Gamma(\frac{n+1}{2})t}{\pi^{\frac{n+1}{2}}(t^2+x^2)^{\frac{n+1}{2}}}. \tag{2.5.7}$$

$u(x,t)$ is harmonic in $\mathbb{R}^n \times (0,\infty)$. For $n=1$, we have

$$u(z) = u(x+\mathrm{i}t) = \int_0^\infty \widehat{f}(\xi) e^{2\pi \mathrm{i} z\xi} \mathrm{d}\xi + \int_{-\infty}^0 \widehat{f}(\xi) e^{2\pi \mathrm{i}\bar{z}\xi} \mathrm{d}\xi. \tag{2.5.8}$$

$u(x,t)$ solves the Dirichlet problem in the upper half plane with boundary data f.

$$\begin{cases} \triangle u = 0, & \text{in } \mathbb{R}\times(0,\infty), \\ u(x,0) = f(x), & x \in \mathbb{R}. \end{cases} \tag{2.5.9}$$

• ***Gauss-Weierstrass summation***: *Define*

$$w(x,t) = \int_{\mathbb{R}^n} \widehat{f}(\xi) e^{2\pi \mathrm{i} x\cdot\xi} e^{-\pi t^2|\xi|^2} \mathrm{d}\xi. \tag{2.5.10}$$

We have

$$w(x,t) = f * G_t(x), \quad G_t(x) = \frac{1}{t^n} e^{-\pi|x/t|^2}. \tag{2.5.11}$$

The function $\widetilde{w}(x,t) := w(x, \sqrt{4\pi t})$ solves the heat equation $\begin{cases} \frac{\partial \widetilde{w}}{\partial t} = \triangle \widetilde{w}, & \text{in } \mathbb{R}^n \times (0,\infty), \\ \widetilde{w}(x,0) = f(x), & x \in \mathbb{R}^n. \end{cases}$

Chapter 3

The Hardy-Littlewood maximal function

3.1 Approximations to the identity

Recall that when we discuss the summability methods, we use the Fejér, Poisson and Gauss-Weierstrass kernels to *try* to reproduce the original function f. In fact this technique can be generalized and is very useful to study the pointwise convergence of Fourier transforms.

Definition 3.1.1. *Let ϕ be an L^1 function on $\mathbb{R}^n$ such that $\int \phi = 1$. For any $t > 0$ we define*

$$\phi_t(x) := t^{-n}\phi(t^{-1}x). \tag{3.1.1}$$

We can understand ϕ_t as a tempered distribution and consider the pairing for a test function $g \in \mathscr{S}(\mathbb{R}^n)$:

$$\langle \phi_t, g\rangle = \int_{\mathbb{R}^n} \phi_t(x)g(x)\mathrm{d}x = \int_{\mathbb{R}^n} t^{-n}\phi(t^{-1}x)g(x)\mathrm{d}x = \int_{\mathbb{R}^n} \phi(y)g(ty)\mathrm{d}y. \tag{3.1.2}$$

And by DCT,

$$\lim_{t\to 0}\langle \phi_t, g\rangle = g(0)\int_{\mathbb{R}^n} \phi(y)\mathrm{d}y = g(0). \tag{3.1.3}$$

Or,

$$\lim_{t\to 0}\langle \phi_t, g\rangle = \langle \delta, g\rangle \quad \text{or} \quad \lim_{t\to 0}\phi_t * g(x) = g(x). \tag{3.1.4}$$

*Since ϕ_t tries to approximate the convolutional identity δ, we call $\{\phi_t\}_{t>0}$ an **approximation to the identity**.*

Using exactly the same argument as we have done for summability methods, we can prove the following theorem.

Theorem 3.1.2. *Let $\{\phi_t\}_{t>0}$ be an approximation to the identity. Then we have the L^p convergence:*

$$\lim_{t\to 0}\|\phi_t * f - f\|_p = 0, \quad 1 \le p < \infty, \quad f \in L^p(\mathbb{R}^n). \tag{3.1.5}$$

And we have the uniform convergence if $f \in C_c(\mathbb{R}^n)$ when $p = \infty$.

Proof. Because ϕ has unit integral, we have

$$\phi_t * f(x) - f(x) = \int_{\mathbb{R}^n} \phi_t(y)(f(x-ty) - f(x))\mathrm{d}y. \tag{3.1.6}$$

Given $\varepsilon > 0$, choose $\delta > 0$ (TBD), such that for any t sufficiently small (TBD), we have by Minkowski's inequality

$$\|\phi_t * f - f\|_p \le \int_{|y|<\delta/t} |\phi(y)| \|f(\cdot - ty) - f(\cdot)\|_p \mathrm{d}y + 2\|f\|_p \int_{|y|\ge\delta/t} |\phi(y)| \mathrm{d}y < \varepsilon. \tag{3.1.7}$$

To this end, we should choose δ such that for any $|h| < \delta$, $\|f(\cdot + h) - f(\cdot)\|_p < \frac{\varepsilon}{2\|\phi\|_1}$ (using the continuity of Lebesgue integral) and t sufficiently small such that $\int_{|y|\ge\delta/t} |\phi(y)| \mathrm{d}y < \frac{\varepsilon}{4\|f\|_p}$. Then we have $\|\phi_t * f - f\|_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. ◻

3.2 Weak-type inequalities and almost everywhere convergence

The motivation of weak-type inequalities is actually from real analysis. Consider an operator $T : L^p \to L^q$, our dream is to have a bounded linear operator, i.e. $\|Tf\|_q \lesssim \|f\|_p$. However, this is not always possible. A good compromise is to have a weak-type inequality, which is a level set estimate, i.e. for any $\alpha > 0$,

$$m(\{x : |Tf(x)| > \alpha\}) \lesssim \left(\frac{\|f\|_p}{\alpha}\right)^q. \tag{3.2.1}$$

This is weaker than the strong-type inequality (follows from Chebyshev's inequality), but still very useful in applications.

The serious definition is as follows.

Definition 3.2.1 (Weak-(p, q) inequality)**.** *Let (X, μ) and (Y, ν) be two measure spaces, T is an operator from $L^p(X, \mu)$ to $\{$measurable functions on $Y\}$. We say T is weak-(p, q) if there exists a constant $C > 0$ such that for any $f \in L^p(X, \mu)$ and any $\alpha > 0$, we have*

$$\nu(\{y \in Y : |Tf(y)| > \alpha\}) \le \left(\frac{C\|f\|_p}{\alpha}\right)^q. \tag{3.2.2}$$

As a comparison, if $Tf \in L^q$ $(q < \infty)$ and $\exists C > 0$ such that $\|Tf\|_q \le C\|f\|_p$, then we say T is strong-(p, q).

We call T is weak-(p, ∞), if T is strong-(p, ∞). We will see later that this convetion will be convenient in most of the statement and proof.

The weak-(p, q) convergence is useful for studying the almost everywhere convergence of operators. This follows from the theory of maximal functions. (A celebrated example is the Lebesgue differentiation theorem.)

Theorem 3.2.2. *Let $\mathcal{A}$ be a (continuous) index set $\subset \mathbb{R}$, $t_0 \in \mathbb{R}$. Let $\{T_t\}_{t\in\mathcal{A}}$ be a family of* ***linear or sublinear*** *operators (e.g. convolution operators) from $L^p(X, \mu)$ to $L^p(X, \mu)$. We define the maximal function by*

$$T^* f(x) := \sup_{t\in\mathcal{A}} |T_t f(x)|. \tag{3.2.3}$$

We can easily prove that, the maximal operator T^* is also sublinear in this case.

If T^* is weak-(p, q) for some $q \ge 1$ (usually $p = q$), then for any $f \in L^p(X, \mu)$, we have the set

$$\{f \in L^p(X, \mu) : \lim_{\mathcal{A}\ni t\to t_0} T_t f(x) = f(x) \text{ for a.e. } x \in X\} \tag{3.2.4}$$

is closed in $L^p(X, \mu)$.

Remark 16. *This theorem suggests that we can extend the almost everywhere convergence from some "nice function" to a large class of functions using the closedness of the set above.*

Remark 17. *We say X is a sublinear operator on a normed vector space if*

$$\|T(f_1 + f_2)\| \leq \|Tf_1\| + \|Tf_2\|, \quad \|T(\lambda f)\| = |\lambda|\|Tf\|. \tag{3.2.5}$$

Proof. Assuming $f_n \to f$ in L^p, such that for every n, $\lim_{t\to t_0} T_t f_n(x) = f_n(x)$ for a.e. $x \in X$. The standard technique in real analysis is to consider the exceptional set

$$W_\lambda := \mu\left(\{x \in X : \limsup_{\mathcal{A}\ni t\to t_0} |T_t f(x) - f(x)| > \lambda\}\right), \quad \lambda > 0. \tag{3.2.6}$$

Assume WLOG $q < \infty$, then for every n, we have

$$\begin{aligned}
W_\lambda &\leq \mu\left(\{x \in X : \limsup_{\mathcal{A}\ni t\to t_0} |T_t(f - f_n)(x) - (f - f_n)(x)| > \lambda\}\right) \\
&\leq \mu\left(\{x \in X : \limsup_{\mathcal{A}\ni t\to t_0} |T_t(f - f_n)(x)| > \lambda/2\} \cup \{x \in X : |(f - f_n)(x)| > \lambda/2\}\right) \\
&\leq \mu\left(\{x \in X : \limsup_{\mathcal{A}\ni t\to t_0} |T_t(f - f_n)(x)| > \lambda/2\}\right) + \mu\left(\{x \in X : |(f - f_n)(x)| > \lambda/2\}\right) \\
&\overset{\text{defn of } T^*}{\leq} \mu\left(\{x \in X : T^*(f - f_n)(x) > \lambda/2\}\right) + \mu\left(\{x \in X : |(f - f_n)(x)| > \lambda/2\}\right) \\
&\overset{\text{weak-}(p,q)\text{ of } T^* \text{ and Chebyshev}}{\leq} \left(\frac{\|f - f_n\|_p}{\lambda}\right)^q + \left(\frac{\|f - f_n\|_p}{\lambda}\right)^q.
\end{aligned} \tag{3.2.7}$$

Taking $n \to \infty$, we have $W_\lambda = 0$ for any $\lambda > 0$, which implies that

$$\mu(\{x \in X : \limsup_{\mathcal{A}\ni t\to t_0} |T_t f(x) - f(x)| > 0\}) \leq \sum_{k=1}^{\infty} W_{1/k} = 0. \tag{3.2.8}$$

□

Remark 18. *Similarly, $\{f \in L^p(X,\mu) : \lim_{\mathcal{A}\ni t\to t_0} T_t f(x) = f(x)$ exists for a.e. $x \in X\}$ is closed in $L^p(X,\mu)$.*

Remark 19. *If $T_t f$ is only defined a.e., do we have the risk that when taking the supremum over $t \in \mathcal{A}$ in the definition of $T^* f$, we may end up with something blowing up?*

In fact, in many applications, $T_t f$ indeed has a canonical definition everywhere, and $T^* f$ will still satisfy the weak type bound.

Proposition 7. *By Theorem 3.1.2, if*

$$\Phi^* f(x) := \sup_{t>0} |\phi_t * f(x)| \tag{3.2.9}$$

is weak-(p,p), then we know that $\lim_{t\to 0} \phi_t * f(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$ for any $f \in L^p(\mathbb{R}^n)$.

3.3 Marcinkiewicz interpolation theorem

Our dream is to prove the weak-(p, p) bound for the maximal function Φ^* of the family of apporimations to the identity $\{\phi_t\}_{t>0}$. A useful tool to help us get focused is the Marcinkiewicz interpolation theorem, which is an analogue of the Riesz-Thorin interpolation theorem for weak-type inequalities.

Theorem 3.3.1 (Marcinkiewicz interpolation theorem)**.** *Let $1 \leq p_0, p_1 \leq \infty$, $\theta \in (0, 1)$. Define p by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}. \tag{3.3.1}$$

If T is a sublinear operator from $L^{p_0} + L^{p_1}$ to the measurable functions on Y such that

$$T \text{ is weak-}(p_0, p_0) \text{ as well as } (p_1, p_1), \tag{3.3.2}$$

then T is strong-(p, p), i.e.

$$\|Tf\|_p \leq C_{p,p_0,p_1}\|f\|_p, \quad \forall f \in L^p. \tag{3.3.3}$$

Remark 20. *The result seems a bit strong, in the sense that we only assume weak type bounds at the endpoints, but we can get strong type bound everywhere in between.*

Remark 21. *The Riesz-Thorin interpolation theorem allows us to interpolate among analytic families of operators, while the Marcinkiewicz interpolation theorem allows us to interpolate among sublinear operators and only need weak-type bounds at the endpoints.*

Lemma 3. *$\phi : [0, \infty) \to [0, \infty)$ increasing, differentiable, $\phi(0) = 0$. Then*

$$\int_X \phi(|f(x)|)\mathrm{d}\mu(x) = \int_0^\infty \phi'(\alpha)\mu(\{x \in X : |f(x)| > \alpha\})\mathrm{d}\alpha. \tag{3.3.4}$$

Proof. By Fubini's theorem, we have

$$\begin{aligned}\int_X \phi(|f(x)|)\mathrm{d}\mu(x) &= \int_X \int_0^{|f(x)|} \phi'(\alpha)\mathrm{d}\alpha\mathrm{d}\mu(x)\\ &= \int_0^\infty \phi'(\alpha) \int_{\{x\in X:|f(x)|>\alpha\}} 1\mathrm{d}\mu(x)\mathrm{d}\alpha\\ &= \int_0^\infty \phi'(\alpha)\mu(\{x \in X : |f(x)| > \alpha\})\mathrm{d}\alpha.\end{aligned} \tag{3.3.5}$$

□

Remark 22. *In particular, if we take $\phi(\alpha) = \alpha^p$, then $\phi'(\alpha) = p\alpha^{p-1}$ and*

$$\|f\|_p^p = \int_0^\infty p\alpha^{p-1}\mu(\{x \in X : |f(x)| > \alpha\})\mathrm{d}\alpha. \tag{3.3.6}$$

Proof of Theorem 3.3.1. The idea is to use the superlevel set estimate to control the L^p norm. Given $f \in L^p$, we decompose it into two parts:

$$f_0 = f \cdot \chi_{\{|f|>\alpha\}}, \quad f_1 = f \cdot \chi_{\{|f|\leq\alpha\}}, \quad f_0 + f_1 = f. \tag{3.3.7}$$

Then we have by elementary estimation:

$$\int_{\{|f|>\alpha\}} |f(x)|^{p_0} \mathrm{d}\mu(x) \overset{\alpha<|f|,\, p_0\le p}{\le} \alpha^{p_0-p} \int_{\{|f|>\alpha\}} |f(x)|^{p} \mathrm{d}\mu(x) \le \alpha^{p_0-p}\|f\|_p^p < \infty \Rightarrow f_0 \in L^{p_0}, \tag{3.3.8}$$

$$\|f_0\|_{p_0}^{p_0} = \int_{\{|f|>\alpha\}} |f(x)|^{p_0} \mathrm{d}\mu(x). \tag{3.3.9}$$

and similarly

$$f_1 \overset{\alpha\ge|f|,\, p_1\ge p}{\in} L^{p_1}, \quad \text{and} \quad \|f_1\|_{p_1}^{p_1} = \int_{\{|f|\le\alpha\}} |f(x)|^{p_1} \mathrm{d}\mu(x). \tag{3.3.10}$$

Therefore, we can apply the weak type bounds,

$$\mu(\{x \in X : |Tf_0(x)| > \alpha\}) \lesssim \left(\frac{\|f_0\|_{p_0}}{\alpha}\right)^{p_0}, \quad \mu(\{x \in X : |Tf_1(x)| > \alpha\}) \lesssim \left(\frac{\|f_1\|_{p_1}}{\alpha}\right)^{p_1}. \tag{3.3.11}$$

Now we have

$$\begin{aligned}
\|Tf\|_p^p &= \int_0^\infty p\alpha^{p-1}\mu(\{x\in X : |Tf(x)|>\alpha\})\mathrm{d}\alpha \\
&\lesssim \int_0^\infty p\alpha^{p-1}\mu(\{x\in X : |Tf_0(x)|>\alpha/2\})\mathrm{d}\alpha + \int_0^\infty p\alpha^{p-1}\mu(\{x\in X : |Tf_1(x)|>\alpha/2\})\mathrm{d}\alpha \\
&\overset{\text{eq. (3.3.11)}}{\lesssim} \int_0^\infty p\alpha^{p-1}\left(\frac{\|f_0\|_{p_0}}{\alpha}\right)^{p_0}\mathrm{d}\alpha + \int_0^\infty p\alpha^{p-1}\left(\frac{\|f_1\|_{p_1}}{\alpha}\right)^{p_1}\mathrm{d}\alpha \\
&\overset{\text{eqs. (3.3.9) and (3.3.10)}}{=} \int_0^\infty p\alpha^{p-1-p_0}\left(\int_{\{|f|>\alpha\}}|f(x)|^{p_0}\mathrm{d}\mu(x)\right)\mathrm{d}\alpha + \int_0^\infty p\alpha^{p-1-p_1}\left(\int_{\{|f|\le\alpha\}}|f(x)|^{p_1}\mathrm{d}\mu(x)\right)\mathrm{d}\alpha \\
&\overset{\text{Fubini}}{=} \int_X |f(x)|^{p_0}\int_0^{|f(x)|} p\alpha^{p-1-p_0}\mathrm{d}\alpha\mathrm{d}\mu(x) + \int_X |f(x)|^{p_1}\int_{|f(x)|}^\infty p\alpha^{p-1-p_1}\mathrm{d}\alpha\mathrm{d}\mu(x) \\
&\overset{\text{compute explicitly}}{=} \frac{p}{p_0-p}\int_X |f(x)|^p\mathrm{d}\mu(x) + \frac{p}{p_1-p}\int_X |f(x)|^p\mathrm{d}\mu(x) \lesssim \|f\|_p^p.
\end{aligned} \tag{3.3.12}$$

□

3.4 The Hardy-Littlewood maximal function

Definition 3.4.1 (Hardy-Littlewood maximal function)**.** *Let $f \in L^1_{loc}(\mathbb{R}^n)$ be a locally integrable function on $\mathbb{R}^n$. The Hardy-Littlewood maximal function Mf is defined by*

$$Mf(x) = \sup_{r>0} \frac{1}{m(B_r)} \int_{B_r} |f(x-y)|\mathrm{d}y = \sup_{r>0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(y)|\mathrm{d}y. \tag{3.4.1}$$

Here, $m(\cdot)$ is the Lebesgue measure, $B_r(x)$ is the ball of radius r centered at x and $B_r = B_r(0)$.

Remark 23. *Though the supremum is taken over uncountably many $r > 0$, but it is a continuous familiy which means that Mf is still a measurable function. This is because (e.g.) we can restrict the supremum to rational $r > 0$.*

Definition 3.4.2 (Related functions)**.** *Define the cube $Q_r(x) := \prod_{i=1}^n [x_i - r, x_i + r]$ and $Q_r = Q_r(0)$. We can similarly define the centered cube maximal function by*

$$M' f(x) = \sup_{r>0} \frac{1}{m(Q_r)} \int_{Q_r} |f(x-y)| \mathrm{d}y = \sup_{r>0} \frac{1}{m(Q_r(x))} \int_{Q_r(x)} |f(y)| \mathrm{d}y. \tag{3.4.2}$$

By converging arguments, we can see that M' and M are equivalent up to a constant depending on n.

We can also define the non-centered maximal function by

$$M'' f(x) = \sup_{Q \ni x,\ Q \text{ cube with edges parallel to the axes}} \frac{1}{m(Q)} \int_Q |f(z)| \mathrm{d}z. \tag{3.4.3}$$

Pointwisely, we have

$$Mf \sim_n M'f \sim_n M''f. \tag{3.4.4}$$

It is often more convenient to work with M.

Theorem 3.4.3. *M is weak-$(1,1)$ and strong-(p,p) for $1 < p \leq \infty$. ($p = \infty$ immediate, and the whole theorem follows from interpolation from $p = 1$, which will be proved in later chapters.)*

Proof. The proof for $n = 1$ is elementary and we will do it here. For $n > 1$, we will use a finite version of the Vitali 3-ball-covering lemma to prove it in later chapters.

As we have mentioned, it suffices to prove the weak-$(1,1)$ bound and the whole theorem follows from Theorem 3.3.1. WLOG suppose $\|f\|_1 = 1$. For any $\lambda > 0$, we need to bound the measure of the superlevel set

$$E_\lambda := m(\{x \in \mathbb{R} : \ Mf(x) > \lambda\}). \tag{3.4.5}$$

For $x \in E_\lambda$, by definition of $Mf(x)$, there exists interval I_x such that

$$\frac{1}{m(I_x)} \int_{I_x} |f(y)| \mathrm{d}y > \lambda. \tag{3.4.6}$$

We are going to find a more efficient subcovering of $\{I_x\}_{x \in E_\lambda}$ to estimate E_λ.

Lemma 4. *For any compact set $K \subset \mathbb{R}$, and any covering of K by a familiy of intervals $\{I_\alpha\}_{\alpha \in \mathcal{A}}$, there exists a subcovering $\{I_j\}_{j=1}^N$ such that*

$$K \subset \bigcup_{j=1}^N I_j, \quad \sum_{j=1}^N \chi_{I_j}(x) \leq 2 \quad \text{for any } x \in \mathbb{R}. \tag{3.4.7}$$

That is, any point occurs in at most two intervals. (First take a finite subcovering, then for any point covered by more than two intervals, we can remove all of them except the leftmost and rightmost ones.)

We go back to estimate E_λ. For any compact subset $K \subset E_\lambda$, we have

$$m(K) \leq \sum_{j=1}^N m(I_j) < \frac{1}{\lambda} \sum_{j=1}^N \int_{I_j} |f(y)| \mathrm{d}y \overset{\sum_{j=1}^N \chi_{I_j}(x) \leq 2}{\leq} \frac{2}{\lambda} \|f\|_1 = \frac{2}{\lambda}. \tag{3.4.8}$$

Since $K \subset E_\lambda$ is arbitrary, we have $m(E_\lambda) \leq \frac{2}{\lambda}$. This completes the proof. $\square$

Proposition 8. *Let ϕ be a positive radial decreasing (in $r \in (0,\infty)$) function on $\mathbb{R}^n$. Then for any $f \in L^1_{loc}(\mathbb{R}^n)$, we have*

$$\sup_{t>0} |f * \phi_t(x)| \leq \|\phi\|_1 Mf(x). \tag{3.4.9}$$

Remark 24. *The correspondence between the "central average" and the maximal function is intuitive, in the sense that if we decompose ϕ into "layers of balls", then the convolution with each layer is controlled by the maximal function.*

Proof. For $\phi = \sum_{j=1}^N a_j \chi_{B_{r_j}}$, $a_j \geq 0, r_j > 0$, we have

$$\phi_t * f(x) = \sum_{j=1}^N a_j m(B_{r_j}) \frac{1}{m(B_{tr_j})} \chi_{B_{tr_j}} * f(x) \leq \left(\sum_{j=1}^N a_j m(B_{r_j}) \right) Mf(x) = \|\phi\|_1 Mf(x). \tag{3.4.10}$$

We can approximate general radial decreasing ϕ by such simple functions from below to get the result. □

Corollary 5. *If $|\phi| \leq \psi$ almost everywhere for ψ radial, decreasing and integrable, then for any $f \in L^1_{loc}(\mathbb{R}^n)$, the maximal function*

$$\Phi^* f(x) := \sup_{t>0} |f * \phi_t(x)| \tag{3.4.11}$$

is weak-$(1,1)$ and strong-(p,p) for $1 < p \leq \infty$.

Corollary 6. *Same assumptions as the above corollary. Then for any $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, or $f \in C_0(\mathbb{R}^n)$ when $p = \infty$, we have*

$$\lim_{t \to 0} f * \phi_t(x) = \left(\int \phi \right) f(x) \text{ for a.e. } x \in \mathbb{R}^n. \tag{3.4.12}$$

In particular, the summability methods we have discussed (Fejér, Poisson, Gauss-Weierstrass), each converges to $f(x)$ a.e. if $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, or $f \in C_0(\mathbb{R}^n)$ when $p = \infty$.

Proof. Since we have the convergence for $f \in \mathcal{S}$, by Theorem 3.2.2, we have the convergence for all $f \in L^p$ or C_0.

The Poisson kernel and the Gauss-Weierstrass kernel are positive radial decreasing functions, so we can apply the previous corollary directly. For the Fejér kernel, we can use the fact that $|F_1(x)| \leq \min(1, (\pi x)^{-2})$. □

3.5 Dyadic maximal function and the weak-$(1,1)$ inequality of the Hardy-Littlewood maximal function

Definition 3.5.1. *Lattice 1-cubes*

$$Q_0 := \{[m_1, m_1+1) \times \cdots \times [m_n, m_n+1) : m_j \in \mathbb{Z}\}. \tag{3.5.1}$$

Lattice 2^{-k}-cubes

$$Q_k := \{[2^{-k}m_1, 2^{-k}(m_1+1)) \times \cdots \times [2^{-k}m_n, 2^{-k}(m_n+1)) : m_j \in \mathbb{Z}\}. \tag{3.5.2}$$

The collection of all dyadic cubes is $\bigcup_{k \in \mathbb{Z}} Q_k$.

Definition 3.5.2. • *We denote the average (or conditional expectation) of f on the 2^{-k} lattice by*

$$E_k f(x) := \sum_{Q \in Q_k} \left(\frac{1}{m(Q)} \int_Q f(y) \mathrm{d}y \right) \chi_Q(x). \tag{3.5.3}$$

Note that Q_k are disjoint, so for all $x \in Q$, $E_k f$ are the same and equal to the average of f on Q.

• *The dyadic maximal function is defined by*

$$M_d f(x) := \sup_{k \in \mathbb{Z}} |E_k f(x)|. \tag{3.5.4}$$

We have the following fundamental properties:

Proposition 9. • *Given k, for any $x \in \mathbb{R}$, x belongs to a unique $Q \in Q_k$.*

• *Any two dyadic cubes either are disjoint or one contains the other.*

• *If $j < k$, then every $Q \in Q_k$ is contained in a unique "ancestor" $Q' \in Q_j$, and contains 2^n "children" in Q_{k+1}.*

Theorem 3.5.3 (The dyadic maximal function is weak $(1,1)$)**.** • *The dyadic maximal function is weak $(1,1)$.*

• *If $f \in L^1_{\mathrm{loc}}$, then $\lim_{k \to \infty} E_k f(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$.*

Proof. • Suppose $f \in L^1$, WTS

$$m(\{x \in \mathbb{R}^n : M_d f(x) > \alpha\}) \le \frac{\|f\|_1}{\alpha}, \quad \forall \alpha > 0. \tag{3.5.5}$$

For such $\alpha > 0$, let

$$\Omega_{k,\alpha} := \{x \in \mathbb{R}^n : |E_k f(x)| > \alpha \text{ but } \left|E_j f(x)\right| \le \alpha,\ \forall j < k\}. \tag{3.5.6}$$

We should understand $x \in \Omega_{k,\alpha}$ as "the stopping time" to hit the level α at time k when we refine the dyadic lattice, then intuitively

$$\{x \in \mathbb{R}^n : M_d f(x) > \alpha\} = \bigsqcup_{k \in \mathbb{Z}} \Omega_{k,\alpha}. \tag{3.5.7}$$

That is because $f \in L^1(\mathbb{R}^n)$ and $\exists k \in \mathbb{Z}$ such that

$$\frac{1}{m(Q)} \int_Q f \le \frac{1}{m(Q)} \|f\|_{L^1} < \alpha \Rightarrow E_k f < \alpha, \quad \text{for some} \quad k \in \mathbb{Z}. \tag{3.5.8}$$

But since E_k takes the same value on each $Q \in Q_k$, thus $\Omega_{k,\alpha}$ is a disjoint union of cubes with side length 2^{-k} in Q_k. Thus

$$\begin{aligned} m\left(\{x \in \mathbb{R}^n : M_d f(x) > \alpha\}\right) = m\left(\bigsqcup_k \Omega_{k,\alpha}\right) &= \sum_k \sum_{Q \in Q_k, Q \subset \Omega_{k,\alpha}} m(Q) \\ &\overset{Q \subset \Omega_k \implies E_k f > \alpha \text{ in } Q}{\le} \sum_k \sum_{Q \in Q_k, Q \subset \Omega_{k,\alpha}} \frac{1}{\alpha} \int_Q |f(y)| \mathrm{d}y \\ &\le \frac{1}{\alpha} \int_{\sqcup \Omega_k} f \mathrm{d}x = \le \frac{1}{\alpha} \|f\|_1. \end{aligned} \tag{3.5.9}$$

- This is a direct application of the maximal function theorem Theorem 3.2.2.

□

The above proof actually has a powerful corollary, namely the Calderón-Zygmund decomposition.

Corollary 7 (Calderón-Zygmund decomposition). *Let $f \in L^1(\mathbb{R}^n)$, $f \geq 0$, $\alpha > 0$. $\exists\{Q_j\}$ disjoint dyadic cubes such that*

- $f(x) \leq \alpha$ *for a.e.* $x \notin \bigsqcup_j Q_j$.
- $m(\bigsqcup_j Q_j) \leq \frac{1}{\alpha}\|f\|_1$. *(volume of "bad cubes")*
- $\alpha < \frac{1}{m(Q_j)}\int_{Q_j} f(x)\mathrm{d}x \leq 2^n\alpha$. *(average on each "bad cube" is still manageable)*

Proof. We just take $\{Q_j\}$ to be $\bigsqcup_k \Omega_{k,\alpha}$ in the previous proof. □

Corollary 8. *Hardy-Littlewood maximal function M' is weak-$(1,1)$. Here the centered cube maximal function is defined by*

$$M'f(x) = \sup_{r>0}\frac{1}{m(Q_r)}\int_{Q_r}|f(x-y)|\mathrm{d}y = \sup_{r>0}\frac{1}{m(Q_r(x))}\int_{Q_r(x)}|f(y)|\mathrm{d}y. \tag{3.5.10}$$

Lemma 5 (comparison between M and M_d). *If $f \geq 0$, then*

$$m\left(\{x \in \mathbb{R}^n : M'f(x) > \alpha\}\right) \leq 2^n m\left(\{x \in \mathbb{R}^n : M_d f(x) > 4^{-n}\alpha\}\right), \quad \forall \alpha > 0. \tag{3.5.11}$$

Proof. **Dyadic idea for estimating the summation.** For every x such that $M'f(x) > \alpha$, there exists $r > 0$ such that $\frac{1}{m(Q_r)}\int_{Q_r(x)} f(y)\mathrm{d}y > \alpha$. Then there exists $k \in \mathbb{Z}$ such that $2^{-k-1} < 2r \leq 2^{-k}$. We consider

$$\{Q_n\} := \text{all dyadic cubes with side length } 2^{-k} \text{ that intersect } Q_r(x). \tag{3.5.12}$$

Then we have at most 2^n such cubes (note that $Q_r(x)$ is a cube with side length $2r$). Thus we have

$$\sum_h \int_{Q_n} f(y)\mathrm{d}y \geq \int_{Q_r(x)} f(y)\mathrm{d}y \tag{3.5.13}$$

$$\Rightarrow \text{for some } h = h_0, \int_{Q_{h_0}} f(y)\mathrm{d}y \geq \frac{1}{2^n}\int_{Q_r(x)} f(y)\mathrm{d}y > \frac{\alpha}{2^n}m(Q_r) \overset{2^{-k}<4r}{\geq} \frac{\alpha}{4^n}m(Q_{h_0}). \tag{3.5.14}$$

On the other hand, since $Q_r(x) \cap Q_{h_0} \neq \varnothing$, we have $x \in 2Q_{h_0}$ (geometrically clear). We know that $\frac{1}{m(Q_{h_0})}\int_{Q_{h_0}} f(y)\mathrm{d}y > \frac{\alpha}{4^n}$, thus

$$Q_{h_0} \subset \left\{x : M_d f(x) > \frac{\alpha}{4^n}\right\}. \tag{3.5.15}$$

From the previous proof of the weak-$(1,1)$ of M_d, we have

$$\left\{x : M_d f(x) > \frac{\alpha}{4^n}\right\} = \bigsqcup_j Q_j, \tag{3.5.16}$$

where Q_j are disjoint dyadic cubes, thus we have $x \in 2Q_{h_0} \subset \bigcup_j 2Q_j$. Therefore, we have shown that

$$\{x : M'f(x) > \alpha\} \subset \bigcup_j 2Q_j \tag{3.5.17}$$

$$\Rightarrow m(\{x : M'f(x) > \alpha\}) \leq \sum_j m(2Q_j) = 2^n \sum_j m(Q_j) = 2^n m\left(\bigsqcup_j Q_j\right) = 2^n m\left(\{x : M_d f(x) > 4^{-n}\alpha\}\right). \tag{3.5.18}$$

□

Proof of the corollary.

$$m\left(\{x \in \mathbb{R}^n : M'f(x) > \lambda\}\right) \leq 2^n m\left(\{x \in \mathbb{R}^n : M_d f(x) > 4^{-n}\lambda\}\right) \overset{M_d \text{ is weak-}(1,1)}{\leq} \frac{8^n}{\lambda}\|f\|_1. \tag{3.5.19}$$

□

Corollary 9 (Lebesgue differentiation theorem)**.** *For any $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, we have*

$$\lim_{r \to 0+} \frac{1}{m(B_r)} \int_{B_r} f(x-y)\mathrm{d}y = f(x) \textit{ for a.e. } x \in \mathbb{R}^n. \tag{3.5.20}$$

From this we also see that $|f| \leq Mf$ a.e. and the same is true for M' or M''.

Proof. Hardy-Littlewood maximal function is weak-$(1,1)$, so we can apply Theorem 3.2.2. □

Remark 25. *Actually we can make it sharper:*

$$\lim_{r \to 0+} \frac{1}{m(B_r)} \int_{B_r} |f(x-y) - f(x)|\mathrm{d}y = 0 \textit{ for a.e. } x \in \mathbb{R}^n. \tag{3.5.21}$$

This follows by Theorem 3.2.2 and the fact that LHS $\leq Mf + |f|$.

If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, and x satisfies eq. (3.5.21), then x is called a Lebesgue point of f.

Corollary 10. *If $f \in L^1$ is not almost everywhere zero, then $Mf \notin L^1$.*

Proof. Suppose we can find $R_0 > 0$ such that

$$\int_{B_{R_0}} |f(x)|\mathrm{d}x > A > 0, \tag{3.5.22}$$

then for $R > 2R_0$ and $|x| = R$, we have $B_{2R}(x) \supset B_{R_0}$. Hence

$$Mf(x) \geq \frac{1}{m(B_{2R})} \int_{B_{2R}(x)} |f(y)|\mathrm{d}y \geq \frac{A}{m(B_{2R})} \int_{B_{R_0}} |f(y)|\mathrm{d}y \gtrsim AR^{-n} = A|x|^{-n}. \tag{3.5.23}$$

Thus $Mf \notin L^1(\mathbb{R}^n)$. □

Though it makes no sense to expect $Mf \in L^1$, we can estimate the integral of Mf on a finite set

Theorem 3.5.4. *If B is a bounded subset of $\mathbb{R}^n$, then*

$$\int_B Mf(x)\mathrm{d}x \leq 2m(B) + C \int_{\mathbb{R}^n} |f(x)| \log^+ |f(x)|\mathrm{d}x. \tag{3.5.24}$$

Here $\log^+ t = \max(\log t, 0)$.

Chapter 4

The Hilbert transform

Motivations: recall the partial integral operator in 1D:

$$S_R f(x) = \int_{-R}^{R} \hat{f}(\xi) e^{2\pi i x\xi} \mathrm{d}\xi, \quad \widehat{S_R f}(\xi) = \hat{f}(\xi)\chi_{[-R,R]}(\xi). \tag{4.0.1}$$

We have

$$S_R f = D_R * f, \quad D_R(x) = \int_{-R}^{R} e^{2\pi i x\xi} \mathrm{d}\xi = \frac{\sin(2\pi R x)}{\pi x} \tag{4.0.2}$$

which is the convolution with the Dirichlet kernel. In Lemma 2, we established a condition for $S_R f$ to converge to f in L^p norm. We also have done some summability methods (Fejér, Poisson, Gauss-Weierstrass) to make $\sigma_R f \to f$ in L^p.

In this chapter, we will prove more general results from which we deduce convergence $S_R f \to f$ in L^p $(1 < p < \infty)$. The key is to study the truncated Hilbert transform and is related to

$$\widehat{Tf}(\xi) = \hat{f}(\xi)\chi_{\{\xi>0\}}(\xi). \tag{4.0.3}$$

This means a multiplier of symbol $\chi_{\{\xi>0\}}(\xi)$, which only leaves the positive frequencies.

4.1 Motivations of conjugate Poisson kernel

This section only gives a heuristic discussion of the Hilbert transform and will not touch any rigourous mathematics.

Given a Schwartz function $f \in \mathscr{S}(\mathbb{R})$, define its **harmonic extension** to the upper half plane $\mathbb{H} = \{(x, y) : x \in \mathbb{R}, y > 0\}$ by

$$u(x,t) = P_t * f(x), \quad P_t(x) = \frac{1}{\pi}\frac{t}{x^2 + t^2}. \tag{4.1.1}$$

If $z = x + \mathrm{i}t$, then

$$u(z) = \int_0^{\infty} \widehat{f}(\xi) e^{2\pi \mathrm{i} z\xi} \mathrm{d}\xi + \int_{-\infty}^{0} \widehat{f}(\xi) e^{2\pi \mathrm{i}\bar{z}\xi} \mathrm{d}\xi. \tag{4.1.2}$$

We define the **harmonic conjugate** of u by

$$\mathrm{i}v(z) = \int_0^{\infty} \widehat{f}(\xi) e^{2\pi \mathrm{i} z\xi} \mathrm{d}\xi - \int_{-\infty}^{0} \widehat{f}(\xi) e^{2\pi \mathrm{i}\bar{z}\xi} \mathrm{d}\xi. \tag{4.1.3}$$

Since f is real-valued, we can see that v is real-valued. Thus $u + \mathrm{i}v = \int_0^\infty \widehat{f}(\xi)e^{2\pi \mathrm{i}z\xi}\mathrm{d}\xi$ (can be understood as "only integrating on the positive frequencies") is holomorphic on $\mathbb{H}$. Also, v is the only harmonic conjugate of u in L^2 s.t. $\|v(\cdot, t)\|_2 < \infty$. In fact we can compute v explicitly:

$$\mathrm{i}v(x,t) = \int_0^\infty \widehat{f}(\xi)e^{2\pi \mathrm{i}x\xi - 2\pi t\xi}\mathrm{d}t - \int_{-\infty}^0 \widehat{f}(\xi)e^{2\pi \mathrm{i}x\xi + 2\pi t\xi}\mathrm{d}t = \int_0^\infty \widehat{f}(\xi)e^{2\pi \mathrm{i}x\xi - 2\pi t|\xi|}\mathrm{d}t - \int_{-\infty}^0 \widehat{f}(\xi)e^{2\pi \mathrm{i}x\xi - 2\pi t|\xi|}\mathrm{d}t. \tag{4.1.4}$$

That is,

$$v(x,t) = \int_{-\infty}^\infty -\mathrm{i}\operatorname{sgn}(\xi)\widehat{f}(\xi)e^{2\pi \mathrm{i}x\xi - 2\pi t|\xi|}\mathrm{d}\xi. \tag{4.1.5}$$

We try to understand this as a "multiplier" (whose official definition will be given later), namely

$$v(x,t) := \int_{-\infty}^\infty \widehat{f}(\xi)\widehat{Q_t}(\xi)e^{2\pi \mathrm{i}x\xi}\mathrm{d}\xi, \quad \widehat{Q_t}(\xi) := -\mathrm{i}\operatorname{sgn}(\xi)e^{-2\pi t|\xi|}. \tag{4.1.6}$$

Note that $\widehat{Q_t}(\xi)$ is L^2 for each fixed $t > 0$, by Theorem 2.4.8 we know that the Fourier inversion formula holds, and $Q_t \in L^2$ can be computed as

$$Q_t(x) = \int_{-\infty}^\infty \widehat{Q_t}(\xi)e^{2\pi \mathrm{i}x\xi}\mathrm{d}\xi = -\mathrm{i}\left(\int_0^\infty e^{-2\pi t\xi}e^{2\pi \mathrm{i}x\xi}\mathrm{d}\xi - \int_{-\infty}^0 e^{2\pi t\xi}e^{2\pi \mathrm{i}x\xi}\mathrm{d}\xi\right) = \frac{1}{\pi}\frac{x}{(x^2+t^2)}. \tag{4.1.7}$$

Then, by the example 10, we know that $\widehat{\underbrace{f}_{\in L^1} * \underbrace{g}_{\in L^2}} = \widehat{f}\cdot\widehat{g}$, thus $\widehat{f * Q_t} = \widehat{f}\cdot\widehat{Q_t} \in L^2$. By the Fourier inversion formula again, we have

$$v(x,t) = Q_t * f(x). \tag{4.1.8}$$

This resembles the Poisson integral formula for $u(x,t)$, where we recall again that

$$u(x,t) = P_t * f(x) = \int_{-\infty}^\infty \widehat{f}(\xi)e^{2\pi \mathrm{i}x\xi - 2\pi t|\xi|}\mathrm{d}\xi. \tag{4.1.9}$$

And we know that this converges to f in L^p norm as $t \to 0$ for $1 \le p < \infty$ since P_t is an approximation of identity. A natural question is whether $v(x,t)$ also converges to something meaningful as $t \to 0$? When we try to do this, we immediately run into an obstacle that Q_t is not even L^1 and $\lim_{t\to 0} Q_t = \frac{1}{\pi x}$ is not even L^1_{loc}.

This naturally leads us to the definition of Hilbert transform. In a nutshell, we should understand the limit of Q_t as $t \to 0$ in the sense of tempered distribution, and the Hilbert transform Hf is just the convolution of a Schwartz function $f \in \mathscr{S}$ with a tempered distribution called "p.v. $\frac{1}{\pi x}$". This definition seems quite restricted, but in fact what we can do is:

- Firstly, this definition immediately extends to L^2 functions since $\widehat{Hf}$ is in fact a well-defined L^2 function when $f \in \mathscr{S}$. Thus $Hf \in L^2$ by the Plancherel theorem, along with $\|Hf\|_2 = \|f\|_2$. Thus H extends to a bounded linear operator on L^2 by the density of $\mathscr{S}$ in L^2. (Moreover, we can even see that H is somehow "self-adjoint".)

- Secondly, we can show by the techniques we have developed (Calderón-Zygmund decomposition) that H is weak-$(1,1)$. By the interpolation together with the L^2 boundedness, we can show that H is strong (p,p) for all $1 < p \le 2$. By the duality argument (the "self-adjointness"), we can extend the strong (p,p) boundedness to all $1 < p < \infty$.

And those above will be the main work of the following two sections.

We should end this introductive section by some remarks on the conjugate Poisson kernel Q_t:

$$(P_t + \mathrm{i}Q_t)(z) = \frac{1}{\pi}\frac{t + \mathrm{i}x}{x^2 + t^2} = \frac{\mathrm{i}}{\pi z}. \tag{4.1.10}$$

This also shows that $P_t + \mathrm{i}Q_t$ is holomorphic on $\mathbb{H}$. Thus Q_t is allows known as the the holomorphic conjugate of the Poisson kernel.

4.2 The conjugate Poisson kernel weakly converges to the principal value of $1/x$

Definition 4.2.1. *The principal value of $1/x$ (p.v. $1/x$) is defined by the following tempered distribution:*

$$\left(p.v.\frac{1}{x}\right)(\varphi) := \lim_{\epsilon\to 0}\left(\int_{|x|>\epsilon}\frac{1}{x}\varphi(x)\mathrm{d}x\right), \quad \forall\varphi\in\mathscr{S}(\mathbb{R}). \tag{4.2.1}$$

Proof. It is well defined since $\varphi\in\mathscr{S}(\mathbb{R})$

$$\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon}\frac{1}{x}\varphi(x)\mathrm{d}x = \lim_{\varepsilon\to 0}\int_{\varepsilon<|x|<1}\frac{1}{x}(\varphi(x)-\varphi(0))\mathrm{d}x + \int_{|x|\geq 1}\frac{1}{x}\varphi(x)\mathrm{d}x \lesssim \|\varphi'\|_\infty + \left\|\varphi(x)\cdot(1+x^2)\right\|_\infty < \infty. \tag{4.2.2}$$

□

Definition 4.2.2. *We define the weak convergence in $\mathscr{S}^*(\mathbb{R})$, it resembles the weak-$*$ convergence in the dual of a Banach space. We say $T_n \to T$ weakly in $\mathscr{S}^*(\mathbb{R})$ if*

$$\lim_{n\to\infty}\langle T_n,\varphi\rangle = \langle T,\varphi\rangle, \quad \forall\varphi\in\mathscr{S}(\mathbb{R}). \tag{4.2.3}$$

Proposition 10. *The Fourier transform $\widehat{\ }$: $\mathscr{S}^*(\mathbb{R}) \to \mathscr{S}^*(\mathbb{R})$ is weakly continuous, since by Theorem 2.4.6 it is in fact "strongly" continuous.*

Proposition 11. *In $\mathscr{S}^*$, Q_t weakly converges to p.v. $1/(\pi x)$.*

Proof. We define

$$\psi_\varepsilon(x) = \chi_{\{|x|>\varepsilon\}}(x)\frac{1}{x}. \tag{4.2.4}$$

We note that ψ_ε defines a tempered distribution by the calculation in the previous proof. It follows at once, by definition that ψ_t weakly converges to p.v. $1/x$ in $\mathscr{S}^*$. Then we just show that $\pi Q_t - \psi_t$ weakly converges to 0 in $\mathscr{S}^*$. Fix $\phi\in\mathscr{S}$, we have

$$\begin{aligned}\langle \pi Q_t - \psi_t, \phi\rangle &= \int_{\mathbb{R}}\frac{x}{x^2+t^2}\phi(x)\mathrm{d}x - \int_{|x|>t}\frac{1}{x}\phi(x)\mathrm{d}x \\ &= \int_{|x|<t}\frac{x\phi(x)}{x^2+t^2}\mathrm{d}x - \int_{|x|>t}\frac{t^2}{x(x^2+t^2)}\phi(x)\mathrm{d}x \\ &= \int_{|x|<1}\frac{x\phi(tx)}{x^2+1}\mathrm{d}x - \int_{|x|>1}\frac{1}{x(x^2+1)}\phi(tx)\mathrm{d}x\end{aligned} \tag{4.2.5}$$

By DCT,

$$\lim_{t\to 0}\langle \pi Q_t - \psi_t, \phi\rangle = \phi(0)\left(\int_{|x|<1}\frac{x}{x^2+1}\mathrm{d}x - \int_{|x|>1}\frac{1}{x(x^2+1)}\mathrm{d}x\right) = 0. \tag{4.2.6}$$

□

Corollary 11. *By the weakly continuity of Fourier transform, we have*

$$\widehat{\left(\frac{1}{\pi}p.v.\frac{1}{x}\right)}(\xi) = (\widehat{\lim_{t\to 0} Q_t})(\xi) = \lim_{t\to 0}\widehat{Q}_t(\xi) = -\mathrm{i}\lim_{t\to 0} sgn(\xi)e^{-2\pi t|\xi|} = -\mathrm{i}sgn(\xi). \tag{4.2.7}$$

This naturally leads to the definition of Hilbert transform. Given a function $f \in \mathscr{S}(\mathbb{R})$, we define its Hilbert transform by

Definition 4.2.3.

$$Hf(x) := \lim_{t\to 0} Q_t * f(x) = \left(\frac{1}{\pi}p.v.\frac{1}{x}\right) * f(x), \tag{4.2.8}$$

or

$$\widehat{Hf}(\xi) = -\mathrm{i}sgn(\xi)\hat{f}(\xi). \tag{4.2.9}$$

Proof. Note that $\widehat{Q}_t(\xi) = -\mathrm{isgn}(\xi)e^{-2\pi t|\xi|}$ converges weakly to $-\mathrm{isgn}(\xi)$ as $t \to 0$ pointwisely, we have $\widehat{Q}_t(\xi)\cdot\widehat{f}(\xi)$ converges to $-\mathrm{isgn}(\xi)\widehat{f}(\xi)$ pointwisely, so this convergence is also in $\mathscr{S}^*$ by DCT. □

Remark 26. *Recall that for general tempered distribution $F \in \mathscr{S}^*$, we can define the convolution $F * f$ for $f \in \mathscr{S}$ by the following pairing:*

$$F * f(x) := \langle F, \underbrace{f(x-\cdot)}_{\in\mathscr{S}}\rangle. \tag{4.2.10}$$

It is obvious that $F * f$ is defined pointwise for all $x \in \mathbb{R}$, and it is actually a continuous function since F is continuous on $\mathscr{S}$.

Remark 27. *Another remark is, we now have a very concrete pointwise expression of $\widehat{Hf}$:*

$$\widehat{Hf}(\xi) = -\mathrm{i}sgn(\xi)\widehat{f}(\xi) \tag{4.2.11}$$

which is in L^2 since $f \in \mathscr{S}$ and $|sgn(\xi)| = 1$. This will actually allow us to extend H to a bounded operator on $L^2(\mathbb{R})$ (will be discussed shortly).

Definition 4.2.4 (Hilbert transform for L^2 functions)**.** *Since $\widehat{Hf}(\xi) = -\mathrm{i}sgn(\xi)\hat{f}(\xi)$ and $|sgn(\xi)| = 1$, we have $\widehat{Hf} \in L^2(\mathbb{R})$ and thus $Hf \in L^2(\mathbb{R})$, along with $\left\|\widehat{Hf}\right\|_2 = \left\|\widehat{f}\right\|_2 = \|Hf\|_2 = \|f\|_2$.*

In this way, H extends to a bounded linear operator on $L^2(\mathbb{R})$ ($\mathscr{S}$ is dense in L^2) with operator norm 1. *We still denote this extension by H and call it the Hilbert transform on $L^2(\mathbb{R})$. It satisfies*

$$H(Hf) = -f, \qquad \int Hf\cdot g = -\int f\cdot Hg. \tag{4.2.12}$$

Proof. Recall we define the notation (cf. Corollary 2)

$$\widetilde{f}(x) = f(-x). \tag{4.2.13}$$

Thus we have

$$\begin{aligned}\int Hf\cdot g &= (Hf,\overline{g})_{L^2} = (\widehat{Hf},\widehat{\overline{g}})_{L^2} = \int \widehat{Hf}\cdot\widehat{\overline{g}}\\ &= -\mathrm{i}\int \mathrm{sgn}(\xi)\widehat{f}(\xi)\overline{\widehat{g}(-\xi)} = \mathrm{i}\int \widehat{f}(\xi)\mathrm{sgn}(-\xi)\overline{\widehat{g}(-\xi)}\\ &= \int \widehat{f}(\xi)\overline{(-\mathrm{i}\widetilde{\mathrm{sgn}(\xi)}\widehat{g}(\xi))} = -(\widehat{f},\widehat{\overline{Hg}})_{L^2} = -(f,\overline{Hg})_{L^2} = -\int f\cdot Hg.\end{aligned} \tag{4.2.14}$$

□

4.3 L^p-boundedness of Hilbert transform: the theorems of M. Riesz and Kolmogorov

Theorem 4.3.1. *For every $f \in (\mathbb{R})$,*

(i) (Kolmogorov) $m(\{x \in \mathbb{R} : |Hf(x)| > \alpha\}) \lesssim \frac{1}{\alpha}\|f\|_1$, $\forall \alpha > 0$.

(ii) (M. Riesz) For $1 < p < \infty$, $\|Hf\|_p \lesssim_p \|f\|_p$. (This can help us extend H to a bounded linear functional on the whole $L^p(\mathbb{R})$ for $1 < p < \infty$)

Remark 28. *Comment: $p = 2$ is already known, which is a very good bonus. The key idea is to apply C-Z decomposition to f and treat the "good part" and "bad part" separately.*

Proof. Replace f with f_+, we will see that it is okay though it may break the smoothness of the Schwartz condition.

- (Kolmogorov) Fixed α, from the C-Z decomposition of f at height α, we obtain disjoint dyadic intervals $\{I_j\}$ such that
$$\begin{cases} \sum_j m(I_j) \lesssim \frac{1}{\alpha}\|f\|_1, \\ \alpha \leq \frac{1}{m(I_j)} \int_{I_j} f(x)\mathrm{d}x \leq 2\alpha, \\ f(x) \leq \alpha \text{ for a.e. } x \notin \bigsqcup_j I_j. \end{cases} \tag{4.3.1}$$
We write $f = g + b$, where
$$g = \begin{cases} f(x), & x \notin \bigsqcup_j I_j, \\ \frac{1}{m(I_j)} \int_{I_j} f(x)\mathrm{d}x, & x \in I_j, \end{cases} \tag{4.3.2}$$
and we have
$$\|g\|_\infty \leq 2\alpha, \quad \|g\|_1 = \|f\|_1, \quad \int_{I_j} b(x)\mathrm{d}x = 0, \tag{4.3.3}$$

Remark 29. $b(x) = \sum_j b_j(x)$, $b_j(x) = \left(f(x) - \frac{1}{m(I_j)} \int_{I_j} f(y)\mathrm{d}y\right)\chi_{I_j}(x)$.

$$\|b\|_1 \leq \|f\|_1 + \sum_j \int_{I_j} \frac{1}{I_j}\left(\int_{I_j} f(x)\mathrm{d}x\right)\mathrm{d}y \leq 2\|f\|_1. \tag{4.3.4}$$

Note that,
$$\|g\|_2 \overset{\text{Hölder}}{\leq} \|g\|_\infty^{\frac{1}{2}}\|g\|_1^{1/2} \lesssim (\alpha)^{1/2}\|f\|_1^{1/2}. \tag{4.3.5}$$
Since $\|H\|_{L^2 \to L^2} \leq 1$, we have
$$m(\{x : |Hg(x)| > \alpha/2\}) \overset{\text{Chebyshev}}{\lesssim} \frac{1}{\alpha^2}\|Hg\|_2^2 \lesssim \frac{\|f\|_1}{\alpha}. \tag{4.3.6}$$
Next we need to control $m(\{x : |Hb(x)| > \alpha/2\})$. Note that
$$\left|\bigcup_j (2I_j)\right| \lesssim \sum_j m(I_j) \lesssim \frac{1}{\alpha}\|f\|_1, \tag{4.3.7}$$

Suffices to check it outside of $\bigcup_j(2I_j)$. Denote c_j the center of I_j, we see that

$$\begin{aligned}
\int_{\mathbb{R}-\bigcup_j(2I_j)} |Hb(x)|\mathrm{d}x &\le \sum_j \int_{\mathbb{R}-2I_j} \left|Hb_j(x)\right|\mathrm{d}x \overset{L^2 \text{ definition of } H}{\lesssim} \sum_j \int_{\mathbb{R}-2I_j} \left|\int_{I_j} \frac{b_j(y)}{x-y}\mathrm{d}y\right|\mathrm{d}x \\
&\overset{\int_{I_j} b_j = 0}{=} \sum_j \int_{\mathbb{R}-2I_j} \left|\int_{I_j} b_j(y)\left(\frac{1}{x-y} - \frac{1}{x-c_j}\right)\mathrm{d}y\right|\mathrm{d}x \\
&= \sum_j \int_{\mathbb{R}-2I_j} \left|\int_{I_j} b_j(y)\frac{y-c_j}{(x-y)(x-c_j)}\mathrm{d}y\right|\mathrm{d}x \\
&\le \sum_j \int_{I_j} \left|b_j(y)\right|\left(\int_{\mathbb{R}-2I_j} \frac{\left|y-c_j\right|}{|x-y|\left|x-c_j\right|}\mathrm{d}x\right)\mathrm{d}y \\
&\le \sum_j \int_{I_j} \left|b_j(y)\right|\mathrm{d}y \int_{\mathbb{R}-2I_j} \frac{\left|I_j\right|}{\left|x-c_j\right|^2}\mathrm{d}x \lesssim \sum_j \int_{I_j} \left|b_j(y)\right|\mathrm{d}y = \|b\|_1 \lesssim \|f\|_1.
\end{aligned} \tag{4.3.8}$$

Hence now we are ready to use Chebyshev to conclude

$$\begin{aligned}
m(\{|Hf(x)| > \alpha\}) &\le m(\{|Hg(x)| > \alpha/2\}) + m(\{|Hb(x)| > \alpha/2\}) \\
&\le m(\{|Hg(x)| > \alpha/2\}) + m(\cup_j(2I_j)) + m(\{x \in \mathbb{R} - \cup_j(2I_j) : |Hb(x)| > \alpha/2\}) \\
&\lesssim \frac{\|f\|_1}{\alpha} + \frac{\|f\|_1}{\alpha} + \frac{\|f\|_1}{\alpha} \lesssim \frac{\|f\|_1}{\alpha}.
\end{aligned} \tag{4.3.9}$$

Remark 30. *Note that this chain of inequalities provides slightly stronger estimate than Young's convolution inequality. Remember we have also tried to prove the L^p convergence of $S_R f$ using Young's convolution inequality, but it fails since $\|D_R\|_1 = \infty$. Here we use the C-Z decomposition to bypass this difficulty: we replace f by b, and on the $\mathbb{R} - \bigcup_j(2I_j)$, we can control the integral of Hb by $\|b\|_1$. This is a very nice trick and will be used frequently in the future.*

- (M. Riesz) Let us use the Marcinkiewicz interpolation theorem to prove this. We already have the weak-$(1,1)$ estimate. We also have the strong-$(2,2)$ estimate. Thus by the Marcinkiewicz interpolation theorem, we have for $1 < p \le 2$,

$$\|Hf\|_p \lesssim_p \|f\|_p. \tag{4.3.10}$$

Remark 31. *Caveat: We only prove the weak-*$(1,1)$ *bound for Schwartz functions, but it is in fact possible to extend this to L^1 space. We can define the Hilbert transform for $f \in L^1$ by approximation: Take $f_n \to f$ in L^1, $f_n \in \mathscr{S}$, then Hf_n is Cauchy **in measure**. We can prove by Fatou that the limit is independent of the choice of f_n, and is weak-*$(1,1)$.

Alternatively, we can also prove it directly by following the proof of Marchinkiewicz.

For $p > 2$, note when $f \in \mathscr{S}$, we use the duality argument:

$$\|Hf\|_p = \sup_{g\in\mathscr{S},\|g\|_{p'}\le 1} \left|\int Hf \cdot g\right| \overset{\text{duality}}{=} \sup_{g\in\mathscr{S},\|g\|_{p'}=1} \left|-\int f \cdot Hg\right| \le \|f\|_p \sup_{g\in\mathscr{S},\|g\|_{p'}=1} \|Hg\|_{p'} \overset{\text{strong } (p',p')}{\lesssim} \|f\|_p. \tag{4.3.11}$$

This argument is very useful when the operator has some self-adjointness properties (Theorem 4.2.4).

□

Remark 32. *H is neither strong-$(1,1)$ nor strong-(∞,∞). In fact, for $f=\chi_{[0,1]}\in L^1\cap L^\infty$, we have*

$$Hf(x)=\frac{1}{\pi}p.v.\int_0^1\frac{1}{x-y}\mathrm{d}y=\frac{1}{\pi}\log\left|\frac{x}{x-1}\right|. \tag{4.3.12}$$

This is not in L^1 nor L^∞.

If we do want to define H on L^∞, we can define it as a map from L^∞ to BMO (will be defined later). But if we do not care that much about a perfect theory and just want to know some examples, we do have chances to find ways to define Hf for some $f\in L^\infty$ (see Problem Set 2, e.g. 9-10)

4.4 Pointwise convergence of $H_\varepsilon f$

In the last section we successfully extended H to a bounded linear operator on $L^p(\mathbb{R})$ for $1<p<\infty$. However, we do this using functional analysis techniques and the definition is given by abstract interpolation or duality arguments. What if we want a more concrete definition, e.g. the pointwise limit of $H_\varepsilon f$ as $\varepsilon\to 0$? Here we will give an affirmative answer to this question using the truncated kernel which we actually have seen in the very beginning point, that is Proposition 11.

As we have learned before, in order to establish pointwise convergence, a powerful to is to show that the maximal function satisfies the weak type estimate. For this, we define the truncated integral and the maximal Hilbert transform:

$$H_\varepsilon(x)=\frac{1}{\pi}\int_{|y|>\varepsilon}\frac{f(x-y)}{y}\mathrm{d}y,\quad H^*f(x)=\sup_{\varepsilon>0}|H_\varepsilon f(x)|. \tag{4.4.1}$$

We will show that H^* is actually weak-$(1,1)$ using the C-Z decomposition again. To do this, we first see that H^* is strong-(p,p) by the following lemma ($1<p<\infty$):

Lemma 6 (Cotlar's inequality)**.** *For $f\in\mathscr{S}(\mathbb{R})$, we have*

$$H^*f(x)\le CMf(x)+MHf(x). \tag{4.4.2}$$

Proof. Suffices to prove it with H^* replaced by arbitrary fixed $\varepsilon>0$, with uniform constant independent of ε.

Choose a non-negative, even, radically decreasing "bump" function ϕ such that

$$\mathrm{supp}\phi\subset[-\frac{1}{2},\frac{1}{2}],\qquad\int\phi=1. \tag{4.4.3}$$

Let $\phi_\varepsilon(x)=\frac{1}{\varepsilon}\phi(\frac{x}{\varepsilon})$. Note that the truncated kernel can be decomposed as

$$\frac{1}{y}\chi_{\{|y|>\varepsilon\}}(y)=\left[\mathrm{p.v.}\frac{1}{x}*\phi_\varepsilon\right](y)+\left(\frac{1}{y}\chi_{\{|y|>\varepsilon\}}(y)-\left[\mathrm{p.v.}\frac{1}{x}*\phi_\varepsilon\right](y)\right). \tag{4.4.4}$$

The contribution of first term is just $Hf*\phi_\varepsilon(x)$, which is pointwisely bounded by $MHf(x)$. For the second term, it is just our dicussion of "$\pi Q_t-\psi_t$" in Proposition 11. For $\varepsilon=1$,

- **tail:** When $|y| > 1$,

$$\left|\frac{1}{y} - \int_{|x|<\frac{1}{2}} \frac{\phi(x)}{y-x}\mathrm{d}x\right|\mathrm{d}y \leq \left|\int_{|x|<\frac{1}{2}} \phi(x)\left(\frac{1}{y} - \frac{1}{y-x}\right)\mathrm{d}x\right| \leq \int_{|x|<\frac{1}{2}} \frac{|x|}{|y(y-x)|}\phi(x)\mathrm{d}x \lesssim_\phi \frac{1}{|y|^2}. \tag{4.4.5}$$

- **close to zero:** When $0 < |y| < 1$, we have

$$\left|\lim_{\delta\to 0}\int_{\delta<|x|<\frac{1}{2}} \frac{\phi(y-x)}{x}\mathrm{d}x\right| \overset{\text{“kill the singularity” strategy}}{=} \left|\int_{|x|<\frac{1}{2}} \frac{\phi(y-x)-\phi(y)}{x}\mathrm{d}x\right| \leq \|\phi'\|_\infty \lesssim_\phi 1. \tag{4.4.6}$$

Hence we proved that for $\varepsilon = 1$,

$$\left(\frac{1}{y}\chi_{\{|y|>1\}} - (\text{p.v. } 1/x) * \phi\right)(y) \lesssim_\phi \begin{cases} \frac{1}{|y|^2}, & |y| > 1, \\ 1, & 0 < |y| < 1 \end{cases} =: \Phi(y), \tag{4.4.7}$$

and for other $\varepsilon > 0$, it is just a rescaled bound $\frac{1}{\varepsilon}\Phi(\frac{y}{\varepsilon})$. Thus by the Proposition 8 we know that the convolution of the second part with f is $\lesssim Mf(x)$.

□

With this lemma, we automatically have the strong-(p, p) boundedness of H^*:

Theorem 4.4.1. *H^*f is strong-(p, p) for $1 < p < \infty$.*

Proof. The Hil bert transform H and the Hardy-Littlewood maximal function M are both strong-(p, p) for $1 < p < \infty$, thus by Cotlar's inequality, we have the strong-(p, p) boundedness of H^*. □

We can then state and prove the weak-$(1, 1)$:

Theorem 4.4.2. *H^*f is weak-$(1, 1)$.*

Proof. We decompose $f = g+b$ using C-Z at height $\alpha > 0$ as before. We recall that we have the following properties: $\begin{cases} \sum_j m(I_j) \lesssim \frac{1}{\alpha}\|f\|_1, \\ f(x) \leq \alpha \text{ for a.e. } x \notin \bigsqcup_j I_j, \\ \alpha \leq \frac{1}{m(I_j)}\int_{I_j} f(x)\mathrm{d}x \leq 2\alpha, \end{cases}$ and

$$g(x) = \begin{cases} f(x), & x \notin \bigsqcup_j I_j, \\ \frac{1}{m(I_j)}\int_{I_j} f(x)\mathrm{d}x, & x \in I_j, \end{cases} \tag{4.4.8}$$

and we have

$$\|g\|_\infty \leq 2\alpha,\ \|g\|_1 = \|f\|_1,\ \int_{I_j} b(x)\mathrm{d}x = 0,\ b(x) = \sum_j b_j(x),\ b_j = \left(f(x) - \frac{1}{m(I_j)}\int_{I_j} f(y)\mathrm{d}y\right)\chi_{I_j}(x). \tag{4.4.9}$$

We want to bound the superlevel set of H^*f by considering H^*g and H^*b separately:

$$H^*f \leq H^*g + H^*b. \tag{4.4.10}$$

For the good part, following from Cotlar's inequality, and the fact that H^* is strong $(2,2)$

$$\|g\|_2 \overset{\text{Hölder}}{\leq} \|g\|_\infty^{1/2}\|g\|_1^{1/2} \lesssim (\alpha)^{1/2}\|f\|_1^{1/2} \Rightarrow \|H^*g\|_2 \lesssim \|g\|_2 \lesssim (\alpha)^{1/2}\|f\|_1^{1/2}, \tag{4.4.11}$$

and thus by Chebyshev,

$$m(\{x : H^*g(x) > \alpha/2\}) \lesssim \frac{1}{\alpha^2}\|H^*g\|_2^2 \lesssim \frac{\|f\|_1}{\alpha}. \tag{4.4.12}$$

For the bad part, it suffices to control it for x outside of $\bigcup_j(2I_j)$. We fix $\varepsilon > 0$ and try to uniformly bound $H_\varepsilon f$. For each $x \notin I_j$, there are three different cases:

- If $\varepsilon \geq d_M$, then $H_\varepsilon b_j(x) = 0$ since the support of b_j is contained in I_j, which is now contained in $\{y : |y - x| < \varepsilon\}$.

- If $\varepsilon \leq d_m$, then $\{y : |y - x| < \varepsilon\}$ and I_j are disjoint, so the truncated integral is actually the full integral: $H_\varepsilon b_j(x) = Hb_j(x)$.

- If $d_m < \varepsilon < d_M$, then

$$\left|H_\varepsilon b_j(x)\right| \leq \frac{1}{\pi}\left|\int_{y\in I_j, |y-x|>\varepsilon} \frac{b_j(y)}{x-y}\mathrm{d}y\right| \leq \frac{1}{\pi}\int_{I_j}\frac{\left|b_j(y)\right|}{|x-y|}\mathrm{d}y \lesssim \int_{I_j}\frac{\left|b_j\right|}{2d_M} \leq \frac{1}{2d_M}\int_{(x-d_M, x+d_M)}\left|b_j\right| \leq Mb_j(x) \tag{4.4.13}$$

  By the weak-$(1,1)$ of M, we have

$$m\left(\left\{x \notin \bigcup_j(2I_j) : Mb_j(x) > \alpha\right\}\right) \lesssim \frac{1}{\alpha}\left\|b_j\right\|_1. \tag{4.4.14}$$

Now we are ready to combine all these estimates and conclude. Note that since all the dyadic intervals are disjoint, at most 2 of these I_j can satisfy $d_m < \varepsilon < d_M$ for fixed x. Thus we have

$$|H_\varepsilon b(x)| \leq \sum_j \left|H_\varepsilon b_j(x)\right| \leq \sum_j \left|Hb_j(x)\right| + 2Mb_j(x), \quad \forall \varepsilon > 0 \tag{4.4.15}$$

Thus we have the following superlevel set estimate:

$$\begin{aligned} m(\{H^*f > \alpha\}) \leq\ & m\left(\left\{H^*g(x) > \frac{\alpha}{2}\right\}\right) + m\left(\bigcup_j 2I_j\right) + m\left(\left\{x \notin \bigcup_j(2I_j) : \sum_j \left|Hb_j(x)\right| > \frac{\alpha}{4}\right\}\right) \\ & + m\left(\left\{x \notin \bigcup_j(2I_j) : \sum_j Mb_j(x) \geq \frac{\alpha}{8}\right\}\right) \lesssim \frac{\|f\|_1}{\alpha}. \end{aligned} \tag{4.4.16}$$

This is because M and H are both weak-$(1,1)$, and $\sum_j \left\|b_j\right\|_1 = \|b\|_1 \lesssim \|f\|_1$.

□

4.5 Multipliers

We consider a bounded function $m \in L^\infty(\mathbb{R}^n)$, we define a bounded operator T_m on $L^2(\mathbb{R}^n)$ by

$$\widehat{T_m f}(\xi) = m(\xi)\widehat{f}(\xi), \tag{4.5.1}$$

and we have

$$\|T_m f\|_2 \le \|m\|_\infty \|f\|_2. \tag{4.5.2}$$

It is easy to see that T_m is $\|m\|_\infty$-bounded on $L^2(\mathbb{R}^n)$. Thus we can always use a bounded function to define a bounded operator on $L^2(\mathbb{R}^n)$ in this way. Sometimes it can be extended from $L^2 \cap L^p$ to a bounded operator on $L^p(\mathbb{R}^n)$ for some $p \neq 2$. If this is the case, we call m an L^p-multiplier.

Example 12. *Let $m(\xi) = -\mathrm{i}sgn(\xi)$, then we can define firstly e.g. for Schwartz function $f \in \mathscr{S}$,*

$$\widehat{T_m f}(\xi) = -\mathrm{i}sgn(\xi)\widehat{f}(\xi). \tag{4.5.3}$$

We know that this is $\widehat{Hf}(\xi) = -\mathrm{i}sgn(\xi)\widehat{f}(\xi)$, thus $T_m f = Hf$ at least for Schwartz functions. Since H is bounded on $L^p(\mathbb{R})$ for $1 < p < \infty$ and $\mathscr{S}$ is L^p-dense in L^p , we see that $m(\xi) = -\mathrm{i}sgn(\xi)$ is an L^p-multiplier for all $1 < p < \infty$.

Example 13. *More generally, for $\chi_{a,b}(\xi)$ the characteristic function on the interval (a, b), we define*

$$\widehat{T_{\chi_{a,b}} f}(\xi) = \chi_{a,b}(\xi)\widehat{f}(\xi). \tag{4.5.4}$$

The operator $T_{\chi_{a,b}}$ defined by this multiplier can also be related to Hilbert transform by

$$S_{a,b} = \frac{\mathrm{i}}{2}\left(M_a H M_{-a} - M_b H M_{-b}\right), \tag{4.5.5}$$

where

$$M_c f(x) = e^{2\pi \mathrm{i} c x} f(x). \tag{4.5.6}$$

Note that the operator M_c will not change any L^p norm, we know that $S_{a,b}$ is also strong-(p, p) for $1 < p < \infty$, by the fact that the Hilbert transform H is strong-(p, p) for $1 < p < \infty$. Thus $\chi_{a,b}(\xi)$ is also an L^p-multiplier for all $1 < p < \infty$.

Example 14. *For an application of the example above, we consider the partial integral operator in 1D, as we have explained in the very beginning of this chapter:*

$$S_R f(x) = \int_{-R}^{R} \widehat{f}(\xi) e^{2\pi \mathrm{i} \xi x} \mathrm{d}\xi. \tag{4.5.7}$$

We now see that by setting $a = -R$, $b = R$, the operator $S_{a,b}$ given by the multiplier $\chi_{-R,R}(\xi)$ is exactly S_R. By the previous example we know that S_R is strong-(p, p) for all $1 < p < \infty$, hence

$$\|S_R f\|_p \lesssim_p \|f\|_p, \quad \forall 1 < p < \infty. \tag{4.5.8}$$

By Lemma 2, we know that $S_R f$ indeed converges to f in L^p norm as $R \to \infty$ for all $f \in L^p(\mathbb{R})$, $1 < p < \infty$:

$$\lim_{R\to\infty} \|S_R f - f\|_p = 0, \quad \forall f \in L^p(\mathbb{R}),\ 1 < p < \infty. \tag{4.5.9}$$

When $p = 1$, we do not have the convergence in norm but only have the convergence in measure:

$$\lim_{R\to\infty} m(\{x \in \mathbb{R} : |S_R f(x) - f(x)| > \alpha\}) = 0, \quad \forall f \in L^1(\mathbb{R}), \text{ and arbitrary fixed } \alpha > 0. \tag{4.5.10}$$

This is because S_R is weak-$(1,1)$. We do this by density argument: for any $f \in L^1(\mathbb{R})$, we can find a truncation $f_M := f\chi_{\{|f|\le M\}\cap[-M,M]}$ such that $\|f - f_M\|_1 \le \varepsilon$, with $M = M(\varepsilon)$. Then we have $f_M \in L^2$, thus

$$\|S_R f_M - f_M\|_1 \lesssim_M \|S_R f_M - f_M\|_2 \to 0 \text{ as } R \to \infty, \tag{4.5.11}$$

thus we can choose $R = R(M,\varepsilon) = R(\varepsilon)$ large enough such that $\|S_R f_M - f_M\|_1 \le \varepsilon$. Finally by the weak-$(1,1)$ of S_R, we have

$$\begin{aligned} m(\{x : |S_R f(x) - f(x)| > \alpha\}) &\le m(\{x : |S_R(f - f_M)(x)| > \alpha/3\}) \\ &\quad + m(\{x : |S_R f_M(x) - f_M(x)| > \alpha/3\}) + m(\{x : |f_M(x) - f(x)| > \alpha/3\}) \\ &\lesssim_\alpha \underbrace{\|f - f_M\|_1}_{\text{weak-}(1,1)} + \underbrace{\|S_R f_M - f_M\|_1}_{\text{Chebyshev}} + \underbrace{\|f - f_M\|_1}_{\text{Chebyshev}} \lesssim \varepsilon + \varepsilon + \varepsilon \lesssim \varepsilon, \end{aligned} \tag{4.5.12}$$

for any large R. Therefore we have the desired convergence in measure.

We summarize the discussion above in the following conclusion:

Theorem 4.5.1. *The partial integral operator S_R in 1D satisfies convergence in measure for all $f \in L^1(\mathbb{R})$, and convergence in L^p norm for all $f \in L^p(\mathbb{R})$, $1 < p < \infty$:*

$$\lim_{R\to\infty} m(\{x \in \mathbb{R} : |S_R f(x) - f(x)| > \alpha\}) = 0, \quad \forall f \in L^1(\mathbb{R}), \text{ and arbitrary fixed } \alpha > 0, \tag{4.5.13}$$

and

$$\lim_{R\to\infty} \|S_R f - f\|_p = 0, \quad \forall f \in L^p(\mathbb{R}),\ 1 < p < \infty. \tag{4.5.14}$$

In particular, for $f \in L^p(\mathbb{R})$, $1 \le p < \infty$, there exists a subsequence $R_k \to \infty$ such that $S_{R_k} f(x) \to f(x)$ for a.e. $x \in \mathbb{R}$, but the subsequence depends on f.

Corollary 12. *If m is a function with bounded variation on $\mathbb{R}$, then it is a multiplier for all $1 < p < \infty$.*

Proof. We assume $\lim_{t\to-\infty} m(t) = 0$ (the limit must exist since m is of BV, and if it is not zero we can add a constant). We normalize m to be right-continuous, then we can denote $\mathrm{d}m$ the corresponding Lebesgue-Stieltjes measure,

$$m(\xi) = \int_{(-\infty,\xi]} \mathrm{d}m(t) = \int_{-\infty}^{\infty} \chi_{(-\infty,\xi]}(t)\mathrm{d}m(t) = \int_{\mathbb{R}} \chi_{(t,\infty)}(\xi)\mathrm{d}m(t). \tag{4.5.15}$$

Therefore,

$$\widehat{T_m f}(\xi) = \int_{\mathbb{R}} \chi_{(t,\infty)}(\xi)\widehat{f}(\xi)\mathrm{d}m(t), \text{ which can be viewed as a linear combination of } \widehat{S_{t,\infty} f}(\xi). \tag{4.5.16}$$

Thus,

$$T_m f(\xi) = \int_{\mathbb{R}} S_{t,\infty} f(\xi)\mathrm{d}m(t). \tag{4.5.17}$$

By Minkowski's integral inequality and the strong-(p,p) of $S_{t,\infty}$, we have

$$\|T_m f\|_p \le \int_{\mathbb{R}} \left\|S_{t,\infty} f\right\|_p \mathrm{d}|m|(t) \lesssim_p \|f\|_p \int_{\mathbb{R}} \mathrm{d}|m|(t) = \|f\|_p \|m\|_{BV(\mathbb{R})}. \tag{4.5.18}$$

Here $\|m\|_{BV(\mathbb{R})}$ is the total variation of m on $\mathbb{R}$, and it is finite. □

We next state a simple extension of multipliers in 1D to higher dimensions, which is a glimpse of the more generalized theory of *singular integrals* which will be discussed in the very next chapter.

If m is a multiplier on $L^p(\mathbb{R})$, then $m(\xi_1)$ is a multiplier on $L^p(\mathbb{R}^n)$.

Example 15. *In fact, if T_m is the one-dimensional multiplier operator associated to m, then for nice enough function $f \in \mathscr{S}(\mathbb{R}^n)$, we can define*

$$\mathbf{T_m} f(x_1, \ldots, x_n) = \left(T_m \left[f(\cdot, x_2, \ldots, x_n)\right]\right)(x_1). \tag{4.5.19}$$

This can be done by first noting

$$\mathbf{m}(\xi_1, \ldots, \xi_n) = m(\xi_1) \cdot 1(\xi_2) \cdots 1(\xi_n). \tag{4.5.20}$$

which is a tensor product structure, and is L^∞, thus it is a genuine multiplier. Moreover, utilizing the tensor product structure we can use Fubini to get the inverse Fourier transform:

$$\check{\mathbf{m}}(x_1, \ldots, x_n) = \check{m}(x_1) \cdot \delta_0(x_2) \cdots \delta_0(x_n), \tag{4.5.21}$$

by noting that the inverse Fourier transform of constant function 1 *is the Dirac delta at* 0*. Thus we have*

$$\widehat{\check{\mathbf{m}} * f}(\xi) = \mathbf{m}(\xi)\widehat{f}(\xi) = m(\xi_1)\widehat{f}(\xi). \tag{4.5.22}$$

Chapter 5

Singular Integrals (Part I)

Recall that for Hilbert transform, we already have the pointwise characterization:

$$Hf = \lim_{\varepsilon \to 0} H_\varepsilon f, \quad \text{a.e.}, \quad \text{where } H_\varepsilon f(x) = \frac{1}{\pi} \int_{|y|>\varepsilon} \frac{f(x-y)}{y} \mathrm{d}y \tag{5.0.1}$$

is the convolution with truncated kernel.

We also the multiplier viewpoint of Hilbert transform from the Fourier side:

$$\widehat{Hf}(\xi) = -\mathrm{i}\,\mathrm{sgn}(\xi)\widehat{f}(\xi). \tag{5.0.2}$$

Singular integrals are generalizations of the Hilbert transform in high dimension. We will study the construction, definition and boundedness of singular integrals in this chapter.

5.1 Definition

We denote $x' = x/|x| \in \mathbb{S}^{n-1}$ for $x \in \mathbb{R}^n \setminus \{0\}$.

Definition 5.1.1.

$$Tf(x) := \left[\left(p.v. \frac{\Omega(x')}{|x|^n}\right) * f\right](x) = \lim_{\varepsilon \to 0} \int_{|y|>\varepsilon} f(x-y) \frac{\Omega(y')}{|y|^n} \mathrm{d}y, \tag{5.1.1}$$

where $\Omega \in L^1(\mathbb{S}^{n-1}, \mathrm{d}\sigma)$ *satisfies the cancellation condition:*

$$\int_{\mathbb{S}^{n-1}} \Omega(x') \mathrm{d}\sigma(x') = 0. \tag{5.1.2}$$

The cancellation property is necessary to ensure the existence of the principal value. We will see this by playing the same trick as in the 1D case:

Proposition 12. *p.v.* $\frac{\Omega(x')}{|x|^n} \in \mathscr{S}^*(\mathbb{R}^n)$ *is a well-defined tempered distribution.*

Proof.

$$\left\langle \mathrm{p.v.} \frac{\Omega(x')}{|x|^n}, \phi \right\rangle = \lim_{\varepsilon \to 0} \int_{|x|>\varepsilon} \frac{\Omega(x')}{|x|^n} \phi(x) \mathrm{d}x \overset{\text{zero mean}}{=} \int_{|x|\le 1} \frac{\Omega(x')}{|x|^n} (\phi(x) - \phi(0)) \mathrm{d}x + \int_{|x|>1} \frac{\Omega(x')}{|x|^n} \phi(x) \mathrm{d}x, \tag{5.1.3}$$

□

Proposition 13. *If the limit* $\lim_{\varepsilon\to 0}\int_{|y|>\varepsilon} f(x-y)\frac{\Omega(y')}{|y|^n}\mathrm{d}y$ *exists pointwise for every Schwartz function (and we do not have a priori assumptions on* Ω*), then* Ω *has a zero mean:*

$$\int_{\mathbb{S}^{n-1}} \Omega(x')\mathrm{d}\sigma(x') = 0. \tag{5.1.4}$$

Proof. Take $\phi \equiv 1$ on the unit ball. If $\int_{\mathbb{S}^{n-1}} \Omega(y')\mathrm{d}y' \neq 0$, then it will blow up and contradict the existence of the limit. □

Definition 5.1.2. *A very important example of singular integral is the Riesz transform:*

$$R_j f(x) := \left(p.v.\ \frac{x_j}{|x|^{n+1}} * f\right)(x) \tag{5.1.5}$$

that is we take the homogeneous function $\Omega(x') = \frac{x_j}{|x|}$ *on* $\mathbb{S}^{n-1}$.

Remark 33. *Other kernels that we may consider include Newtonian potential, logarithmic potential, etc.*

5.2 Formula of the Fourier transform of the kernel p.v. $\Omega(x')/|x|^n$

For Hilbert transform, i.e. the singular integral with kernel p.v. π/x, we have already computed its Fourier transform in Theorem 4.2.4:

$$\widehat{\mathrm{p.v.}\ \pi/x}(\xi) = -\mathrm{i}\pi\mathrm{sgn}(\xi). \tag{5.2.1}$$

We can generalize this to higher dimensions.

We note that the kernel p.v.π/x can be viewed as a homogeneous function of degree -1 in $\mathbb{R}$, and the Fourier transform of it is $-\mathrm{i}\pi\mathrm{sgn}(\xi)$ which is homogeneous of degree 0 in $\mathbb{R}$. This is not a coincidence, in fact we have the following general result:

Theorem 5.2.1. *If* $T \in \mathscr{S}^*(\mathbb{R}^n)$ *is a homogeneous distribution of degree a, then* $\widehat{T}$ *is a homogeneous distribution of degree* $-n-a$.

Here, we define the homogeneity of distributions as follows:

Definition 5.2.2. *We say that a tempered distribution* $T \in \mathscr{S}^*(\mathbb{R}^n)$ *is homogeneous of degree a if for any* $\lambda > 0$, *we have*

$$\langle T, \phi_\lambda\rangle = \lambda^a \langle T, \phi\rangle, \quad \forall \phi \in \mathscr{S}(\mathbb{R}^n), \tag{5.2.2}$$

where $\phi_\lambda(x) = \lambda^{-n}\phi(\lambda^{-1}x)$.

Example 16. *The kernel p.v.* $\frac{\Omega(x')}{|x|^n}$ *is homogeneous of degree* $-n$.

Proof.

$$\begin{aligned}\left\langle \mathrm{p.v.}\frac{\Omega(x')}{|x|^n}, \phi_\lambda\right\rangle &= \lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} \frac{\Omega(x')}{|x|^n}\lambda^{-n}\phi(\lambda^{-1}x)\mathrm{d}x \overset{\text{change of variable}}{=} \lim_{\varepsilon\to 0}\int_{|y|>\lambda^{-1}\varepsilon} \frac{\Omega(y')}{\lambda^{-n}|y|^n}\lambda^{-n}\phi(y)\lambda^{-n}\mathrm{d}y \\ &= \lambda^{-n}\lim_{\varepsilon\to 0}\int_{|y|>\lambda^{-1}\varepsilon}\frac{\Omega(y')}{|y|^n}\phi(y)\mathrm{d}y = \lambda^{-n}\left\langle \mathrm{p.v.}\frac{\Omega(x')}{|x|^n}, \phi\right\rangle.\end{aligned} \tag{5.2.3}$$

□

Proof of the theorem. For any $\phi \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\begin{aligned}\left\langle \widehat{T}, \phi_\lambda \right\rangle = \lambda^{-n} \left\langle T, \widehat{\phi(\lambda^{-1}\cdot)} \right\rangle &\overset{\text{scaling property of FT}}{=} \left\langle T, \widehat{\phi}(\lambda \cdot) \right\rangle = \lambda^{-n} \left\langle T, [\widehat{\phi}]_\lambda \right\rangle \\ &\overset{\text{homogeneity of } T}{=} \lambda^{-n-a} \left\langle T, \widehat{\phi} \right\rangle = \lambda^{-n-a} \left\langle \widehat{T}, \phi \right\rangle\end{aligned} \tag{5.2.4}$$

which shows that $\widehat{T}$ is homogeneous of degree $-n-a$. □

Example 17 (The Fourier transform of $|x|^{-\alpha}$, $0 < \alpha \leq n$)**.** *$f(x) = |x|^{-\alpha}$ is a homogeneous function of degree $-\alpha$ in $\mathbb{R}^n$. By the theorem above, we know that $\widehat{f}$ is a homogeneous distribution of degree $-n+\alpha$.*

For $\frac{1}{n} < \alpha < n$, by cutoff argument we know that $|x|^{-\alpha} \in L^1 + L^2$, thus $\widehat{f} \in L^\infty + L^2$ can be pointwise defined. Now we have that $\widehat{f}$ is rotationally invariant, L^1_{loc} and homogeneous of degree $-n+\alpha$, thus for some constant $C_{\alpha,n}$,

$$\widehat{f}(\xi) = C_{\alpha,n} |\xi|^{\alpha-n}. \tag{5.2.5}$$

By pairing with Gaussian function and computing the constant, we have

$$\left\langle f, e^{-\pi|x|^2} \right\rangle = \left\langle \widehat{f}, e^{-\pi|\xi|^2} \right\rangle \tag{5.2.6}$$

which means that

$$\int_{\mathbb{R}^n} |x|^{-\alpha} e^{-\pi|x|^2} \mathrm{d}x = C_{\alpha,n} \int_{\mathbb{R}^n} |\xi|^{\alpha-n} e^{-\pi|\xi|^2} \mathrm{d}\xi. \tag{5.2.7}$$

By polar coordinate change

$$\int_0^\infty r^{n-1-\alpha} e^{-\pi r^2} \mathrm{d}r = C_{\alpha,n} \int_0^\infty r^{\alpha-1} e^{-\pi r^2} \mathrm{d}r, \tag{5.2.8}$$

Note that

$$\int_0^\infty e^{-\pi r^2} r^b \mathrm{d}r \overset{u=\pi r^2}{=} \int_0^\infty e^{-u} \frac{u^{(b-1)/2}}{2\pi^{(b+1)/2}} \mathrm{d}u = \frac{\Gamma\left(\frac{b+1}{2}\right)}{2\pi^{(b+1)/2}}. \tag{5.2.9}$$

We see that

$$C_{\alpha,n} = \frac{\Gamma(\frac{n-\alpha}{2})}{\Gamma(\frac{\alpha}{2})} \pi^{\alpha - \frac{n}{2}}. \tag{5.2.10}$$

This is also true for $\frac{n}{2} < \operatorname{Re}\alpha < n$, and by continuity of $\widehat{|x|^{-\alpha}}$ in the sense of distributions, we can extend this to $\frac{n}{2} \leq \operatorname{Re}\alpha < n$. By $\widehat{\widehat{f}} = \widetilde{f}$, it is also true for $0 < \operatorname{Re}\alpha < n$.

The main theorem of this section is the following formula of the Fourier transform of the kernel of singular integrals:

Theorem 5.2.3.

- *The Fourier transform of the kernel of singular integral $\widehat{\text{p.v.}\frac{\Omega(x')}{|x|^n}}$ is a homogeneous function of degree 0 in $\mathbb{R}^n$.*
- *Therefore, it is enough to know its value on the unit sphere $\mathbb{S}^{n-1}$. The formula is given by*

$$m(\xi) = \int_{\mathbb{S}^{n-1}} \Omega(u) \left[\log \frac{1}{|u \cdot \xi'|} - \mathrm{i}\frac{\pi}{2} \operatorname{sgn}(u \cdot \xi') \right]. \tag{5.2.11}$$

Proof. We play the same trick as in the 1D case but with polar coordinates. For $\xi \in \mathbb{S}^{n-1}$, we have

$$\begin{aligned} m(\xi) &\overset{\text{by defn}}{=} \int_{\varepsilon<|y|<\frac{1}{\varepsilon}} \frac{\Omega(y')}{|y|^n} e^{-2\pi \mathrm{i} y\cdot\xi} \mathrm{d}\xi \\ &\overset{\text{polar coord}}{=} \int_{\mathbb{S}^{n-1}} \Omega(u) \left(\int_{\varepsilon}^{\frac{1}{\varepsilon}} \frac{e^{-2\pi \mathrm{i} r(u\cdot\xi)}}{r} \mathrm{d}r \right) \mathrm{d}\sigma(u) \\ &= \int_{\mathbb{S}^{n-1}} \Omega(u) \left(\int_{\varepsilon}^{1} [e^{-2\pi \mathrm{i} ru\cdot\xi} - 1] \frac{\mathrm{d}r}{r} + \int_{1}^{\frac{1}{\varepsilon}} e^{-2\pi \mathrm{i} ru\cdot\xi} \frac{\mathrm{d}r}{r} \right) \mathrm{d}\sigma(u) \\ &=: I_1 - \mathrm{i} I_2, \end{aligned} \tag{5.2.12}$$

where

$$I_1 = \int_{\mathbb{S}^{n-1}} \Omega(u) \left(\int_{\varepsilon}^{1} [\cos(2\pi ru\cdot\xi) - 1] \frac{\mathrm{d}r}{r} + \int_{1}^{\frac{1}{\varepsilon}} \cos(2\pi ru\cdot\xi) \frac{\mathrm{d}r}{r} \right) \mathrm{d}\sigma(u), \tag{5.2.13}$$

and

$$I_2 = \int_{\mathbb{S}^{n-1}} \Omega(u) \left(\int_{\varepsilon}^{\frac{1}{\varepsilon}} \sin(2\pi ru\cdot\xi) \frac{\mathrm{d}r}{r} \right) \mathrm{d}\sigma(u). \tag{5.2.14}$$

Here, the inner integral of I_2 is

$$\int_{\varepsilon}^{\frac{1}{\varepsilon}} \sin(2\pi ru\cdot\xi) \frac{\mathrm{d}r}{r} = \int_{2\pi|u\cdot\xi|\varepsilon}^{2\pi|u\cdot\xi|/\varepsilon} \sin(s)\,\mathrm{sgn}(u\cdot\xi) \frac{\mathrm{d}s}{s} \to \frac{\pi}{2}\,\mathrm{sgn}(u\cdot\xi) \text{ as } \varepsilon \to 0, \tag{5.2.15}$$

and the inner integral of I_1 is

$$\int_{2\pi|u\cdot\xi|\varepsilon}^{1} [\cos(s) - 1] \frac{\mathrm{d}s}{s} + \int_{1}^{2\pi|u\cdot\xi|/\varepsilon} \cos(s) \frac{\mathrm{d}s}{s} - \int_{1}^{2\pi|u\cdot\xi|} \frac{\mathrm{d}s}{s} \to \text{constant} - \log \frac{1}{|u\cdot\xi|} \text{ as } \varepsilon \to 0. \tag{5.2.16}$$

But the constant part vanishes after integrating against $\Omega(u)$ on $\mathbb{S}^{n-1}$ due to the zero mean property. Thus we have

$$m(\xi) = \int_{\mathbb{S}^{n-1}} \Omega(u) \left[\log \frac{1}{|u\cdot\xi'|} - \mathrm{i}\frac{\pi}{2}\,\mathrm{sgn}(u\cdot\xi') \right]. \tag{5.2.17}$$

□

Now that we have the formula of the Fourier transform of the kernel of singular integrals, we can use it to study the L^p boundedness of singular integrals in the next two sections. We will separate the odd and even kernels,

$$\Omega = \Omega_o + \Omega_e, \quad \Omega_o(x') = \frac{\Omega(x') - \Omega(-x')}{2}, \quad \Omega_e(x') = \frac{\Omega(x') + \Omega(-x')}{2}. \tag{5.2.18}$$

Then by noting that $\log \frac{1}{|u\cdot\xi'|}$ is even in u, and $\mathrm{sgn}(u\cdot\xi')$ is odd in u, we have

$$m(\xi) = m_o(\xi) + m_e(\xi), \tag{5.2.19}$$

where

$$m_o(\xi) = -\mathrm{i}\frac{\pi}{2} \int_{\mathbb{S}^{n-1}} \Omega_o(u)\,\mathrm{sgn}(u\cdot\xi') \mathrm{d}\sigma(u), \tag{5.2.20}$$

and

$$m_e(\xi) = \int_{\mathbb{S}^{n-1}} \Omega_e(u) \log \frac{1}{|u \cdot \xi'|} \mathrm{d}\sigma(u). \tag{5.2.21}$$

Recall that for Hilbert transform in 1D, we strongly use the L^∞ boundedness of the multiplier. Here, to make the multiplier $m(\xi)$ bounded, we assume that

$$\Omega_o \in L^1(\mathbb{S}^{n-1}), \quad \Omega_e \in L^q(\mathbb{S}^{n-1}), \text{ for some } q > 1. \tag{5.2.22}$$

Then, by Hölder's inequality, we have

$$\|m\|_\infty \lesssim \|\Omega_o\|_1 + \|\Omega_e\|_q < \infty, \tag{5.2.23}$$

which ensures the L^2 boundedness of the singular integral T by Plancherel, where

$$Tf := \left(\text{p.v.} \frac{\Omega(x')}{|x|^n}\right) * f, \quad \|Tf\|_2 \lesssim \left(\|\Omega_o\|_1 + \|\Omega_e\|_q\right) \|f\|_2. \tag{5.2.24}$$

5.3 L^p boundedness of odd singular integrals by the method of rotation

If we are given a 1D linear or sublinear operator T, we can define the corresponding directional operator T_u on $L^p(\mathbb{R}^n)$ for any $u \in \mathbb{S}^{n-1}$ by

$$T_u f(x) := T[f(\cdot u + \overline{x})](x_1), \quad (x_1, \overline{x}) \in \mathbb{R} \times \mathrm{Span}(u)^\perp, \tag{5.3.1}$$

where $\overline{x} = x - x_1 u$ is the projection of x onto the hyperplane orthogonal to u. By Fubini (the integrability in the product measure space implies the integrability in each direction for a.e. slice).

Some examples include:

Example 18. • *Directional Hardy-Littlewood maximal function:*

$$M_u f(x) := \sup_{h>0} \frac{1}{2h} \int_{-h}^{h} f(x + tu) \mathrm{d}t. \tag{5.3.2}$$

• *Directional Hilbert transform in direction u:*

$$H_u f(x) := \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{|t|>\varepsilon} \frac{f(x - tu)}{t} \mathrm{d}t. \tag{5.3.3}$$

• *Directional maximal Hilbert transform in direction u:*

$$H_u^* f(x) := \sup_{\varepsilon>0} \left| H_{\varepsilon,u} f(x) \right|. \tag{5.3.4}$$

Note that M, H, H^* are all strong-(p, p) for $1 < p < \infty$ in 1D, thus by Fubini's theorem we know that M_u, H_u, H_u^* ***are also strong-(p, p) for $1 < p < \infty$ on $L^p(\mathbb{R}^n)$, with the same operator norm as in 1D.***

Proposition 14. *Given one-dimensional operator T and $\Omega \in L^1(\mathbb{S}^{n-1})$ with zero mean, we can define the operator*

$$T_\Omega f(x) := \int_{\mathbb{S}^{n-1}} \Omega(u) T_u f(x) \mathrm{d}\sigma(u). \tag{5.3.5}$$

This can be viewed as a superposition of directional operators T_u. (Remark: If Ω is chosen as a probability distribution i.e. is nonnegative and integrates to 1, then this it is a convex superposition.) is bounded on L^p if T is bounded on $L^p(\mathbb{R})$.

Proof. By Minkowski's integral inquality, together with the strong-(p, p) of T_u. □

Theorem 5.3.1. *Let Ω be an odd function on $\mathbb{S}^{n-1}$ with $\Omega \in L^1(\mathbb{S}^{n-1})$. Then the odd singular integral operator is bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$.*

Proof. The only work left is to express the odd singular integral as a superposition of directional Hilbert transforms. By polar coordinates, we have

$$\begin{aligned}
\left(\mathrm{p.v.}\frac{\Omega(x')}{|x|^n}\right) * f(x) &= \lim_{\varepsilon\to 0}\int_{|y|>\varepsilon} f(x-y)\frac{\Omega(y')}{|y|^n}\mathrm{d}y \overset{\text{polar coord}}{=} \lim_{\varepsilon\to 0}\int_{\mathbb{S}^{n-1}}\Omega(u)\left(\int_{r>\varepsilon} f(x-ru)\frac{\mathrm{d}r}{r}\right)\mathrm{d}\sigma(u) \\
&\overset{\Omega \text{ is odd}}{=} \frac{1}{2}\lim_{\varepsilon\to 0}\int_{\mathbb{S}^{n-1}}\Omega(u)\left(\int_{|t|>\varepsilon} f(x-tu)\frac{\mathrm{d}t}{t}\right)\mathrm{d}\sigma(u) \\
&\overset{\Omega \text{ has zero mean}}{=} \frac{1}{2}\int_{\mathbb{S}^{n-1}}\Omega(u)\underbrace{\left[\lim_{\varepsilon\to 0}\int_{\varepsilon<|t|<1}\frac{f(x-tu)-f(x)}{t} + \int_{|t|>1}\frac{f(x-tu)}{t}\mathrm{d}t\right]}_{\pi H_u f(x)}\mathrm{d}\sigma(u) \\
&= \frac{\pi}{2}\int_{\mathbb{S}^{n-1}}\Omega(u) H_u f(x)\mathrm{d}\sigma(u).
\end{aligned} \tag{5.3.6}$$

Since $H_u f$ is strong-(p, p) with the same operator norm as in 1D (independent of u), by Proposition 14 we know that the odd singular integral is also strong-(p, p) for all $1 < p < \infty$. □

Remark 34. *A natural question is that, whether the method of rotation gives weak-$(1, 1)$ or not? Short answer: there can be much more work to do!*

We will see in the next chapter that, if we have some axiomatic properties of T, then we can get the weak-$(1, 1)$ boundedness.

Corollary 13. *Similarly, the maximal odd singular integral*

$$T^* f(s) = \sup_{\varepsilon>0}\left|\int_{|y|>\varepsilon} f(x-y)\frac{\Omega(y')}{|y|^n}\mathrm{d}y\right| \tag{5.3.7}$$

is strong-(p, p) for all $1 < p < \infty$.

Corollary 14. *The Riesz transform*

$$R_j f(x) = c_n\left(p.v.\frac{x_j}{|x|^{n+1}} * f\right)(x) \tag{5.3.8}$$

is strong-(p, p) for all $1 < p < \infty$.

Proof. This is because the kernel $\frac{x_j}{|x|^{n+1}}$ corresponds to $\Omega(x') = x'_j = x_j/|x|$, which is an odd function on $\mathbb{S}^{n-1}$. ◻

Remark 35.
- *In the definition we need the normalization constant c_n to ensure*

$$\widehat{R_j f}(\xi) = -\mathrm{i}\frac{\xi_j}{|\xi|}\widehat{f}(\xi), \tag{5.3.9}$$

and so that

$$\sum_{j=1}^{n} R_j^2 f = -f. \tag{5.3.10}$$

- *The c_n can be determined by noting*

$$\frac{\partial}{\partial x_j}\left(|x|^{-n+1}\right) = -(1-n)p.v.\frac{x_j}{|x|^{n+1}}. \tag{5.3.11}$$

This can be verified by the definition of the distributional derivative:

$$\left\langle \frac{\partial}{\partial x_j}|x|^{-n+1}, \phi \right\rangle = -\left\langle |x|^{-n+1}, \frac{\partial \phi}{\partial x_j} \right\rangle = -\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} |x|^{-n+1}\frac{\partial \phi}{\partial x_j}\mathrm{d}x. \tag{5.3.12}$$

By integration by parts, and noting that the boundary term vanishes as $\varepsilon \to 0$, we can verify that this gives the p.v. formula above. Thus, by the Fourier transform of $|x|^{-\alpha}$, we have

$$p.v.\widehat{\frac{x_j}{|x|^{n+1}}}(\xi) = \frac{1}{1-n}\widehat{\frac{\partial}{\partial x_j}|x|^{-n+1}}(\xi) \overset{\textit{Property of FT}}{=} \frac{2\pi \mathrm{i}\xi_j}{1-n}\widehat{|x|^{-n+1}}(\xi) \overset{\textit{example 17}}{=} \frac{2\pi\mathrm{i}\xi_j}{1-n}\cdot\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{n-1}{2})}\pi^{\frac{n}{2}-1}|\xi|^{-1}. \tag{5.3.13}$$

Therefore it follows that

$$c_n = \Gamma\left(\frac{n+1}{2}\right)\pi^{-\frac{n+1}{2}}. \tag{5.3.14}$$

5.4 L^p boundedness of even singular integrals by the Riesz transform

As previously mentioned, in this section we will be working on the case when

$$\Omega \text{ is even and } \Omega \in L^q(\mathbb{S}^{n-1}), \quad q > 1. \tag{5.4.1}$$

The key idea is use the following algebraic identity involving Riesz transforms:

$$Tf = -\sum_{j=1}^{n} R_j^2 Tf \overset{\text{understood as}}{=} -\sum_{j=1}^{n} R_j\left(R_j T\right)f. \tag{5.4.2}$$

We then reduce the L^p boundedness of T to that of R_jT, which is **an odd singular integral operator** (this is the key observation!). However, the operator R_jT is formally a "double convolution" operator, which can be complicated to analyze. Our strategy here is to first approximate it by the truncated kernel. We define the truncated kernel of T by

$$K_\varepsilon(x) := \chi_{|x|>\varepsilon}\frac{\Omega(x')}{|x|^n}, \quad T_\varepsilon f := K_\varepsilon * f. \tag{5.4.3}$$

We will have the following lemma:

Lemma 7. *$K_\varepsilon \in L^r$ ($1 < r < q$).*

Proof. By Young's convolution inequality, since $K_\varepsilon \in L^{1+\delta}$ for any $\delta > 0$ (due to the integrability near 0 and ∞), we have that $K_\varepsilon *$ is a well-defined bounded operator on L^p for all $1 < p < \infty$. □

Lemma 8. *$R_j(K_\varepsilon * f) = (R_j K_\varepsilon) * f$ for all Schwartz function f.*

Proof. This follows from the associativity of convolution and the fact that $R_j f = \text{p.v.}\frac{x_j}{|x|^{n+1}} * f$. □

The following lemma is the key ingredient to show the L^p boundedness of even singular integrals.

Lemma 9. *There exists $\widetilde{K}_j$ odd, homogeneous of degree $-n$ such that it is the limit kernel of $R_j K_\varepsilon$:*

$$\widetilde{K}_j|_K = \lim_{\varepsilon\to 0} R_j K_\varepsilon|_K, \quad \text{in } L^\infty \tag{5.4.4}$$

on every compact set K that $\not\ni 0$.

Remark 36. *Intuitively, $\widetilde{K}_j$ is just the kernel of $R_j T$. But we need to make it rigorous. We will omit the technical notation of the restriction on compact set in the future.*

Proof. We will look at $R_j K_\varepsilon - R_j K_\nu$ ($0 < \varepsilon < \nu$ fixed) and prove a Cauchy-sequence argument. Keeping in mind that x is away from 0 and by direct computation, we have

$$\begin{aligned} R_j K_\varepsilon(x) - R_j K_\nu(x) &= c_n \int_{\varepsilon<|y|<\nu} \frac{(x_j - y_j)}{|x-y|^{n+1}} \frac{\Omega(y')}{|y|^n} \mathrm{d}y \overset{\text{zero mean property}}{=} c_n \int_{\varepsilon<|y|<\nu} \left[\frac{x_j - y_j}{|x-y|^{n+1}} - \frac{x_j}{|x|^{n+1}}\right] \frac{\Omega(y')}{|y|^n} \mathrm{d}y \\ &\overset{\text{differentiating by MVT, choosing } |x|>2\nu}{\lesssim} \frac{1}{|x|^{n+1}} \int_{\varepsilon<|y|<\nu} |y| \frac{|\Omega(y')|}{|y|^n} \mathrm{d}y = \frac{1}{|x|^{n+1}} \int_{\varepsilon<|y|<\nu} \frac{|\Omega(y')|}{|y|^{n-1}} \mathrm{d}y \\ &\overset{\text{polar coord}}{=} \frac{1}{|x|^{n+1}} \int_{\mathbb{S}^{n-1}} |\Omega(u)| \left(\int_\varepsilon^\nu 1 \mathrm{d}r\right) \mathrm{d}\sigma(u) \le \frac{\nu}{|x|^{n+1}} \|\Omega\|_1, \end{aligned} \tag{5.4.5}$$

therefore, $R_j K_\varepsilon(x)$ is a Cauchy sequence in L^∞ on every compact set not containing 0. So for almost every x, we can define the pointwise limit

$$K_j^*(x) := \lim_{\varepsilon\to 0} R_j K_\varepsilon(x). \tag{5.4.6}$$

The function $R_j K_\varepsilon$ is odd for every $\varepsilon > 0$, so, by modifying the value of K_j^* on a measure zero set if necessary, we can assume that K_j^* is also odd.

Intuitively, K_j^* is the desired limit kernel $\widetilde{K}_j$. The only thing left is to show that $\widetilde{K}_j$ is homogeneous of degree $-n$. We begin with the observation that for fixed $\varepsilon > 0$,

$$\begin{aligned} R_j K_\varepsilon(\lambda x) &= \lim_{\delta\to 0} c_n \int_{|\lambda x - y|>\delta} \frac{\lambda x_j - y_j}{|\lambda x - y|^{n+1}} K_\varepsilon(y) \mathrm{d}y = \lim_{\delta\to 0} c_n \int_{|x-z|>\delta/\lambda} \frac{\lambda x_j - \lambda z_j}{|\lambda x - \lambda z|^{n+1}} K_\varepsilon(\lambda z) \lambda^n \mathrm{d}z \\ &= \lim_{\delta\to 0} c_n \lambda^{-n} \int_{|x-z|>\delta/\lambda} \frac{x_j - z_j}{|x-z|^{n+1}} K_{\varepsilon/\lambda}(z) \mathrm{d}z = \lambda^{-n} R_j K_{\varepsilon/\lambda}(x). \end{aligned} \tag{5.4.7}$$

Hence, by taking the limit $\varepsilon \to 0$, we have $K_j^*(\lambda x) = \lambda^{-n} K_j^*(x)$ for almost every x, and the null set may depend on λ. But K_j^* is a measurable function, thus the set

$$D = \{(x,\lambda) \in \mathbb{R}^n \times (0,\infty) : K_j^*(\lambda x) \ne \lambda^{-n} K_j^*(x)\} \tag{5.4.8}$$

is also zero measure in $\mathbb{R}^n \times (0,\infty)$. By Fubini's theorem, for almost every $x \in \mathbb{R}^n$, the section

$$D_x = \{\lambda \in (0,\infty) : (x,\lambda) \in D\} \tag{5.4.9}$$

is also measure zero in $(0,\infty)$. Thus there exists $\rho \in (0,\infty)$ such that

$$\text{measure}_{S_\rho \times \mathbb{R}}(D \cap (S_\rho \times \mathbb{R})) = 0, \tag{5.4.10}$$

where S_ρ is the sphere of radius ρ.

Now we redefine $\widetilde{K}_j$ as follows:

$$\widetilde{K}_j(x) := \frac{\rho^n}{|x|^n} K_j^*\left(\rho \frac{x}{|x|}\right), \quad x \neq 0, \tag{5.4.11}$$

as well as $\widetilde{K}_j(0) = 0$. It is easy to see that $\widetilde{K}_j$ is homogeneous of degree $-n$, and it is equal to K_j^* almost everywhere. Thus we have finished the proof. □

Lemma 10. *The limit kernel $\widetilde{K}_j \in L^1(\mathbb{S}^{n-1}, \mathrm{d}\sigma)$ satisfies*

$$\int_{\mathbb{S}^{n-1}} \left|\widetilde{K}_j(x')\right| \mathrm{d}\sigma(x') \lesssim \|\Omega\|_{L^q(\mathbb{S}^{n-1})}, \tag{5.4.12}$$

furthermore, if we consider the truncated kernel $\widetilde{K}_{j,\varepsilon} = \chi_{|x|>\varepsilon}\widetilde{K}_j(x)$, then

$$\Delta_\varepsilon := R_j K_\varepsilon - \widetilde{K}_{j,\varepsilon} \in L^1(\mathbb{R}^n), \quad \textit{and} \quad \|\Delta_\varepsilon\|_1 \lesssim \|\Omega\|_{L^q(\mathbb{S}^{n-1})}. \tag{5.4.13}$$

Proof.

$$\begin{aligned}\int_{\mathbb{S}^{n-1}} \left|\widetilde{K}_j(u)\right| \mathrm{d}\sigma(u) \sim \int_{1<|x|<2} \left|\widetilde{K}_j(x)\right| \mathrm{d}x &\leq \int_{1<|x|<2} \left|\widetilde{K}_j(x) - R_j K_{\frac{1}{2}}(x)\right| \mathrm{d}x + \int_{1<|x|<2} \left|R_j K_{\frac{1}{2}}(x)\right| \mathrm{d}x \\ &\leq \underbrace{\int_{1<|x|<2} \left|\widetilde{K}_j(x) - R_j K_{\frac{1}{2}}(x)\right| \mathrm{d}x}_{(I)} + \underbrace{\int_{1<|x|<2} \left|R_j K_{\frac{1}{2}}(x)\right| \mathrm{d}x}_{(II)}.\end{aligned} \tag{5.4.14}$$

(I) is bounded by $C\|\Omega\|_1 \leq C\|\Omega\|_q$ pointwise by the proof of the L^q-boundedness of R_j.

$$\text{(II)} \overset{\text{Hölder}}{\lesssim} \left\|R_j K_{\frac{1}{2}}\right\|_{L^q} \overset{L^q\text{-boundedness of } R_j}{\lesssim} \left\|K_{\frac{1}{2}}\right\|_q \overset{\text{polar coord}}{\lesssim} \|\Omega\|_{L^q(\mathbb{S}^{n-1})}. \tag{5.4.15}$$

For the second part, it suffices to show that $\|\Delta_1\|_1 \lesssim \|\Omega\|_{L^q(\mathbb{S}^{n-1})}$. Since,

$$\|\Delta_1\|_1 = \int \left|R_j K_1 - \widetilde{K}_{j,1}\right| \mathrm{d}x \leq \underbrace{\int_{|x|<2} \left|R_j K_1(x)\right| \mathrm{d}x}_{(III)} + \underbrace{\int_{1<|x|<2} \left|\widetilde{K}_j(x)\right| \mathrm{d}x}_{(IV)} + \underbrace{\int_{|x|>2} |\Delta_1| \mathrm{d}x}_{(V)}, \tag{5.4.16}$$

and

$$\text{(III)} \overset{\text{Hölder}}{\lesssim} \left\|R_j K_1\right\|_{L^q} \overset{L^q\text{-boundedness of } R_j}{\lesssim} \|K_1\|_q \overset{\text{polar coord}}{\lesssim} \|\Omega\|_{L^q(\mathbb{S}^{n-1})}, \tag{5.4.17}$$

$$\text{(IV)} \overset{\text{the same as (I)}}{\lesssim} \|\Omega\|_{L^q(\mathbb{S}^{n-1})}, \tag{5.4.18}$$

$$\text{(V)} \overset{\text{by the proof of the boundedness of } R_jK_\varepsilon - R_jK_\nu}{\lesssim} \|\Omega\|_1 \int_{|x|>2} |x|^{-(n+1)} \mathrm{d}x \lesssim \|\Omega\|_{L^q(\mathbb{S}^{n-1})}, \tag{5.4.19}$$

thus we have finished the proof. □

Now we are ready to wrap up the lemmas above and prove the main theorem of this section:

Theorem 5.4.1. *Let Ω be an even function on $\mathbb{S}^{n-1}$ with $\Omega \in L^q(\mathbb{S}^{n-1})$ for some $q > 1$. Then the even singular integral operator is bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$.*

Proof. It suffices to prove the L^p boundedness for Schwartz functions f. By the algebraic identity, we have

$$T_\varepsilon f = K_\varepsilon * f = -\sum_j R_j R_j (K_\varepsilon * f) \overset{\text{Lemma 8}}{=} -\sum_j R_j((R_j K_\varepsilon) * f) = -\sum_j R_j((\widetilde{K}_{j,\varepsilon} + \Delta_\varepsilon) * f). \quad (5.4.20)$$

Note that, R_j is L^p bounded, we can pass everything through it and consider the two terms $\widetilde{K}_{j,\varepsilon} * f$ and $\Delta_\varepsilon * f$ separately. For the first term, since $\widetilde{K}_j$ is odd and homogeneous of degree $-n$, and is in $L^1(\mathbb{S}^{n-1})$ (Lemma 10). By the main theorem of the previous section, we know that the first term is strong-(p, p). For the second term, since the Δ_ε is $L^1(\mathbb{R}^n)$, by Young's convolution inequality and Lemma 10, we have

$$\|\Delta_\varepsilon * f\|_p \le \|\Delta_\varepsilon\|_1 \|f\|_p \lesssim \|\Omega\|_{L^q(\mathbb{S}^{n-1})} \|f\|_p. \quad (5.4.21)$$

Finally, since we have the pointwise limit $\lim_{\varepsilon\to 0} \Delta_\varepsilon * f = Tf$, we can conclude by the Fatou's lemma that

$$\|Tf\|_p = \left\| \liminf_{\varepsilon\to 0} |T_\varepsilon f|^p \right\|_1^{1/p} \le \liminf_{\varepsilon\to 0} \|T_\varepsilon f\|_p \lesssim \|f\|_p. \quad (5.4.22)$$

□

5.5 An operator algebra

We have already made it through the most technical part of the singular integral theory: the L^p boundedness of the singular integral operators:

$$Tf = \left(\text{p.v.} \frac{\Omega(x')}{|x|^n}\right) * f, \quad (5.5.1)$$

where Ω satisfies the right integrability (odd in L^1, even in L^q for some $q > 1$) and zero mean condition.

This section will be devoted to studying the composition of them. We have already attempted this in the previous section, i.e. the composition of Riesz transform and even singular integrals. But on the Fourier side, if we try to compose two multipliers without any regularity assumption, it can cause trouble. But with some smoothness assumption, we can express them in a nicer way. Another motivation of the theory is to study some constant-coefficient partial differential operators:

$$P(\zeta) = \sum_{|\alpha|\le m} b_\alpha \zeta^\alpha, \quad \zeta \in \mathbb{C}^n, \quad (5.5.2)$$

is a polynomial of degree m in n variables, we can define the corresponding differential operator

$$P(D)f := \sum_{|\alpha|\le m} b_\alpha D^\alpha f, \quad (5.5.3)$$

if f is a Schwartz function. By the Fourier transform, we have

$$\widehat{P(D)f}(\xi) = P(2\pi i\xi)\widehat{f}(\xi). \quad (5.5.4)$$

So if we try to study this family of operators, it all boils down to studying the "half-wave" operators:

$$\widehat{\Lambda f}(\zeta) = 2\pi|\zeta|\widehat{f}(\zeta). \tag{5.5.5}$$

Formally, we can write

$$\Lambda = \sqrt{-\triangle}. \tag{5.5.6}$$

If P is homogeneous of degree m, then we can write

$$P(D)f = T(\Lambda^m f), \tag{5.5.7}$$

where T is the singular integral operator, with multiplier (will be shown rigorously later)

$$\widehat{Tf}(\xi) = \mathrm{i}^m \frac{P(\xi)}{|\xi|^m}\widehat{f}(\xi). \tag{5.5.8}$$

and from which we see that the study of singular integrals can help us understand constant-coefficient partial differential operators. We will prove: $\Lambda^m f \in L^p$ implies $P(D)f \in L^p$. This is a quick corollary of the following theorem:

Theorem 5.5.1. *If $m \in C^\infty(\mathbb{R}^n \setminus \{0\})$ is homogeneous of degree 0, then the corresponding operator, defined by*

$$\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi), \tag{5.5.9}$$

is a singular integral operator up to a diagonal part. That is, there exists $a \in \mathbb{C}$ and $\Omega \in C^\infty(\mathbb{S}^{n-1})$ with zero mean, such that

$$Tf = af + \left(p.v.\frac{\Omega(x')}{|x|^n}\right) * f. \tag{5.5.10}$$

It suffices to show:

Lemma 11. *If m in the theorem also has zero mean on the sphere $\mathbb{S}^{n-1}$, then there exists $\Omega \in C^\infty(\mathbb{S}^{n-1})$ with zero mean, such that*

$$\widehat{m}(x) = p.v.\frac{\Omega(x')}{|x|^n}. \tag{5.5.11}$$

Remark 37. *In the preceding part of the section, we have been working on the case that somehow related: given an Ω, we can find the corresponding multiplier m. Here we are doing the reverse direction: given a multiplier m, we can find that it must come from some kernel Ω.*

Proof. Note that m is smooth, so $\widehat{m} \in \mathscr{S}^*$, and for every j we have

$$\widehat{\left(\frac{\partial^n m}{\partial x_j^n}\right)}(\xi) = C\xi_j^n\widehat{m}(\xi). \tag{5.5.12}$$

The function $\frac{\partial^n m}{\partial x_j^n}$ is also a function in $C^\infty(\mathbb{R}^n \setminus \{0\})$, and is a homogeneous tempered distribution of degree $-n$, which can be shown by direct computation.

Exercise: show that the n-th derivative of a degree 0 homogeneous function smooth away from 0 must have zero mean on the sphere. (Hint: test with a radial approximation of identity $\phi(x/R)$, pass on the derivative to ϕ and try to show some blow-up behavior as $R \to \infty$ if the mean is not zero.)

Hence, by homework 2.5, we know that

$$\frac{\partial^n m}{\partial x_j^n} = \text{p.v.}\frac{\partial^n m}{\partial x_j^n} + \sum_{|\alpha|\le K} c_\alpha D^\alpha \delta_0, \tag{5.5.13}$$

by homogeneity consideration, only the δ_0 (but not its derivatives) can appear on the right-hand side. Taking the Fourier transform, we have

$$C\xi_j^n \widehat{m}(\xi) = \text{p.v.}\widehat{\frac{\partial^n m}{\partial x_j^n}}(\xi) + c_0, \tag{5.5.14}$$

Therefore, we see that $\widehat{m}(\xi)$ coincides with a C^∞ homogeneous function of degree $-n$. (To see C^∞, one can consider the *spherical harmonics*, or consider *convolution on the rotation group*.)

Define $\widehat{m} = \Omega$ (away from 0). Claim: Ω has zero mean on the sphere. This can be seen by testing against a positive radial function supported on e.g. $1 < |x| < 2$. Then pass the Fourier transform to the test function, and use the fact that m has zero mean on the sphere.

Consider $\widehat{m} - \text{p.v.}\frac{\Omega(x')}{|x|^n}$, it is supported at $\{0\}$, and is homogeneous of degree 0, thus it must be a constant multiple of δ_0:

$$\widehat{m} = \text{p.v.}\frac{\Omega(x')}{|x|^n} + a\delta_0. \tag{5.5.15}$$

But by taking the Fourier transform again, we have

$$m(x) = a + \text{p.v.}\widehat{\frac{\Omega(x')}{|x|^n}}(x), \tag{5.5.16}$$

and both m and $\text{p.v.}\widehat{\frac{\Omega(x')}{|x|^n}}$ have zero mean on the sphere, so a must be 0. This finishes the proof. □

Theorem 5.5.2.

$$\mathcal{A} := \left\{T : \widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi), m \in C^\infty(\mathbb{R}^n \setminus \{0\}) \text{ homogeneous of degree } 0\right\} \tag{5.5.17}$$

Moreover, $T_m \in \mathcal{A}$ is invertible if and only if m is never zero.

Proof. $T_{m_1 m_2} = T_{m_1} \circ T_{m_2}$, and T_1 is the identity. □

Remark 38. *When T defined by the multiplier $m(\xi) = \mathrm{i}^m \frac{P(\xi)}{|\xi|^m}$ as in the beginning, then T is invertible in the algebra $\mathcal{A}$, then $P(D)u = f$ is equivalent to*

$$(-\triangle)^{m/2} u = \Lambda^m u = T^{-1} f. \tag{5.5.18}$$

So the $P(D)u = f$ boils down to studying the elliptic type equation $\Lambda^m u = g$, which is much easier to handle.

5.6 The singular integral with variable kernels

We will give the pseudo-differential operator a glimpse from the viewpoint of harmonic analysis in this section.

Motivation: the story of this section comes from the variable coefficient partial differential operator:

$$P(x,D)f=\sum_{|\alpha|=m} b_\alpha(x)D^\alpha f, \tag{5.6.1}$$

we can *almost* view this as a convolution. For f Schwartz, by the Fourier inversion formula, we have

$$P(x,D)f(x)=\int_{\mathbb{R}^n} P(x,2\pi\mathrm{i}\xi)\widehat{f}(\xi)\,\mathrm{d}\xi. \tag{5.6.2}$$

Remember that we will try to reduce the study of $P(x,D)$ to that of $\Lambda^m f$,

$$P(x,D)f=T(\Lambda^m f), \tag{5.6.3}$$

$$Tf(x)=\int_{\mathbb{R}^n}\sigma(x,\xi)\widehat{f}(\xi)e^{2\pi\mathrm{i}x\cdot\xi}\mathrm{d}\xi, \tag{5.6.4}$$

where

$$\sigma(x,\xi)=\frac{P(x,\mathrm{i}\xi)}{|\xi|^m} \tag{5.6.5}$$

is a homogeneous function of degree 0 in ξ for each fixed x.

At least formally, T is now a "convolution" operator with variable kernel:

$$Tf(x)=\int_{\mathbb{R}^n}K(x,x-y)f(y)\mathrm{d}y, \tag{5.6.6}$$

where

$$K(x,z)=\int_{\mathbb{R}^n}\sigma(x,\xi)e^{2\pi\mathrm{i}z\cdot\xi}\mathrm{d}\xi. \tag{5.6.7}$$

By the theory in the previous section, for fixed x,

$$K(x,z)=a(x)\delta_0(z)+\mathrm{p.v.}\frac{\Omega(x,z')}{|z|^n}, \tag{5.6.8}$$

which motivates us to study the variable kernel singular integral operator:

$$Tf(x)=\left(\mathrm{p.v.}\frac{\Omega(x,(x-y)')}{|x-y|^n}\right)*f(y)=\lim_{\varepsilon\to0}\int_{|y|>\varepsilon}f(x-y)\frac{\Omega(x,y')}{|y|^n}\mathrm{d}y, \tag{5.6.9}$$

Theorem 5.6.1. *Let $\Omega(x,y)$ be a function homogeneous in y of degree 0, such that*

- *(For this section, we only consider the odd case) For every fixed x, $\Omega(x,y)$ is an odd function in y with zero mean on $\mathbb{S}^{n-1}$.*
- *We have a maximal function control uniformly over x, i.e.*

$$\Omega^*(u):=\sup_x|\Omega(x,u)| \tag{5.6.10}$$

is in $L^1(\mathbb{S}^{n-1})$.

Then T is strong-(p,p) for all $1<p<\infty$.

Proof. The proof is nothing but a method of rotation.

$$Tf(x) = \frac{\pi}{2}\int_{\mathbb{S}^{n-1}} \Omega(x,u)H_u f(x)\mathrm{d}\sigma(u), \tag{5.6.11}$$

which is pointwise controlled by

$$|Tf(x)| \leq \frac{\pi}{2}\int_{\mathbb{S}^{n-1}} \Omega^*(u)|H_u f(x)|\mathrm{d}\sigma(u). \tag{5.6.12}$$

Then it follows from the L^p boundedness of H_u and Minkowski's integral inequality that

$$\|Tf\|_p \leq \frac{\pi}{2}\int_{\mathbb{S}^{n-1}} \Omega^*(u)\|H_u f\|_p \mathrm{d}\sigma(u) \lesssim \|f\|_p \int_{\mathbb{S}^{n-1}} \Omega^*(u)\mathrm{d}\sigma(u). \tag{5.6.13}$$

□

Sometimes letting the uniform control to be in L^1 is too strong, we can relax it to the following milder condition:

Theorem 5.6.2. *We keep the same odd condition, and replace the uniform L^1 control by the following condition:*

$$\sup_x \left(\int_{\mathbb{S}^{n-1}} |\Omega(x,u)|^q\right)^{1/q} \mathrm{d}\sigma(u) =: B_q < \infty. \tag{5.6.14}$$

for some $q > 1$. Then T is strong-(p,p) for all $q' \leq p < \infty$, where q' is the conjugate exponent of q.

Remark 39. *This condition only requires the "uniform integrability" of $\Omega(x,u)$ in x, which is much milder than the previous one.*

Proof. We apply Hölder to the method of rotation formula:

$$|Tf(x)| \leq \frac{\pi}{2}\left(\int_{\mathbb{S}^{n-1}} |\Omega(x,u)|^q \mathrm{d}\sigma(u)\right)^{1/q}\left(\int_{\mathbb{S}^{n-1}} |H_u f(x)|^{q'} \mathrm{d}\sigma(u)\right)^{1/q'} \leq \frac{\pi}{2}B_q\left(\int_{\mathbb{S}^{n-1}} |H_u f(x)|^{q'} \mathrm{d}\sigma(u)\right)^{1/q'}. \tag{5.6.15}$$

We integrate on both side to the power q', changing the order of integration, and apply the (q',q')-boundedness of the directional Hilbert transform to conclude that

$$\int |Tf|^{q'} \lesssim B_q^{q'} \int_{\mathbb{S}^{n-1}}\int_{\mathbb{R}^n} |H_u f(x)|^{q'} \mathrm{d}x\mathrm{d}\sigma(u) \lesssim B_q^{q'} \|f\|_{q'}^{q'}. \tag{5.6.16}$$

Thus the operator is strong-(q',q'). The general case follows by twisting the exponent and we can still apply Hölder to conclude. □

Chapter 6

Singular Integrals (Part II)

Last chapter: use methods of rotations + Riesz transforms to prove L^p boundedness of singular integral operators. But the methods of rotations are in general very rigid, in the sense that:

- In the case of singular intgral with variable kernels, we can only handle the perfectly odd kernels.
- It is very hard to prove weak-$(1,1)$ boundedness by this method.

In this chapter we will try to develop a more abstract but general theory of singular integrals, which can handle (1) weak-$(1,1)$ boundedness, (2) general singular integral operators (not necessarily a convolution operator). This will be another time when the Calderón-Zygmund decomposition technique comes to the stage. (Recall that we have used it to prove the weak-$(1,1)$ boundedness of the Hardy-Littlewood maximal function.)

6.1 The Calderón-Zygmund theorem

Theorem 6.1.1 (Calderón-Zygmund theorem: an axiomatic theorem for the boundedness of singular integrals)**.** *Let $K \in \mathscr{S}^*$ be a locally integrable function away from 0, satisfying the following conditions:*

- $\left|\widehat{K}(\xi)\right| \leq A$ *for some constant $A > 0$.*
- *(Hörmander condition) There exists $B > 0$ such that for all $R > 0$,*

$$\int_{|x|>2|y|} |K(x-y) - K(x)| \mathrm{d}x \leq B. \tag{6.1.1}$$

  Remark 40. *This condition is some sort of smoothness condition of K away from 0. For example, in the case of Hilbert transform, we have*

$$\left|\frac{1}{x-y} - \frac{1}{x}\right| = \frac{y}{x(x-y)} \lesssim \frac{1}{x^2}, \quad \textit{for } |x| > 2|y|, \tag{6.1.2}$$

Then the singular integral operator $Tf := K * f$ is strong-(p,p) for all $1 < p < \infty$, and weak-$(1,1)$.

Remark 41. *Totally similar to the proof of the case of Hilbert transform.*

Outline. Using the Calderón-Zygmund decomposition, we can decompose $f = g + b$, and by Hölder's inequality, we have for the good part g,

$$\|g\|_2 \leq \|g\|_\infty^{\frac{1}{2}} \|g\|_1^{\frac{1}{2}} \lesssim \alpha^{\frac{1}{2}} \|f\|_1^{\frac{1}{2}}, \tag{6.1.3}$$

where, since $\widehat{K}$ is bounded, we have that T is strong-$(2,2)$, thus

$$m\left(\left\{x : |Tg| > \frac{\alpha}{2}\right\}\right) \lesssim \frac{1}{\alpha^2}\|g\|_2^2 \lesssim \frac{1}{\alpha}\|f\|_1. \tag{6.1.4}$$

For the bad part $b = \sum_j b_j$ with b_j supported on Q_j, we have the vanishing mean property $\int b_j = 0$. For the part inside $\bigcup_j 2Q_j$, we have

$$\left|\bigcup_j 2Q_j\right| \lesssim \sum_j |Q_j| \overset{\text{Calderón-Zygmund decomposition}}{\lesssim} \frac{1}{\alpha}\|f\|_1. \tag{6.1.5}$$

For the part outside $\bigcup_j 2Q_j$, by the Hörmander condition, we have

$$\begin{aligned}
\int_{\mathbb{R}^n \setminus \bigcup_j 2Q_j} |Tb(x)| \mathrm{d}x &\leq \sum_j \int_{\mathbb{R}^n \setminus 2Q_j} \left|Tb_j(x)\right| \mathrm{d}x \overset{\text{vanishing mean}}{=} \sum_j \int_{\mathbb{R}^n \setminus 2Q_j} \left|\int_{Q_j} (K(x-y) - K(x)) b_j(y) \mathrm{d}y\right| \mathrm{d}x \\
&\overset{\text{Fubini}}{\leq} \sum_j \int_{Q_j} \left(\int_{|x-c_j| > 2\sqrt{n} l(Q_j)} |K(x-y) - K(x)| \mathrm{d}x\right) \left|b_j(y)\right| \mathrm{d}y \\
&\leq B \sum_j \left\|b_j\right\|_1 \overset{\text{Calderón-Zygmund decomposition}}{\lesssim} B\|f\|_1.
\end{aligned} \tag{6.1.6}$$

Here, c_j is the center of Q_j, and $l(Q_j)$ is the side length of Q_j. Thus by Chebyshev's inequality, we pass to the superlevel set estimate to conclude that

$$\begin{aligned}
|\{x : |Tf(x)| > \alpha\}| &\leq \left|\left\{x : |Tg(x)| > \frac{\alpha}{2}\right\}\right| + \left|\left\{x : |Tb(x)| > \frac{\alpha}{2}\right\}\right| \\
&\leq \left|\left\{x : |Tg(x)| > \frac{\alpha}{2}\right\}\right| + \left|\bigcup_j 2Q_j\right| + \left|\left\{x \in \mathbb{R}^n \setminus \bigcup_j 2Q_j : |Tb(x)| > \frac{\alpha}{2}\right\}\right| \\
&\lesssim \frac{1}{\alpha}\|f\|_1 + \frac{2}{\alpha}\int_{\mathbb{R}^n \setminus \bigcup_j 2Q_j} |Tb(x)| \mathrm{d}x \lesssim \frac{1}{\alpha}\|f\|_1.
\end{aligned} \tag{6.1.7}$$

Therefore, the operator is weak-$(1,1)$. The strong-(p,p) boundedness for $1 < p \leq 2$ follows from the Marcinkiewicz interpolation theorem. Other cases follow from duality. □

Remark 42. *A widely considered sufficient condition for the Hörmander condition is the following gradient estimate:*

$$|\nabla K(x)| \leq \frac{C}{|x|^{n+1}}, \tag{6.1.8}$$

away from zero, then the Hörmander condition is satisfied by applying the MVT:

$$\begin{aligned}
\int_{|x|>2|y|} |K(x-y) - K(x)| \mathrm{d}x &= \int_{|x|>2|y|} \left|\int_0^1 \nabla K(x - ty) \cdot (-y) \mathrm{d}t\right| \mathrm{d}x \leq |y| \int_0^1 \int_{|x|>2|y|} |\nabla K(x-ty)| \mathrm{d}x \mathrm{d}t \\
&\lesssim \int_0^1 \int_{|x|>2|y|} \frac{|y|}{|x|^{n+1}} \mathrm{d}x \mathrm{d}t \lesssim \int_{2|y|}^\infty \frac{|y|}{r^{n+1}} r^{n-1} \mathrm{d}r = \frac{|y|}{2|y|} = \frac{1}{2}.
\end{aligned} \tag{6.1.9}$$

An application of the abstract Calderón-Zygmund theorem to the rough kernel singular integral operator is the "Dini's criterion":

Theorem 6.1.2 (Dini's criterion). *For $\Omega \in L^1(\mathbb{S}^{n-1})$ homogeneous of degree $-n$ with zero mean, define the modulus of continuity*

$$\omega_\infty(t) = \sup_{\substack{x',y'\in\mathbb{S}^{n-1}\\ |x'-y'|<t}} |\Omega(x') - \Omega(y')| \tag{6.1.10}$$

satisfies

$$\int_0^1 \frac{\omega_\infty(t)}{t}\mathrm{d}t < \infty, \tag{6.1.11}$$

then the singular integral operator defined by p.v.$\frac{\Omega(x')}{|x|^n}*$ satisfies the Hörmander condition.

Remark 43. *This will be satisfied if Ω is α-Hölder continuous for some $\alpha > 0$ on the sphere. Quite similar to the Dini's criterion for the pointwise convergence of Fourier series.*

Proof. When $|x| > 2|y|$, we have

$$\begin{aligned}|K(x-y) - K(x)| &= \left|\frac{\Omega((x-y)')}{|x-y|^n} - \frac{\Omega(x')}{|x|^n}\right| \leq \left|\frac{\Omega((x-y)') - \Omega(x')}{|x-y|^n}\right| + |\Omega(x')|\left|\frac{1}{|x|^n} - \frac{1}{|x-y|^n}\right| \\ &\overset{\text{observation } |(x-y)'-x'|\leq 4\frac{|y|}{|x|}}{\lesssim} \omega_\infty\left[C\cdot\frac{|y|}{|x|}\right]\cdot|x|^{-n} + |\Omega(x')|\cdot\frac{|y|}{|x|^{n+1}}.\end{aligned} \tag{6.1.12}$$

Integral this with polar coordinates, we get

$$\begin{aligned}\int_{|x|>2|y|} |K(x-y) - K(x)|\mathrm{d}x &\lesssim \int_{|x|>2|y|} |x|^{-n}\omega_\infty\left(C\cdot\frac{|y|}{|x|}\right)\mathrm{d}x \\ &\overset{\text{polar coord.}}{\lesssim} \int_{2|y|}^\infty \frac{1}{r}\omega_\infty\left(C\frac{|y|}{r}\right)\mathrm{d}r \\ &\overset{\text{change of variable } s=1/r,\text{ invariance of the Haar measure}}{=} \int_0^{1/2|y|} \frac{\omega_\infty(C|y|s)}{s}\mathrm{d}s \\ &\lesssim \int_0^1 \frac{\omega_\infty(t)}{t}\mathrm{d}t < \infty.\end{aligned} \tag{6.1.13}$$

$\square$

Remark 44. *For the multiplicative group $\mathbb{R}_+^\times$, the Haar measure is given by*

$$\mathrm{d}^\times r = \frac{\mathrm{d}r}{r}, \tag{6.1.14}$$

it is invariant under:

- *Scaling: for any $a > 0$,*

$$\int_{\mathbb{R}_+^\times} f(ar)\mathrm{d}^\times r = \int_{\mathbb{R}_+^\times} f(r)\mathrm{d}^\times r. \tag{6.1.15}$$

- *Inversion:*

$$\int_{\mathbb{R}_+^\times} f\left(\frac{1}{r}\right)\mathrm{d}^\times r = \int_{\mathbb{R}_+^\times} f(r)\mathrm{d}^\times r. \tag{6.1.16}$$

Corollary 15. *If Ω is a function on $\mathbb{S}^{n-1}$ with zero mean satisfying the Dini's criterion, then*

$$Tf = \left(p.v. \frac{\Omega(x')}{|x|^n}\right) * f \tag{6.1.17}$$

is strong-(p,p) for all $1 < p < \infty$ and weak-$(1,1)$.

Proof. We easily see that Ω is bounded on the sphere, thus it is both L^1 and L^q integrable for some $q > 1$ on $\mathbb{S}^{n-1}$. In the previous section we have the formula of the Fourier transform of the kernel:

$$\widehat{K}(\xi) = \widehat{K}_o(\xi) + \widehat{K}_e(\xi) = -\mathrm{i}\frac{\pi}{2}\int_{\mathbb{S}^{n-1}} \Omega_o(u)\,\mathrm{sgn}(u\cdot\xi')\mathrm{d}\sigma(u) + \int_{\mathbb{S}^{n-1}} \Omega_e(u)\log\frac{1}{|u\cdot\xi'|}\mathrm{d}\sigma(u). \tag{6.1.18}$$

Using the L^1 integrability of Ω_o and the L^q integrability of Ω_e, we can use Hölder's inequality to conclude that $\widehat{K}(\xi)$ is bounded. By Calderón-Zygmund theorem and Dini's criterion, we conclude the proof. □

Remark 45. *The Dini's criterion is very strong and directly implies the boundedness of the kernel $\widehat{K}$, and this directly implies the strong-(p,p) by the previous chapter. However, what is new here is the weak-$(1,1)$ boundedness.*

Remark 46. *From this example, we also see that we strongly used the expression of $\widehat{K}$, if K is given by the principal value integral. But in general how can we deal with other operators?*

6.2 Truncated integrals and the principal value

In the previous section (in fact, as well as the previous chapter), we have already seen that for singular integral operators with Ω being sufficiently regular, the multiplier $\widehat{K}$ will be bounded. This can also be characterized by some property of the integral kernel K.

Proposition 15 (The truncated kernels are bounded)**.** *If K is locally integrable away from 0, such that*

- $\left|\int_{a<|x|<b} K(x)\mathrm{d}x\right| \le A$ *for all $a, b > 0$. (cancellation property)*
- $\int_{a<|x|<2a} |K(x)|\mathrm{d}x \le B$ *for all $a > 0$. (some technical regularity that mitigates the error term)*
- $\int_{|x|>2|y|} |K(x-y) - K(x)|\mathrm{d}x \le C$ *for all $y \in \mathbb{R}^n$. (Hörmander or smooth condition)*

Then:

$$\left|\widehat{K}_{\varepsilon,R}\right| \le M, \tag{6.2.1}$$

where M is independent of $\varepsilon, R > 0$, and $K_{\varepsilon,R} = K \cdot \int_{\varepsilon<|x|<R}$ is the truncated kernel.

Proof. We first consider the cases where $\varepsilon < 1/|\xi| < R$, then

$$\widehat{K_{\varepsilon,R}} = \int_{\varepsilon<|x|<R} K(x)e^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{d}x = \underbrace{\int_{\varepsilon<|x|<1/|\xi|} K(x)e^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{d}x}_{\mathrm{I}} + \underbrace{\int_{1/|\xi|<|x|<R} K(x)e^{-2\pi\mathrm{i}x\cdot\xi}\mathrm{d}x}_{\mathrm{II}}. \tag{6.2.2}$$

For the first part, by the cancellation property, we have

$$|(\mathrm{I})| \le \left|\int_{\varepsilon<|x|<1/|\xi|} K(x)\mathrm{d}x\right| + \left|\int_{\varepsilon<|x|<1/|\xi|} K(x)(1-e^{-2\pi\mathrm{i}x\cdot\xi})\mathrm{d}x\right| \lesssim A + \int_{\varepsilon<|x|<1/|\xi|} |K(x)|\cdot 2\pi|x\cdot\xi|\mathrm{d}x. \tag{6.2.3}$$

By the cake-layer estimate, we have

$$|(\mathrm{I})| \leq A + \sum_{2^j \in (\varepsilon/100, 100|\xi|^{-1})} \int_{2^j<|x|<2^{j+1}} |K(x)|\mathrm{d}x \cdot 2^j|\xi| \lesssim A + B. \tag{6.2.4}$$

For the second part, we consider a small shift of the integral variable:

$$(\mathrm{II}) = -\int_{|\xi|^{-1}<|x-z|<R} K(x-z)e^{-2\pi \mathrm{i}(x-z)\cdot\xi}\mathrm{d}x = -\int_{|\xi|^{-1}<|x|<R} K(x-z)e^{-2\pi \mathrm{i}x\cdot\xi}\mathrm{d}x, \tag{6.2.5}$$

with $z = $ e.g. $\frac{1}{2}\xi^{-1}$.

Then we have

$$\begin{aligned}
|(\mathrm{II})| &= \frac{1}{2}|(\mathrm{II}) + (\mathrm{II})| = \frac{1}{2}\left|\int_{|\xi|^{-1}<|x|<R} K(x)e^{-2\pi \mathrm{i}x\cdot\xi}\mathrm{d}x - \int_{|\xi|^{-1}<|x|<R} K(x-z)e^{-2\pi \mathrm{i}x\cdot\xi}\mathrm{d}x\right| \\
&\lesssim \int_{|\xi|^{-1}<|x|<R} |K(x)-K(x-z)|\mathrm{d}x + \int_{\frac{1}{2}|\xi|^{-1}<|x|<\frac{3}{2}|\xi|^{-1}} |K(x)|\mathrm{d}x + \int_{R-\frac{1}{2}|\xi|^{-1}<|x|<R+\frac{1}{2}|\xi|^{-1}} |K(x)|\mathrm{d}x \\
&\lesssim C + B + B.
\end{aligned} \tag{6.2.6}$$

Here, the latter two terms come from the "trash areas" that are not overlapped after the shift. Combining the estimates of (I) and (II), we conclude the proof in the case $\varepsilon < 1/|\xi| < R$. For other case, we only have one of (I) or (II), and the estimate is even easier.

□

Now we can apply the Calderón-Zygmund theorem to the truncated singular integral operator,

Corollary 16. *If K satisfies the conditions in the previous proposition, then the truncated singular integral operator $T_{\varepsilon,R}f = K_{\varepsilon,R} * f$ is strong-(p,p) for all $1 < p < \infty$ and weak-$(1,1)$.*

Proof. The previous proof implies the boundedness of the multiplier, and thus implies the strong-$(2,2)$ boundedness. The Hörmander condition is already assumed. Thus by Calderón-Zygmund theorem, we conclude the proof. □

Now the remaining issue is what happens at the origin (the $R \to \infty$ limit is in general no problem, at least in the weak sense). For this, we need an additional condition which is actually both necessary and sufficient.

Theorem 6.2.1. *Given a function satisfying the second condition in the previous proposition, then the principal value p.v.K exists, i.e.*

$$\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} K(x)f(x)\mathrm{d}x \quad \textit{exists for all } f \in \mathscr{S}, \tag{6.2.7}$$

if and only if

$$\lim_{\varepsilon\to 0}\int_{\varepsilon<|x|<1} K(x)\mathrm{d}x \quad \textit{exists}. \tag{6.2.8}$$

Proof. This is actually nothing but what we have played with in the Hilbert transform case. We only need to test the $\phi \equiv 1$ on $B(0,1)$ to see the necessity. For the sufficiency, we can write

$$\int_{|x|>\varepsilon} K\phi = \underbrace{\int_{|x|>1} K\phi}_{\text{(I)}} + \underbrace{\int_{\varepsilon<|x|<1} K(x)(\phi(x)-\phi(0))\mathrm{d}x}_{\text{(II)}} + \phi(0)\underbrace{\int_{\varepsilon<|x|<1} K(x)\mathrm{d}x}_{\text{(III)}}. \tag{6.2.9}$$

(I) is absolutely convergent since ϕ is Schwartz. (III) converges by assumption. For (II), by the MVT, we have

$$|\text{(II)}| \leq \int_{\varepsilon<|x|<1} |K(x)| \cdot \||x|\nabla\phi\|_\infty \mathrm{d}x, \tag{6.2.10}$$

which converges as $\varepsilon \to 0$ by the condition on K. □

Corollary 17. *If the integral kernel K satisfies the three conditions in the previous proposition, together with the existence of the limit*

$$\lim_{\varepsilon\to 0} \int_{\varepsilon<|x|<1} K(x)\mathrm{d}x, \tag{6.2.11}$$

then

$$Tf(x) = \lim_{\varepsilon\to 0} \int_{|y|>\varepsilon} K(x-y)f(y)\mathrm{d}y \tag{6.2.12}$$

is strong-(p,p) for all $1 < p < \infty$ and weak-$(1,1)$. (first define it for Schwartz functions, then extend to L^p functions by density)

Example 19 ($K(x) = |x|^{-n-\mathrm{i}t}$)**.** *Briefly speaking, using the previous results, we can see that the truncated kernel singular integral operator is both strong-(p,p) and weak-$(1,1)$. However, when we try to define the principal value integral for kernels with complex homogeneity, we will need to choose a suitable sequence to approach 0, and the resulting p.v.K will not be a homonegeneous tempered distribution anymore. However, by considering the analytic continuation definition of the Fourier transform, we can find a "canonical" definition of this singular integral operator in the sense of homogeneity, which differs from the arbitrariness defined principal value integral by a multiple of the delta distribution at 0.*

We consider the kernel with complex homogeneity:

$$K(x) = \frac{1}{|x|^{n+\mathrm{i}t}}, \quad t \in \mathbb{R} \setminus \{0\}. \tag{6.2.13}$$

By direct computation, we note that K is locally integrable away from 0 and satisfies the three conditions in Proposition 15. Therefore, the truncated singular integral operator $K_{\varepsilon,R}*$ is strong-(p,p) for all $1 < p < \infty$ and weak-$(1,1)$. However, if we want to define the principal value integral, we need to check the existence of the limit eq. (6.2.8)

$$\int_{\varepsilon<|x|<1} \frac{1}{|x|^{n+\mathrm{i}t}}\mathrm{d}x = \int_\varepsilon^1 r^{-\mathrm{i}t-1}\mathrm{d}r \cdot \int_{\mathbb{S}^{n-1}} \mathrm{d}\sigma = \sigma(\mathbb{S}^{n-1}) \cdot \frac{1-\varepsilon^{-\mathrm{i}t}}{\mathrm{i}t}, \tag{6.2.14}$$

which unfortunately does not converge. But if we choose an appropriate sequence $\varepsilon_k \to 0$, e.g. $\varepsilon_k = e^{-2\pi k/t}$, then the contribution $\int_{\varepsilon_k<|x|<1} K(x)\mathrm{d}x$ vanishes along this sequence. Thus the limit exists along this sequence, by the trick of subtracting the zero mean part:

$$\lim_{k\to\infty, R\to\infty} K_{\varepsilon_k,R} * f(x) = \int_{|y|<1} \frac{f(x-y)-f(x)}{|y|^{n+\mathrm{i}t}}\mathrm{d}y + \int_{|y|>1} \frac{f(x-y)}{|y|^{n+\mathrm{i}t}}\mathrm{d}y. \tag{6.2.15}$$

This makes sense for Schwartz functions so we can define an integral operator as follows (we still call it the principal value, although it actually depends on the choice of the sequence):

$$p.v. \frac{1}{|x|^{n+\mathrm{i}t}}(\phi) = \int_{|x|<1} \frac{\phi(x) - \phi(0)}{|x|^{n+\mathrm{i}t}} \mathrm{d}x + \int_{|x|>1} \frac{\phi(x)}{|x|^{n+\mathrm{i}t}} \mathrm{d}x. \tag{6.2.16}$$

So far so good, but the issue is now this distribution is NOT homogeneous of order $-n - \mathrm{i}t$*. What is missing here? We consider an indirect way to see the Fourier transform. We consider the kernel* $\frac{1}{|x|^z}$ *for* $\operatorname{Re} z < n$*, and try to define a tempered distribution homogeneous of order* $-z$ *directly. Again we test for* $\phi \in \mathscr{S}$*,*

$$(1/|x|^z)(\phi) = \int_{|x|<1} \frac{\phi(x) - \phi(0)}{|x|^z} \mathrm{d}x + \int_{|x|>1} \frac{\phi(x)}{|x|^z} \mathrm{d}x + \frac{\sigma(\mathbb{S}^{n-1})}{n - z} \phi(0). \tag{6.2.17}$$

Very interestingly, this definition makes perfect sense for any $0 < \operatorname{Re} z < n + 1$ *unless* $z = n$*. Thus this gives a meromorphic function in the strip with a simple pole* $z = n$*. If we let* $z = n + \mathrm{i}t$*, then away from the pole* $(t = 0)$*, we have*

$$(1/|x|^z)\,(\phi)|_{z=n+\mathrm{i}t} = p.v. \frac{1}{|x|^{n+\mathrm{i}t}}(\phi) - \frac{1}{\mathrm{i}t}\langle \delta_0, \phi\rangle, \tag{6.2.18}$$

which means that they ***differ by a multiple of the delta distribution at*** 0 *(this also implies that the analytic definition is canonical in the sense that it is the unique distribution that is homogeneous of degree* $-n - \mathrm{i}t$ *that coincides away from the origin with the locally integrable function* $K(x) = \frac{1}{|x|^{n+\mathrm{i}t}}$*). Thus using the Fourier transform of the homogeneous distribution* $|x|^{-z}$ *(see example 17), we get the Fourier transform of eq.* (6.2.16)

$$\left(p.v. \frac{1}{|x|^{n+\mathrm{i}t}}\right)^{\widehat{}} (\xi) = \pi^{\frac{n}{2}+\mathrm{i}t} \frac{\Gamma(-\mathrm{i}t/2)}{\Gamma((n + \mathrm{i}t)/2)} |\xi|^{\mathrm{i}t} + \frac{1}{\mathrm{i}t} \sigma(\mathbb{S}^{n-1}). \tag{6.2.19}$$

6.3 Standard kernels and generalized Calderón-Zygmund operators

In fact, we can generalized the Calderón-Zygmund theorem to non-convolution type singular integral operators. The main theorem is the following:

Theorem 6.3.1. *Let* T *be an* L^2*-bounded operator (which is analogous to the boundedness of* $\widehat{K}$ *in the convolution case, at the very beginning of the Calderón-Zygmund theorem).*

Let K *be a function on* $\mathbb{R}^n \times \mathbb{R}^n$ *away from the diagonal* $\Delta\{(x, x) : x \in \mathbb{R}^n\}$*, such that*

(1) *T is given by the following integral representation for all* $f \in L^2(\mathbb{R}^n)$ *compactly supported and for* $x \notin \operatorname{supp} f$*:*

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y) \mathrm{d}y, \tag{6.3.1}$$

(2) *T satisfies the following generlized Hörmander condition:*

$$\begin{cases} \int_{|x-y|>2|z-y|} |K(x, y) - K(x, z)| \mathrm{d}x \le C, \\ \int_{|x-y|>2|x-w|} |K(x, y) - K(w, y)| \mathrm{d}x \le C. \end{cases} \tag{6.3.2}$$

(2′) *Sometimes we can replace the generalized Hörmander condition by the following standard kernel condition. We say that K is a standard kernel if there exists $C > 0$ such that for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \setminus \Delta$,*

$$\begin{cases} |K(x,y)| \le \frac{C}{|x-y|^n}, \\ |K(x,y) - K(x,z)| \le C\frac{|y-z|^\delta}{|x-y|^{n+\delta}}, \quad \text{for } |x-y| > 2|y-z|, \\ |K(x,y) - K(w,y)| \le C\frac{|x-w|^\delta}{|x-y|^{n+\delta}}, \quad \text{for } |x-y| > 2|x-w|, \end{cases} \tag{6.3.3}$$

Remark 47. *It is spiritually similar to the case of convolution kernel satisfying the α-Hölder continuity condition, which implies the Hörmander condition.*

Then, T is strong-(p, p) for all $1 < p < \infty$ and weak-$(1, 1)$.

Example 20. • ***The Cauchy integral along a Lipschitz curve.*** *A Lipschitz function, $\Gamma : \mathbb{R} \to \mathbb{C}$ defined by $\Gamma(t) = t + \mathrm{i}A(t)$ a plane curve. With this parametrization, we can view any function f defined on Γ as a function of t. The Cauchy integral operator along Γ, can be viewed as a singular integral operator that makes sense in the open set on one side of the curve:*

$$C_\Gamma f(z) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma \frac{f(\zeta)}{\zeta - z} \mathrm{d}\zeta \overset{\text{parametrization}}{=} \frac{1}{2\pi \mathrm{i}} \int_{\mathbb{R}} \frac{f(t)}{t + \mathrm{i}A(t) - z} (1 + \mathrm{i}A'(t))\mathrm{d}t. \tag{6.3.4}$$

Here A' exists almost everywhere and is an L^∞ fuction. $C_\Gamma f$ makes perfect sense on

$$\Omega_+ = \{x + \mathrm{i}y : y > A(x)\}. \tag{6.3.5}$$

What about the boundary values on Γ? It is given by the following principal value integral (a limit of truncated integrals):

$$\frac{1}{2}\left[f(x) + \frac{\mathrm{i}}{\pi} \lim_{\varepsilon \to 0} \int_{|x-t|>\varepsilon} \frac{f(t)}{x - t + \mathrm{i}(A(x) - A(t))} (1 + \mathrm{i}A'(t))\mathrm{d}t \right] \tag{6.3.6}$$

which can be viewed as a singular integral operator with kernel

$$K(x,y) = \frac{1}{\pi} \cdot \frac{1 + \mathrm{i}A'(y)}{x - y + \mathrm{i}(A(x) - A(y))}. \tag{6.3.7}$$

This kernel is a standard kernel with $\delta = 1$, thus by the generalized Calderón-Zygmund theorem, the Cauchy integral operator is strong-(p, p) for all $1 < p < \infty$ and weak-$(1, 1)$.

• ***The Calderón commutator.***

Definition 6.3.2. *We say that a linear operator T is a generalized Calderón-Zygmund operator (CZO), if*

(1) T is bounded on $L^2(\mathbb{R}^n)$.

(2) There exists a standard kernel K such that for all $f \in L^2(\mathbb{R}^n)$ compactly supported and for $x \notin$ supp f, Tf is given by the integral representation

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y) f(y) \mathrm{d}y. \tag{6.3.8}$$

6.4 Calderón-Zygmund singular integrals

At the very end of the previous chapter we defined the CZO to be the singular integral operator with standard kernel and for $f \in L^2(\mathbb{R}^n)$ compactly supported and for $x \notin \operatorname{supp} f$. Just as in the singular convolution operator case, we also want to ask: does

$$Tf(x) := \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} K(x,y) f(y) \mathrm{d}y \tag{6.4.1}$$

make sense? We have already seen that (1) the limit may not exist (e.g. the complex homogeneity kernel case); (2) even if the limit exists, the identity may not hold. An easy example is, I is a CZO with $K(x,y) \equiv 0$, but clearly $If = 0$ if $x \notin \operatorname{supp} f$, while the limit $\lim_{\varepsilon \to 0} T_\varepsilon f(x)$ is just 0.

Nonetheless, we have the similar property as Theorem 6.2.1

Proposition 16. *$\lim_{\varepsilon \to 0} T_\varepsilon f(x)$ exists for every $f \in C_c^\infty$, iff $\lim_{\varepsilon \to 0} \int_{\varepsilon < |x-y| < 1} K(x,y) \mathrm{d}y$ exists for almost every $x \in \mathbb{R}^n$.*

Proof. Identical to the proof of Theorem 6.2.1. □

Proposition 17. *If two CZOs are associated with the same standard kernel K, then there difference is a pointwise multiplication operator by an L^∞ function.*

Proof. Let T_1, T_2 be two CZOs associated with the same standard kernel K. Define $S = T_1 - T_2$, which is also L^2-bounded, since T_1 and T_2 are both L^2-bounded. Note that for any $f \in L^2(\mathbb{R}^n)$ compactly supported and for $x \notin \operatorname{supp} f$, we have $Sf(x) = \int_{\mathbb{R}^n} (K(x,y) - K(x,y)) f(y) \mathrm{d}y = 0$. We define

$$\mu(E) := \langle S\chi_E, \chi_E \rangle, \quad E \subset \mathbb{R}^n \text{ measurable and of finite measure.} \tag{6.4.2}$$

Then by the L^2-boundedness of S, we have

$$|\mu(E)| \leq \|S\|_{L^2 \to L^2} |E|. \tag{6.4.3}$$

Therefore, μ is absolutely continuous with respect to the Lebesgue measure, thus by the Radon-Nikodym theorem, there exists b a measurable function such that for all measurable set E with finite measure,

$$\mu(E) = \int_E b(x) \mathrm{d}x. \tag{6.4.4}$$

By Lebesgue differentiation theorem, we have for almost every $x \in \mathbb{R}^n$,

$$|b(x)| \leq \limsup_{Q \ni x,\ |Q| \to 0} \frac{\left| \int_Q b \right|}{|Q|} \leq \limsup_{Q \ni x,\ |Q| \to 0} \frac{\|S\|_{L^2 \to L^2} |Q|}{|Q|} = \|S\|_{L^2 \to L^2}. \tag{6.4.5}$$

Therefore, $b \in L^\infty$.

We next show that Sf is given by the multiplication operator $b(x)f(x)$ for any $f \in L^2$ a.e. x. We claim that for any E, F measurable with finite measure, we have

$$S(\chi_E \chi_F) = \chi_F S(\chi_E). \tag{6.4.6}$$

This is because for a.e. $x \notin E$, both sides are 0 since $x \notin \operatorname{supp} \chi_E$ and $\notin \operatorname{supp} \chi_E \chi_F$. For a.e. $x \in E \setminus F$, both sides are also 0 since $\chi_F(x) = 0$. For a.e. $x \in E \cap F$, both sides are equal to $S(\chi_E)(x)$ since $\chi_F(x) = 1$. Thus the claim holds. Hence,

$$\langle S(\chi_E), \chi_F \rangle = \int \chi_F(x) S(\chi_E(x)) = \langle S(\chi_E \chi_F), \chi_E \chi_F \rangle \overset{\text{by def. of } \mu}{=} \int_{E \cap F} b(x) \mathrm{d}x = \int_{\mathbb{R}^n} b(x) \chi_E(x) \chi_F(x) \mathrm{d}x. \tag{6.4.7}$$

Therefore for any simple functions $\widetilde{f}$ and $\widetilde{g}$, we have by linearity

$$\langle S(\widetilde{f}), \widetilde{g} \rangle = \int_{\mathbb{R}^n} b(x) \widetilde{f}(x) \overline{\widetilde{g}(x)} \mathrm{d}x. \tag{6.4.8}$$

Since simple functions are dense in L^2, by the L^2-boundedness of S, we conclude that for any $f, g \in L^2$,

$$\langle S(f), g \rangle = \int_{\mathbb{R}^n} b(x) f(x) \overline{g(x)} \mathrm{d}x = \langle bf, g \rangle, \tag{6.4.9}$$

thus $Sf = b(x) f(x)$ for a.e. x. $\square$

Remark 48. *We see that, we still strongly use the L^2-boundedness of the operator.*

Definition 6.4.1 (Calderón-Zygmund singular integral)**.** *We call a CZO T a Calderón-Zygmund singular integral (should not be confused with CZO itself), if for any Schwartz function $f \in \mathscr{S}$, we have*

$$\lim_{\varepsilon \to 0} T_\varepsilon f(x) = Tf(x). \tag{6.4.10}$$

Remark 49. *Again, the existence of the limit does not guarantee that the identity holds. Thus this definition is non-trivial.*

We will do the same technique as the Hilbert transform case to prove that this is also an a.e. pointwise convergence for $f \in L^p$, $1 \leq p < \infty$. Again, the key is to use the maximal function and the Cotlar-type inequality.

Theorem 6.4.2. *If T os a Calderón-Zygmund singular integral, then T^* is weak-$(1,1)$ and strong-(p,p) for all $1 < p < \infty$.*

Lemma 12 (Cotlar-type inequality)**.** *If T is a CZO, then for any $\nu \in (0,1)$ and $f \in C_c^\infty$,*

$$T^* f(x) \leq C_\nu [(M|Tf|^\nu)(x)^{\frac{1}{\nu}} + Mf(x)]. \tag{6.4.11}$$

Lemma 13 (Kolmogorov inequality)**.** *Given a weak-$(1,1)$-type operator S, then for any $0 < \nu < 1$, there exists $C_\nu > 0$ such that for any measurable set E with finite measure and for any $f \in L^1$*

$$\int_E |Sf(x)|^\nu \mathrm{d}x \leq C_\nu |E|^{1-\nu} \|f\|_{L^1}^\nu. \tag{6.4.12}$$

The proof is an easy application of the layer-cake representation and the definition of weak-$(1,1)$ boundedness.

Proof of the Cotlar-type inequality. WLOG we translate x to 0. For $\varepsilon > 0$, we let $Q = B(0, \varepsilon/2)$ and let $f_1 = f \cdot \chi_{2Q}$ and $f_2 = f - f_1$. We know that $T_\varepsilon f(0) = Tf_2(0)$ by definition of the trucated operator. For $z \in Q$,

$$|Tf_2(z) - Tf_2(0)| = \left|\int_{|y|>\varepsilon} (K(z,y) - K(0,y))f(y)\mathrm{d}y\right| \overset{\text{std. ker.}}{\lesssim} |z|^\delta \int_{|y|>\varepsilon} \frac{|f(y)|}{|y|^{n+\delta}}\mathrm{d}y \overset{\text{Proposition 8}}{\lesssim} Mf(0). \tag{6.4.13}$$

Hence

$$|T_\varepsilon f(0)| \le |Tf_2(z)| + CMf(0) \le |Tf(z)| + |Tf_1(z)| + CMf(0). \tag{6.4.14}$$

WLOG, $|T_\varepsilon f(0)| > 0$. Fix arbitrary $\lambda > 0$ and $0 < \lambda < |T_\varepsilon f(0)|$, and we let

$$Q_1 := \{z \in Q : |Tf(z)| > \frac{\lambda}{3}\}, \quad Q_2 := \{z \in Q : |Tf_1(z)| > \frac{\lambda}{3}\}. \tag{6.4.15}$$

Therefore, we have either $CMf(0) \ge \frac{\lambda}{3}$ or $Q = Q_1 \cup Q_2$. But

$$|Q_1| \le \frac{3}{\lambda}\int_Q |Tf| \le \frac{3}{\lambda}|Q|M|Tf|(0), \quad |Q_2| \overset{\text{weak-(1,1)}}{\le} \frac{1}{\lambda}\|f\|_{L^1(Q)} \le \frac{1}{\lambda}|Q|Mf(0). \tag{6.4.16}$$

Hence, we either have $\lambda \le 3CMf(0)$ or have

$$|Q| \le |Q_1| + |Q_2| \lesssim \frac{|Q|}{\lambda}\left[M|Tf|(0) + Mf(0)\right] \tag{6.4.17}$$

i.e.

$$\lambda \lesssim M|Tf|(0) + Mf(0). \tag{6.4.18}$$

Therefore, by taking e.g. $\lambda = \frac{1}{2}|T_\varepsilon f(0)|$, we get

$$|T_\varepsilon f(0)| \lesssim M|Tf|(0) + Mf(0). \tag{6.4.19}$$

This is for $\nu = 1$. For general $\nu \in (0,1)$ we can repeat the above argument by raising everything in eq. (6.4.14) to the power ν, and then averaging over Q to conclude

$$|T_\varepsilon f(0)| \lesssim \left[Mf(0) + M(|Tf|^\nu)(0)^{\frac{1}{\nu}} + \left(\frac{1}{|Q|}\int_Q |Tf_1|^\nu\right)^{\frac{1}{\nu}}\right]. \tag{6.4.20}$$

□

Proof of the theorem. By the Cotlar-type inequality, with $\nu = 1$, we have strong-(p,p) boundedness since both M and T are. We can argue weak-$(1,1)$ by the same way as in the Hilbert transform case. Alternatively we can also use the Cotlar type inequality with $\nu < 1$, by noting that the superlevel set of T^*f can be controlled by the superlevel sets of $M|Tf|^\nu$ and Mf. For the first term, we first pass to the dyadic maximal function, then use the following property of dyadic maximal function:

$$|\{x : M_d g(x) > \lambda\}| \le \frac{1}{\lambda}\|g\|_{L^1}. \tag{6.4.21}$$

Then use Kolmogorov to control it (We will get something like $|E| \lesssim \frac{1}{\lambda^\nu}|E|^{1-\nu}\|f\|_{L^1}^\nu$). For the second term we control it directly by the weak-$(1,1)$ boundedness of M. Then the boundedness for $1 < p < 2$ follows by Marchinkiewicz interpolation, and for $p > 2$ follows by duality. □

6.5 Vector value generalizations

Let B be a seperable Banach space over $\mathbb{C}$. We say $F : \mathbb{R}^n \to B$ is strongly measurable, iff for every $b' \in B^*$, $x \mapsto \langle F(x), b' \rangle$ is measurable.

Example 21. *If F is measurable, then $x \mapsto \|F(x)\|_B$ is measurable, since it is the supremum of countably many measurable functions (by separability of B).*

Definition 6.5.1. *$L^p(B)$ is the set of measurable functions $F : \mathbb{R}^n \to B$ such that*

$$\|F\|_{L^p(B)} := \left(\int_{\mathbb{R}^n} \|F(x)\|_B^p \mathrm{d}x \right)^{\frac{1}{p}} < \infty. \tag{6.5.1}$$

In particular,

$$L^\infty(B) = \{F : \mathbb{R}^n \to B : \operatorname{esssup}_{x \in \mathbb{R}^n} \|F(x)\|_B < \infty\}. \tag{6.5.2}$$

Proposition 18 (Finite rank vector-valued functions are dense in $L^p(B)$)**.** *We next want to deal with the theory of duality, we first start with finite rank functions. For $\mathbb{C}$-valued $f \in L^p(\mathbb{R}^n)$, we define*

$$(f \cdot b)(x) := f(x)b, \quad b \in B. \tag{6.5.3}$$

Then $(f \cdot b)$ is called a rank-1 *function $\mathbb{R}^n \to B$. Moreover,*

$$\|f \cdot b\|_{L^p(B)} = \left(\int_{\mathbb{R}^n} \|f(x)b\|_B^p \mathrm{d}x \right)^{\frac{1}{p}} = \left(\int_{\mathbb{R}^n} |f|^p \right)^{\frac{1}{p}} \|b\|_B = \|f\|_{L^p(\mathbb{R}^n)} \|b\|_B. \tag{6.5.4}$$

A finite linear combination of rank-1 functions is called a finite rank *function. The set of finite rank functions is denoted by*

$$L^p \otimes B := \left\{ \sum_{i=1}^N f_i \cdot b_i : N \in \mathbb{N}, f_i \in L^p(\mathbb{R}^n), b_i \in B \right\}. \tag{6.5.5}$$

We have: $L^p \otimes B$ is dense in $L^p(B)$ for all $1 \le p < \infty$.

Next we will discuss the duality of $L^p \otimes (B)$. This will be very useful for the study of Littlewood-Paley theory. For example, if we have (r, r)-boundedness, then by duality we have (r', r')-boundedness for the dual operator, then we can do interpolation to get all the (p, p)-boundedness.

Proposition 19. *$F \in L^p(B)$, $G \in L^{p'}(B^*)$, then $\langle F, G \rangle$ is integrable, and*

$$\|G\|_{L^{p'}(B^*)} = \sup_{\|F\|_{L^p(B)} \le 1} \left| \int_{\mathbb{R}^n} \langle \underbrace{F(x)}_{\in B}, \underbrace{G(x)}_{\in B^*} \rangle \mathrm{d}x \right|. \tag{6.5.6}$$

$L^{p'}(B^*) \subset (L^p(B))^*$, and not necessarily equal. The equality holds if B is reflexive and $1 < p < \infty$. (This will be the case in the application in Littlewood-Paley theory.)

Now we can discuss the generalized Calderón-Zygmund theory in the vector-valued setting.

We denote by K an $\mathcal{L}(A, B)$-valued function (A, B are Banach spaces) on $\mathbb{R}^n \times \mathbb{R}^n \setminus \triangle$ and T is an operator associated with K. For $f \in L^\infty(A)$ with compact support, we define

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y) \mathrm{d}y, \quad x \notin \operatorname{supp} f. \tag{6.5.7}$$

So we consider the boundedness of T e.g. from $L^p(A)$ to $L^p(B)$.

Theorem 6.5.2. *$T : L^r(A) \to L^r(B)$ bounded fo some $1 < r < \infty$, satisfying the generalized Hörmander condition*

$$\begin{cases} \int_{|x-y|>2|y-z|} \|K(x,y) - K(x,z)\|_{\mathcal{L}(A,B)} \mathrm{d}x \leq C, \\ \int_{|x-y|>2|x-w|} \|K(x,y) - K(w,y)\|_{\mathcal{L}(A,B)} \mathrm{d}y \leq C, \end{cases} \tag{6.5.8}$$

then T is (p,p)-bounded for any $1 < p < \infty$ and weak-$(1,1)$.

Proof. (sketch) Mostly same as before (for the weak $(1,1)$ and interpolation)

□

Chapter 7

H^1 and BMO

7.1 Atomic H^1 space

Definition 7.1.1. *We say that an L^∞ function a is an atom, if supp $a \subset Q$ for some cube Q, $\|a\|_{L^\infty} \le |Q|^{-1}$, and $\int a(x)\mathrm{d}x = 0$ (zero mean).*

The atom satisfies a key property: the L^1 norm of Ta will be uniformly bounded for any Calderón-Zygmund singular integral T.

Proposition 20 (L^1-uniform boundedness of CZ singular integrals on atoms)**.** *Let T be an integral satisfying the condition of the generalized Calderón-Zygmund theorem Theorem 6.3.1, then there exists $C > 0$ such that for any atom a, we have*

$$\|Ta\|_{L^1} \le C. \tag{7.1.1}$$

Proof. We choose a cube Q^* with the same center c_Q as Q but with side length $2\sqrt{n}$ times of Q. Firstly,

$$\int_{Q^*} |Ta(x)|\mathrm{d}x \overset{\text{Cauchy-Schwarz}}{\le} |Q^*|^{\frac{1}{2}}\|Ta\|_{L^2} \overset{L^2 \text{ boundedness of } T}{\le} C|Q|^{\frac{1}{2}}\|a\|_{L^2} \le C|Q|^{\frac{1}{2}}|Q|^{-\frac{1}{2}} = C. \tag{7.1.2}$$

Next, for $x \notin Q^*$, by the zero mean property of a, we have

$$\begin{aligned}
\int_{\mathbb{R}^n\setminus Q} |Ta(x)|\mathrm{d}x &= \int_{\mathbb{R}^n\setminus Q^*} \left|\int_Q (K(x,y) - K(x,c_Q))a(y)\mathrm{d}y\right|\mathrm{d}x \\
&\overset{\text{Fubini}}{\le} \int_Q \left(\int_{\mathbb{R}^n\setminus Q^*} \left|K(x,y) - K(x,c_Q)\right|\mathrm{d}x\right)|a(y)|\mathrm{d}y \\
&\overset{\text{generalized Hörmander cond., } |x-y| > 2\left|y-c_Q\right| \text{ for any } x\notin Q^*}{\le} C\|a\|_{L^\infty}|Q| \le C.
\end{aligned} \tag{7.1.3}$$

∎

We denote the atomic Hardy space $H^1_{\mathrm{at}}(\mathbb{R}^n)$ to be

$$H^1_{\mathrm{at}}(\mathbb{R}^n) := \left\{ f \in L^1(\mathbb{R}^n) : f = \sum_{j=1}^{\infty} \lambda_j a_j,\ a_j \text{ are atoms}, \sum_{j=1}^{\infty} \left|\lambda_j\right| < \infty \right\}, \tag{7.1.4}$$

and clearly, $H^1_{\rm at}(\mathbb{R}^n)$ is contained in $L^1(\mathbb{R}^n)$. We define the norm

$$\|f\|_{H^1_{\rm at}} := \inf\left\{\left\|(\lambda_j)\right\|_{\ell^1} := \sum_{j=1}^{\infty}\left|\lambda_j\right| : f = \sum_{j=1}^{\infty}\lambda_j a_j,\ a_j \text{ are atoms}\right\}. \tag{7.1.5}$$

With this definition, $H^1_{\rm at}(\mathbb{R}^n)$ is a Banach space. With this, if we view T as an operator from $H^1_{\rm at}(\mathbb{R}^n)$ to $L^1(\mathbb{R}^n)$, then it follows immediately from Proposition 20 that T is bounded from $H^1_{\rm at}(\mathbb{R}^n)$ to $L^1(\mathbb{R}^n)$.

Corollary 18. *With the same assumption as in Proposition 20, T is a bounded operator from $H^1_{\rm at}(\mathbb{R}^n)$ to $L^1(\mathbb{R}^n)$.*

The atomic Hardy space is the "largest" space such that all Calderón-Zygmund singular integrals are bounded from it to $L^1(\mathbb{R}^n)$. More precisely, we have the following theorem:

Theorem 7.1.2. $H^1(\mathbb{R}^n) = H^1_{\rm at}(\mathbb{R}^n)$ *with equivalent norms.*
Here,

$$H^1(\mathbb{R}^n) := \{f \in L^1(\mathbb{R}^n) : R_j f \in L^1(\mathbb{R}^n)\}, \tag{7.1.6}$$

where R_j ($j = 1, \dots, n$) are the Riesz transforms (Hilbert transform for $n = 1$). The norm is given by

$$\|f\|_{H^1} := \|f\|_{L^1} + \sum_{j=1}^{n}\left\|R_j f\right\|_{L^1}. \tag{7.1.7}$$

7.2 BMO space

The sharp maximal function $M^\sharp f$ is defined as the supremum of the *mean oscillation* of f over all cubes containing x.

$$M^\sharp f(x) := \sup_{Q\ni x}\frac{1}{|Q|}\int_Q\left|f(y) - f_Q\right|\mathrm{d}y, \quad f_Q := \frac{1}{|Q|}\int_Q f(z)\mathrm{d}z. \tag{7.2.1}$$

The BMO space is defined as the set of functions with bounded mean oscillation.

Definition 7.2.1.

$$BMO = \{f \in L^1_{\rm loc}(\mathbb{R}^n) : \|f\|_{BMO} := \left\|M^\sharp f\right\|_{L^\infty} < \infty\}, \tag{7.2.2}$$

We define $\|f\|_* = \left\|M^\sharp f\right\|_{L^\infty}$, which is not properly a norm since **a.e. constants** have zero mean oscillation. In fact, $BMO/\sim$ (where $f \sim g$ iff $f - g$ is a.e. constant) is a Banach space with the norm $\|\cdot\|_*$. We customarily think of BMO as the quotient space $BMO/\sim$.

Some basic properties of BMO:

Proposition 21.

- *$M^\sharp f \lesssim_n Mf$ pointwise for all locally integrable f.*
- $$\frac{1}{2}\|f\|_* \le \sup_Q \inf_{a\in\mathbb{C}}\frac{1}{|Q|}\int_Q|f(x) - a|\mathrm{d}x \le \|f\|_*, \tag{7.2.3}$$
- $$M^\sharp|f|(x) \le 2M^\sharp f(x). \tag{7.2.4}$$

- By definition, $M^\sharp f$ is dominated by the Hardy-Littlewood maximal function Mf. (Note that if we consider the cube maximal function $M''f$, then $M^\sharp f \le 2M''f$ and the constant does not depend on dimension.)

- The second inequality is immediate by taking $a = f_Q$ where f_Q is the average of f over Q. For the first inequality, we use triangle inequality to get

$$\int_Q \left|f(x) - f_Q\right| \mathrm{d}x \le \int_Q |f(x) - a| \mathrm{d}x + \int_Q \left|a - f_Q\right| \mathrm{d}x \le 2 \int_Q |f(x) - a| \mathrm{d}x. \tag{7.2.5}$$

  Now, devide by $|Q|$ and take supremum over all cubes Q containing x, then take infimum over all $a \in \mathbb{C}$ to get the desired result.

- By taking $a = \left|f_Q\right|$, and note that $\left||f(x)| - \left|f_Q\right|\right| \le \left|f(x) - f_Q\right|$.

Remark 50. *The second inequality gives us an equivalent norm for BMO, without using the average f_Q. Specifically, $f \in BMO$ if and only if there exists $C > 0$ such that for any cube Q, there exists $a_Q \in \mathbb{C}$ such that*

$$\frac{1}{|Q|} \int_Q \left|f(x) - a_Q\right| \mathrm{d}x \le C. \tag{7.2.6}$$

The third property shows that if $f \in BMO$, then $|f| \in BMO$ as well. But the converse is not true in general. This also suggests that being in BMO is not just a property of the size of the function, but also of its oscillation. Clearly $L^\infty \subset BMO$ (e.g. by taking $a \in \mathbb{R}$, $a < 0$ in the equivalent norm), but there are also unbounded functions in BMO.

Example 22.

$$f(x) = \begin{cases} \log \frac{1}{|x|}, & |x| < 1, \\ 0, & |x| \ge 1. \end{cases} \tag{7.2.7}$$

is in $BMO(\mathbb{R})$. But $sgn(x) \cdot f(x)$ is not in $BMO(\mathbb{R})$ (even though f is its absolute value).

Chapter 8

Littlewood-Paley theory

General idea: if we only keep *some* Fourier frequencies of a function, will the generalized function change a lot in terms of L^p?

The answer is, briefly, smooth dyadic modification of a function on the Fourier side is *mild*.

Theorem 8.0.1 (informal)**.** *Let $\widetilde{S}_j$ be smooth dyadic cutoffs on the Fourier side, e.g.*

$$\widehat{\widetilde{S}_j f}(\xi) = \psi(2^{-j}\xi)\widehat{f}(\xi), \quad \psi \in C_c^\infty(\mathbb{R}^n),\ \psi \equiv 1 \text{ on } \{\frac{1}{2} \le |\xi| \le 2\},\ \operatorname{supp}\psi \subset \{\frac{1}{4} \le |\xi| \le 4\}. \tag{8.0.1}$$

8.1 Vector-valued inequalities and Littlewood-Paley theory in 1D

We recall from Theorem 6.5.2 that if T is a vector-valued operator satisfying the generalized Hörmander condition, then T is (p,p)-bounded for all $1 < p < \infty$ and weak-$(1,1)$. We now apply this to obtain: (note for this chapter, we will again go back to the convolution kernel setting, instead of going to the general Calderón-Zygmund kernel setting)

Theorem 8.1.1. *If T is a convolution operator bounded on $L^2(\mathbb{R}^n)$, w. the associated kernel K satisfying the Hörmander condition, then we have*

$$\left\|\left(\sum_j |Tf_j|^r\right)\right\|_p \lesssim_{p,r} \left\|\left(\sum_j |f_j|^r\right)\right\|_p, \quad 1 < p < \infty,\ 1 < r < \infty. \tag{8.1.1}$$

Moreover, $(f_j) \in L^1(\ell^r)$ implies that $(Tf_j) \in L^{1,\infty}(\ell^r)$ with weak-$(1,1)$ bound.

Remark 51. *How to understand this vector-valued inequality? From the first part of the condition, we know that T is strong-(r,r) and weak-$(1,1)$ as operators between scalar functions (by the Calderón-Zygmund theorem for scalar cases). Therefore, for $(f_j) \in L^p(\ell^r)$, we know entry-wise Tf_j is well-defined and is in L^p. A natural attempt is to first look at the $p = r$ case, then we have*

$$\left\|\left(\sum_j |Tf_j|^r\right)^{\frac{1}{r}}\right\|_{L^r}^r = \int_{\mathbb{R}^n} \underbrace{\sum_j |Tf_j(x)|^r}_{\ell^r\text{-norm of }(Tf_j(x))_j} \mathrm{d}x \overset{\text{Fubini}}{=} \sum_j \|Tf_j\|_{L^r}^r \overset{T \text{ is strong-}(r,r)}{\lesssim_r} \sum_j \|f_j\|_{L^r}^r = \left\|\left(\sum_j |f_j|^r\right)^{\frac{1}{r}}\right\|_{L^r}^r. \tag{8.1.2}$$

Now we just consider the operator in the vector-valued setting:

$$\text{“}T^{\otimes\mathbb{N}}\text{”} : (f_j)_{j\in\mathbb{N}} \mapsto (Tf_j)_{j\in\mathbb{N}} \tag{8.1.3}$$

whose kernel is given by $K(x)\cdot I$ where I is identity. Obviously since K satisfies Hörmander, $K(x)\cdot I$ also satisfies the generalized Hörmander condition. Therefore by Theorem 6.5.2, we have the desired boundedness for all $1 < p < \infty$ and weak-$(1,1)$.

Caution: not easy to analyze $(f_j) \mapsto (T_j f_j)$ in general if T_j's are unrelated. In this case uniform boundedness of T_j is not enough, because what is missing here is something like a “uniform” version of Hörmander condition:

$$\int_{|x|>2|y|} \sup_j \left|K_j(x-y) - K_j(x)\right| \mathrm{d}x \le C \tag{8.1.4}$$

Interestingly, for “sharp cutoffs” on the Fourier side in 1D, we can have the $(f_j) \mapsto (T_j f_j)$ boundedness. This is because in 1D, the sharp cutoff operators are all related to the Hilbert transform.

Proposition 22. *Let S_j be defined by*

$$\widehat{S_j f}(\xi) = \chi_{I_j}(\xi)\widehat{f}(\xi), \tag{8.1.5}$$

where I_j are intervals in $\mathbb{R}$. Then we have for any $1 < p < \infty$, $1 < r < \infty$,

$$\left\|\left(\sum_j \left|S_j f_j\right|^r\right)^{\frac{1}{r}}\right\|_{L^p} \lesssim_{p,r} \left\|\left(\sum_j \left|f_j\right|^r\right)^{\frac{1}{r}}\right\|_{L^p}. \tag{8.1.6}$$

Proof. Let $I_j = [a_j, b_j]$. We can write

$$S_j f = \frac{\mathrm{i}}{2}\left[e^{2\pi \mathrm{i} b_j x} H(e^{-2\pi \mathrm{i} b_j x} f) - e^{2\pi \mathrm{i} a_j x} H(e^{-2\pi \mathrm{i} a_j x} f)\right]. \tag{8.1.7}$$

(Similar expression for half-infinite intervals, only one term.) Then use Theorem 8.1.1 for the Hilbert transform H to conclude. □

With this we can already show the Littlewood-Paley theory in 1D with dyadic sharp cutoffs.

$$\Delta_j := (-2^{j+1}, -2^j] \cup [2^j, 2^{j+1}). \tag{8.1.8}$$

Let

$$\widehat{S_j f}(\xi) = \chi_{\Delta_j}(\xi)\widehat{f}(\xi). \tag{8.1.9}$$

We will prove $\left\|(\sum_j \left|S_j f\right|^2)^{\frac{1}{2}}\right\|_{L^p} \sim_p \|f\|_{L^p}$ for all $1 < p < \infty$. The $p = 2$ case is already known by Plancherel.

Theorem 8.1.2 (Smooth dyadic cutoffs in 1D)**.** *Let $f \in L^p(\mathbb{R})$ for some $1 < p < \infty$. Let the smooth dyadic cutoff operators $\widetilde{S}_j$ be defined by*

$$\widehat{\widetilde{S}_j f}(\xi) = \psi(2^{-j}\xi)\widehat{f}(\xi), \quad \psi \in C_c^\infty(\mathbb{R}),\ \psi \equiv 1 \text{ on } \{\frac{1}{2} \le |\xi| \le 2\},\ \text{supp}\,\psi \subset \{\frac{1}{4} \le |\xi| \le 4\}. \tag{8.1.10}$$

Then we have

$$\left\|\left(\sum_j \left|\widetilde{S}_j f\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} \sim_p \|f\|_{L^p}. \tag{8.1.11}$$

Proof. (sketch) $p = 2$ true directly by Plancherel. We can check that the kernel of $\widetilde{S}_j$ satisfies Hörmander condition uniformly in j (By noting that the kernel $\Psi_j(x) = 2^j\check{\psi}(2^jx)$ satisfies $\left\|\Psi_j'\right\|_{\ell^2} \leq C|x|^{-2}$. This is because$|\psi'(x)| \leq C\min(1, |x|^{-3})$ with C independent of j since $\Psi \in \mathscr{S}$.), and we have the same proposition Proposition 22 for $\widetilde{S}_j$'s as the vector-valued inequality in sharp cutoff case. Therefore, by the same argument as before, we have the desired boundedness for all $1 < p < \infty$. The reverse inequality follows by duality. □

Theorem 8.1.3 (Dyadic sharp cutoffs in 1D)**.** *Let $f \in L^p$ $(1 < p < \infty)$, and let S_j be defined as the dyadic sharp cutoff operators above. Then we have*

$$\left\|\left(\sum_j \left|S_jf\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} \sim_p \|f\|_{L^p}. \tag{8.1.12}$$

Proof. We note a simple identity: $S_j\widetilde{S}_j = S_j$, which means that

$$\left\|\left(\sum_j \left|S_jf\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} = \left\|\left(\sum_j \left|S_j\widetilde{S}_jf\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} \lesssim_p \left\|\left(\sum_j \left|\widetilde{S}_jf\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} \sim_p \|f\|_{L^p}. \tag{8.1.13}$$

The reverse inequality follows by duality. □

There are two natural ways to generalize the above Littlewood-Paley theory to higher dimensions:

- Consider the smooth dyadic cutoffs on the annuli $2^j \leq |\xi| \leq 2^{j+1}$ in $\mathbb{R}^n$.
- Consider the product of sharp diadic intervals in each coordinate, i.e. consider the rectangles of the form
$$R_{\mathbf{j}} = \prod_{k=1}^n \left([-2^{j_k+1}, -2^{j_k}] \cup [2^{j_k}, 2^{j_k+1})\right), \quad \mathbf{j} = (j_1, \ldots, j_n) \in \mathbb{Z}^n. \tag{8.1.14}$$
- We **cannot** consider sharp cutoffs on the annuli in higher dimensions!

8.2 Littlewood-Paley theory in higher dimensions

Theorem 8.2.1 (Smooth dyadic cutoffs in higher dimensions)**.** *$\psi \in \mathscr{S}(\mathbb{R}^n)$ with $\psi(0) = 0$ and $S_j(j \in \mathbb{Z})$ defined by*

$$\widehat{S_jf}(\xi) = \psi(2^{-j}\xi)\widehat{f}(\xi). \tag{8.2.1}$$

Then for any $1 < p < \infty$, we have

$$\left\|\left(\sum_j \left|S_jf\right|^2\right)^{\frac{1}{2}}\right\|_{L^p(\mathbb{R}^n)} \lesssim_p \|f\|_{L^p(\mathbb{R}^n)}. \tag{8.2.2}$$

Furtherfore, if for all $\xi \neq 0$ we have

$$\sum_{j\in\mathbb{Z}} \left|\psi(2^{-j}\xi)\right|^2 = C < \infty, \tag{8.2.3}$$

then we also have the reverse inequality

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim_p \left\| \left(\sum_j \left| S_j f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)}. \tag{8.2.4}$$

Proof. (sketch) Since $\psi \in \mathscr{S}$ and $\psi(0) = 0$, it satisfies

$$\sum_j \left| \psi(2^{-j}\xi) \right|^2 \leq C. \tag{8.2.5}$$

Therefore the proof of the first inquality is exactly the proof of Theorem 8.1.2. For the second part, it follows again from duality since by the equality we have

$$\int_{\mathbb{R}^n} \sum_j S_j f(x) \overline{S_j g(x)} \mathrm{d}x = C \int_{\mathbb{R}^n} f(x) \overline{g(x)} \mathrm{d}x. \tag{8.2.6}$$

Then duality argument holds as follows:

$$\begin{aligned} \|f\|_p &= \sup_{\|g\|_{p'} \leq 1} \left| \int_{\mathbb{R}^n} f\overline{g} \right| = \frac{1}{C} \sup_{\|g\|_{p'} \leq 1} \left| \int_{\mathbb{R}^n} \sum_j S_j f \overline{S_j g} \right| \\ &\leq \frac{1}{C} \left\| \left(\sum_j \left| S_j f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \sup_{\|g\|_{p'} \leq 1} \left\| \left(\sum_j \left| S_j g \right|^2 \right)^{\frac{1}{2}} \right\|_{L^{p'}} \lesssim_p \left\| \left(\sum_j \left| S_j f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p}. \end{aligned} \tag{8.2.7}$$

□

Remark 52. *It is not hard to construct a smooth cutoff function ψ satisfying the above conditions (see e.g. J. Duoandikoetxea's book, section 8.3)*

Theorem 8.2.2 (Rectangle dyadic sharp cutoffs in higher dimensions)**.** *Let $R_{\mathbf{j}}$ be defined as above for $\mathbf{j} \in \mathbb{Z}^n$, and define*

$$\widehat{S_{\mathbf{j}} f}(\xi) = \chi_{R_{\mathbf{j}}}(\xi) \widehat{f}(\xi). \tag{8.2.8}$$

Then for any $1 < p < \infty$, we have

$$\left\| \left(\sum_{\mathbf{j} \in \mathbb{Z}^n} \left| S_{\mathbf{j}} f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)} \sim_p \|f\|_{L^p(\mathbb{R}^n)}. \tag{8.2.9}$$

Proof. We only prove for $\mathbb{R}^2$ for simplicity. The general case is similar. Note that in this case $S_{\mathbf{j}}$ is a product of 1D dyadic sharp cutoff operators:

$$S_{\mathbf{j}} = S^{(1)}_{j_1} \circ S^{(2)}_{j_2}, \tag{8.2.10}$$

thus the vector-valued inequality Proposition 22 implies:

$$\left\| \left(\sum_{j,k} \left| \widetilde{S}^{(2)}_j f_k \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^2)} \lesssim_p \left\| \left(\sum_k |f_k|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^2)}. \tag{8.2.11}$$

This holds by Fubini for f_k being a function of two variables. Furthermore we also have

$$S_j^{(\alpha)}\widetilde{S}_j^{(\alpha)} = S_j^{(\alpha)}, \quad \alpha = 1, 2, \tag{8.2.12}$$

then we can apply the vector-valued inequality twice again to get

$$\left\| \left(\sum_{j,k} \left| S_j^{(1)} S_k^{(2)} f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^2)} \lesssim_p \left\| \left(\sum_{k} \left| \widetilde{S}_k^{(2)} f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^2)} \lesssim_p \|f\|_{L^p(\mathbb{R}^2)}. \tag{8.2.13}$$

□

8.3 The Hörmander multiplier theorem

Perhaps one of the most remarkable applications of Littlewood-Payley theory is the Hörmander multiplier theorem, which gives a sufficient condition for a Fourier multiplier to be bounded on L^p for all $1 < p < \infty$.

$$L_a^2(\mathbb{R}^n) = \left\{ g : \langle\xi\rangle^a \widehat{g}(\xi) \in L^2(\mathbb{R}^n) \right\}, \tag{8.3.1}$$

where $\langle\xi\rangle = (1 + |\xi|^2)^{\frac{1}{2}}$ is the Japanese bracket. The space $L_a^2(\mathbb{R}^n)$ is a Sobolev space with smoothness index a measured in L^2 sense.

Proposition 23. *If $a > \frac{n}{2}$ and $g \in L_a^2(\mathbb{R}^n)$, then $\widehat{g} \in L^1$ and inparticular, g is continuous and bounded.*

Proof.

$$\int |\widehat{g}| \leq \left(\int |h|^2 \right)^{\frac{1}{2}} \left(\int \langle\xi\rangle^{-a} \right)^{\frac{1}{2}} \leq C_a \|g\|_{L_a^2}, \tag{8.3.2}$$

where $h := \langle\xi\rangle^a \widehat{g}(\xi) \in L^2$. Note that since we have $a > \frac{n}{2}$, C_a would be a finite constant. □

Therefore, if m is in L_a^2 with $a > \frac{n}{2}$, then m is bounded and thus in L^∞. In fact, if $\widehat{Tf} = m\widehat{f}$, then $Tf = K * f$ with $K \in L^1$. Hörmander theorem tells us thta m is a multiplier on L^p, under a much weaker condition.

Theorem 8.3.1 (Hörmander multiplier theorem)**.** *Let ψ be the radial function supported on $\frac{1}{2} \leq |\xi| \leq 2$ such that $\sum_{j\in\mathbb{Z}} \psi(2^{-j}\xi) = 1$ for all $\xi \neq 0$ (cf. the end of the last section).*

Let $m \in L^\infty(\mathbb{R}^n)$ such that

$$\sup_{j\in\mathbb{Z}} \left\| m(2^j \cdot)\psi \right\|_{L_a^2(\mathbb{R}^n)} < \infty, \tag{8.3.3}$$

for some $a > \frac{n}{2}$. Then the multiplier operator T_m defined by $\widehat{T_m f}(\xi) = m(\xi)\widehat{f}(\xi)$ is bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$.

The proof is based on the Littlewood-Paley theory with smooth dyadic cutoffs. We first give a technical lemma which is a weighted inequality

Lemma 14. *Let $m \in L_a^2$ for some $a > \frac{n}{2}$, $\lambda > 0$, we define the operator T_λ by*

$$\widehat{T_\lambda f}(\xi) = m(\lambda\xi)\widehat{f}(\xi). \tag{8.3.4}$$

Then

$$\int_{\mathbb{R}^n} |T_\lambda f(x)|^2 u(x) \mathrm{d}x \lesssim_{m,n,a} \int_{\mathbb{R}^n} |f(x)|^2 Mu(x) \mathrm{d}x, \tag{8.3.5}$$

where the constant is independent of λ and u, and M is the Hardy-Littlewood maximal function.

Proof. Let $\widehat{K} = m$, $R := \langle\xi\rangle^a \widehat{K}(\xi) \in L^2$ with norm $\|R\|_{L^2} = \|m\|_{L^2_a}$. Then the kernel of T_λ is given by $\lambda^{-n}K(\lambda^{-1}x)$, thus

$$\begin{aligned}\int_{\mathbb{R}^n} |T_\lambda f|^2 u &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \frac{\lambda^{-n}R(\lambda^{-1}(x-y))}{\langle\left|\lambda^{-1}(x-y)\right|\rangle^a} \mathrm{d}y \right|^2 u(x)\mathrm{d}x \\ &\overset{\text{Cauchy-Schwarz}}{\leq} \lambda^{-n}\|R\|_{L^2}^2 \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|f(y)|^2}{\langle\left|\lambda^{-1}(x-y)\right|\rangle^{2a}} u(x)\mathrm{d}y\mathrm{d}x \\ &\overset{\langle\rangle^{-a} \text{ is radial and decreasing, Proposition 8}}{\lesssim}{}_{n,a}\ \lambda^{-n}\|m\|_{L^2_a}^2 \int_{\mathbb{R}^n} |f(y)|^2 Mu(y)\mathrm{d}y.\end{aligned} \tag{8.3.6}$$

□

Proof of Hörmander multiplier theorem. Let $\widetilde{\psi}$ be the smooth cutoff function and $\widetilde{S}_j$ be the associated smooth dyadic cutoff operators. Define S_j associated to ψ as before. Note that we have the similar identity $S_j\widetilde{S}_j = S_j$ for all j. Now both S_j and $\widetilde{S}_j$ satisfies the inequality in Theorem 8.2.1. Thus

$$\|Tf\|_p \leq C\left\|\left(\sum_j \left|S_jTf\right|^2\right)^{\frac{1}{2}}\right\|_p = C\left\|\left(\sum_j \left|S_jT\widetilde{S}_jf\right|^2\right)^{\frac{1}{2}}\right\|_p. \tag{8.3.7}$$

The multiplier of S_jT is given by $\psi(2^{-j}\xi)m(\xi)$, thus by $\sup_j \left\|m(2^j\cdot)\psi\right\|_{L^2_a} < \infty$ and Lemma 14, we have

$$\int_{\mathbb{R}^n} \left|S_jTf\right|^2 u \leq C\int_{\mathbb{R}^n} |f|^2 Mu. \tag{8.3.8}$$

Here, C is independent of j. Using this inequality and the vector-valued inequality argument, we can prove that for $p \geq 2$,

$$\left\|\left(\sum_j \left|S_jTg_j\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} \leq C_p\left\|\left(\sum_j \left|g_j\right|^2\right)^{\frac{1}{2}}\right\|_{L^p}. \tag{8.3.9}$$

Taking $g_j = \widetilde{S}_jf$, we have

$$\|Tf\|_p \leq C\left\|\left(\sum_j \left|\widetilde{S}_jf\right|^2\right)^{\frac{1}{2}}\right\|_p \lesssim C\|f\|_p, \tag{8.3.10}$$

where the last inequality follows from Theorem 8.2.1 for $\widetilde{S}_j$. The case $1 < p < 2$ follows by duality. □

The Hörmander multiplier theorem is usually stated in the following way:

Corollary 19. $k = \lfloor n/2 \rfloor + 1$. *If* $m \in C^k(\mathbb{R}^n \setminus \{0\})$ *satisfies*

$$\sup_R R^{|\beta|}\left(\frac{1}{R^n}\int_{R<|\xi|<2R} \left|D^\beta m(\xi)\right|^2 \mathrm{d}\xi\right)^{\frac{1}{2}} < \infty. \tag{8.3.11}$$

Then, m is a multiplier on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$. In particular, if m satisfies the Mikhlin condition

$$\left|D^\beta m(\xi)\right| \leq C_\beta|\xi|^{-|\beta|}, \quad \forall|\beta| \leq k, \tag{8.3.12}$$

then m is a multiplier on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$.

Example 23.

- *$m(\xi) = |\xi|^{\mathrm{i}t}$ satisfies Mikhlin condition with $C_\beta \lesssim |t|^{|\beta|}$.*
- *If m is homogeneous of degree 0 and C^k on the unit sphere for some $k > \frac{n}{2}$, then m satisfies the Mikhlin condition.*