Harmonic Analysis

Author: Hui Sun

Date: October 10, 2023

Contents

Logistics

0.1 Lecture 1

Here we go.

0.1.1 Logistics

OH: Wednesday 1-2pm virtual, 2-4pm Evans 813.

Textbook: Fourier Analysis by J. Duoandikoetxea (plan to cover chapter 1-6, and sections 1-4 of chapter 8); other texts: *Introduction to Fourier Analysis on Euclidean Spaces* by Stein and Weiss, *Singular Integrals and Differential Properties of Functions* by Stein, *Harmonic Analysis* by Stein.

Topics: Fourier series, Fourier transform, maximal functions, Hilbert transform, singular integrals, Littlewood-Paley theorem, multipliers, oscillatory integrals

Grading: The grading will be entirely dependent on 3 problem sets given throughout the semester (with an ample amount of optional problems).

0.1.2 Course Overview

We will begin by defining what "Harmonic" means in the context of Math 258: to us, this word harmonic refers to "Euclidean Fourier analysis." And more specifically, we will study Fourier analysis on the n-dimensional torus, in $\mathbb{R}^n$. One justification for studying on/in these spaces is that many are equipped with translation invariance, which among other things, gives us nice behaving eigenfunctions.

Consider the function $e(x) := e^{2\pi i n x}$, and consider the translation operator $f_t(x) = f(x + t)$, we have

$$f_t(e(x)) = e^{2\pi i n(x+t)} = e^{2\pi i n t} \cdot e^{2\pi i n x}$$

Here, $e^{2\pi i n x}$ can be seen as an eigenfunction of translations. Another obvious example is differentiation. Consider the differentiation operator on $e(x)$, we have

$$\partial_x(e(x)) = 2\pi i n(e^{2\pi i n x})$$

Again, $e^{2\pi i n x}$ is an eigenfunction. In the forseeable future, we will see $\{e^{2\pi i n x}\}_{n \in \mathbb{N}}$ forms a basis of funcitons on the 1-dim torus, $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ i.e. functions on the circle. Likewise, we have $\{e^{2\pi i \sum n_i x_i}\}_{n_i \in \mathbb{Z}}$ as the basis of functions in the n-dim torus, defined as $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$. They have the nice properties of diagonalizing translation, differentiation operators as they are eigenfunctions. Similarly, we have $\{e^{2\pi i \sum n_i x_i}\}_{n_i \in \mathbb{Z}}$ for $\mathbb{R}^n$, and we say they are "almost in L^2," or L^2-wannabes as they are not far from L^2, but not quite in $L^2(\mathbb{R}^n)$.

Remark This property gives them the importance of in studying differential operators with constant coefficients

We will go through various technical things along the way, one of them being "cancellation." In the most general sense, using triangle inequality for everything is quite of a waste, for example, for highly oscillatory functions. We would like to exploit whenever we can, such as the oscillations of functions, kernels of operators, etc. More importantly, we will use different methods for different parts, to treat different issues. In other words, one should go to the dentist when they broke their ankle.

We will first study the question when do partial sums of a Fourier series (of functions on the cirlce) or Fourier transform of functions in the Euclidean space converge, and converge in what sense. Convergence usually has two "senses:" pointwise convergence and L^p norm convergence. We will study both.

Chapter 1 Fourier Series and Fourier transform

1.1 Lecture 2

More logistics: OH's have been updated as follows: Wednesday 10:30am-11:30am, 2-3pm.

We now begin Chapter 1 of our text.

Recall we define the 1-d torus as $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and the functions on the torus are naturally identified with the functions on the unit interval. Fourier analysis began when Fourier asked the following question: given a function f on the circle, can we find a_k, b_k such that the following is true:

$$f(x) = \sum_{k=0}^{\infty} a_k \cos(2\pi kx) + b_k \sin(2\pi kx)$$

The modern Fourier analysis asks the following, can we find c_k such the following is true:

$$f(x) = \sum_{k=0}^{\infty} c_k e^{2\pi ikx}$$

The above two questions are identical if we take $a_k = c_k, b_k = ic_k$.

One intuitioin for having $2\pi k$ in the $\cos, \sin$ is that we would like to have periodic functions with period 1 to approximate f. Now we introduce the Fourier coefficients, and we first motivate this using trigonometric polynomials of the form

$$f(x) = \sum_{k=0}^{N} c_k e^{2\pi ikx}$$

We only know f is a finite sum of $e^{2\pi ikx}$, yet we would like to know the c_k's. And we do this by exploiting the orthogonality of $\{e^{2\pi ikx}\}$. Notice we have the following:

$$\int_0^1 e^{2\pi ik_1 x}\overline{e^{2\pi ik_2 x}}dx = \begin{cases} 1, k_1 = k_2 \\ 0, k_1 \neq k_2 \end{cases}$$

We therefore have, for any fixed k,

$$\int_0^1 f(x)e^{-2\pi ikx}dx = \int_0^1 \left(\sum_{k=0}^{n} c_k e^{2\pi ikx} \right) e^{-2\pi ikx}dx = c_k$$

Definition 1.1 (Fourier coefficients, Fourier series)

Given $f \in L^1(\mathbb{T})$, or any periodic function on $\mathbb{R}$ with period 1 that is locally integrable, we define its k-th Fourier coefficient as follows:

$$\hat{f}(k) = \int_0^1 f(x)e^{-2\pi ixk}dx$$

Given the Fourier coefficients, we write f's Fourier series as follows:

$$f(x) \sim \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{2\pi ikx}$$

♣

For trigonometric polynomials, as we saw, its Fourier series agrees exactly with itself, and its Fourier series will only have finitely many terms. However, more often than not, arbitrary f's Fourier series will have infinitely many terms, and now we go back to the question: when will f's Fourier series converge and converge in what sense.

We define partial sums of Fourier series as $S_N(f)(x) = \sum_{|k|\leq N} \hat{f}e^{2\pi ikx}$, and pointwise convergence asks the question: for fixed x, when do we have

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

We now introduce two theorems on pointwise convergence (that will be proved in the next lecture).

Theorem 1.1 (Dini's criterion)

Fix $x \in \mathbb{T}$, if we have

$$\int_{|t|<\delta} \left| \frac{f(x+t)f(x)}{t} \right| < \infty$$

Then we have

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

For example, Lipshitz functions would have their Fourier series pointwisely converge. More generally, if a function blows up around a point only slightly, then we would have its Fourier series converge to it at that point.

Theorem 1.2 (Jordan's criterion)

Fix $x \in \mathbb{T}$, and some $\delta > 0$, if f is of bounded variation in $(x - \delta, x + \delta)$, then we have

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

Again, we delay the proof till next time.

To study pointwise convergence, we first note that $S_N(f)$ is a convolution operator.

$$S_N(f)(x) = \sum_{|k|\leq N} \hat{f}(k)e^{2\pi ikx}$$

$$= \sum_{k} \left(\int_0^1 f(t)e^{-2\pi ikt}dt \right) e^{2\pi ikx}$$

$$= \int_0^1 f(t) \sum_k e^{2\pi ik(x-t)}dt$$

$$= f(x) * D_N(x)$$

where $D_N(x) = \sum_{|k|\leq N} e^{2\pi ikx}$, and this kernel is called the Dirichlet kernel.

Remark To study the pointwise convergence of functions on the circle, it suffices to study the point $x = 0$, and by translation invariance, we have the same conclusion hold for $x = x_0$.

The convolution can be thought of as "redistribution of mass," and we show the total mass of the Dirichlet kernel is 1.

$$\int_0^1 D_N(t)dt = \int_0^1 e^{2\pi ikt}dt = 1$$

We now introduce a simple expression for the Dirichlet's kernel.

Proposition 1.1 (Dirichelt kernel)

We have

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

Proof $D_N(t) = \sum_{k=-N}^N e^{2\pi ikt}$ is a geometric series, with the ratio $= e^{-2\pi it}$, hence by the formula of partial sums, we have

$$LHS = \frac{e^{2\pi i2Nt}(1 - e^{-2\pi it(2N+1)})}{1 - e^{-2\pi it}}$$

Now we examine the RHS. $\sin(t) = (e^{it} - e^{-it})/2i$, hence we have

$$RHS = \frac{e^{(2N+1)\pi it} - e^{-(2N+1)\pi it}}{e^{i\pi t} - e^{-i\pi t}}$$

Dividing top and bottom by $e^{i\pi t}$ gives us the LHS.

$\square$

Can we comment on the bound of D_N? If one draws out a picture, then it is clear that D_N have a blow-up at $t = 0$,

and $|D_N(t)| \to 2N + 1$ as $t \to 0$, one could also see this using the expression above. Also we have the following:

$$|D_N(t)| \leq \frac{1}{\sin(\pi t)}, 0 < t \leq \frac{1}{2}$$

To prove the above two theorems, we introduce some tools.

> **Lemma 1.1 (Riemann-Lebesgue Lemma)**
>
> *If $f \in L^1(\mathbb{T})$, then as $|k| \to \infty$, the Fourier coefficent tends to 0, i.e. we have*
>
> $$\lim_{|k| \to \infty} |\hat{f}(k)| = \lim_{|k| \to \infty} \left| \int_0^1 f(x) e^{-2\pi i k x} \right| = 0$$
>
> ♡

Proof For $f \in L^1(\mathbb{T})$, there exists g, h such that $f = g + h$, where g is a simple function of the form $g = \sum_{i=1}^{n} c_i \chi_{E_i}$, and $\|h\|_{L^1} < \epsilon$.

For g simple function, the Fourier coefficent decays with rate $O(1/k)$ by integration by parts.

$$|\hat{g}(k)| = \left| \int_0^1 \sum_i c_i \chi_{E_i} e^{-2\pi i k x} \right| \leq \sum |c_i| \chi_{E_i} \left| \int_0^1 e^{-2\pi i k x} \right| = \sum |c_i| \chi_{E_i} \frac{1}{2\pi i k} \left| \int_0^1 x e^{-2\pi i k x} \right| \lesssim O(1/|k|)$$

For $\hat{h}(k)$, we have,

$$\left| \hat{h}(k) \right| = \left| \int_0^1 h(x) e^{-2\pi i k x} \right| \leq \int_0^1 |h(x)| dx < \epsilon$$

Hence we have $|\hat{f}(k)| \leq O(1/k)$, and as $|k| \to \infty$, we have $\hat{f}(k) \to 0$ if $f \in L^1(\mathbb{T})$.

□

Next we a result that guarantees pointwise convergence.

> **Theorem 1.3 (Riemann Localization principle)**
>
> *If $f = 0$ in a neighborhood of x, say, $(x - \delta, x + \delta)$, then we have*
>
> $$\lim_{N \to \infty} S_N(f)(x) = 0 = f(x)$$
>
> ♡

Proof If we write $S_N(f)$ as $D_N * f$, where $D_N(t) = \frac{\sin((2N+1)\pi t)}{\sin(\pi t)}$, then we see the badness of the denominator is avoided around x by $f = 0$ around x, and if we are far away from x, then nice bound ensues. We now do explicit computation.

$$S_N(f) = \int_0^1 D_N(t) f(x - t) dt$$
$$\int_{\delta \leq t \leq 1} \frac{\sin((2N+1)\pi t)}{\sin(\pi t)} f(x - t) dt$$
$$= \int_{\delta \leq t \leq 1} \frac{f(x-t)}{\sin(\pi t)} \frac{e^{i(2N+1)\pi t} - e^{i(2N+1)\pi t}}{2i} dt$$
$$= \frac{1}{2i} \widehat{\frac{f(x-t)}{\sin(\pi t)} e^{i\pi t}}(-N) + \widehat{\frac{f(x-t)}{\sin(\pi t)} e^{-i\pi t}}(N)$$

And we note that both $\frac{f(x-t)}{\sin(\pi t)} e^{i\pi t}$ and $\frac{f(x-t)}{\sin(\pi t)} e^{-i\pi t}$ are in $L^1(\mathbb{T})$. Hence by the Riemann-Lebesgue lemma, their Fourier coefficients tend to 0 as $N \to \infty$.

□

1.2 Lecture 3

Recall Riemann localization principle.

Remark Given the localization principle, if we'd like to examine whether Fourier series converges at a particular point, we can assume $f = 0$ outside a small neighborhood of that point.

We now will prove Dini's criterion.

> **Theorem 1.4 (Dini's Criterion)**
>
> *If $x \in \mathbb{T}$, $\delta > 0$ such that*
> $$\int_{|t| < \delta} \left| \frac{f(x+t) - f(x)}{t} \right| < \infty$$
> *Then the partial sum converges at x.*

Proof Assume $\delta < \frac{1}{2}$, and we would like to show $S_N(f)(x) - f(x) \to 0$.

$$S_N(f(x) - f(x)) = \int_{|t| \leq 1/2} (f(x-t) - f(x)) D_N(t) dt$$

$$= \int_{|t| \leq \delta} (f(x-t) - f(x)) D_N(t) dt + \int_{\delta \leq |t| \leq 1/2} (f(x-t) - f(x)) D_N(t) dt$$

We notice that

$$\left| \int_{|t| \leq \delta} (f(x-t) - f(x)) D_N(t) dt \right| \leq \int_{|t| \leq \delta} \frac{|f(x-t) - f(x)|}{|\sin(t)|} < \infty$$

By the condition that the integral is ∞, we can choose a even smaller $\delta_1 < \delta$ such that $\int_{|t| < \delta} < \epsilon$. Finishing the proof.

$\square$

Remark For Holder continuous functions, they all satisfy the the assumptions of Dini. Note, a function being continuosu does not guarantee that the partial sum of Fourier series would converge.

Now we begin Jordan's Criterion.

> **Theorem 1.5 (Jordan's Criterion)**
>
> *If $x \in \mathbb{T}$, and f is of bounded variation in $(x - \delta, x + \delta)$, then we have*
> $$\lim_{N \to \infty} S_N f(x) = \frac{1}{2}(f(x+) + f(x-))$$

Remark We don't expect f to be continuous or even defined at $f(x)$.

Proof WLOG, we assume $\delta < 1/2$, and by the Riemann localization theorem, we assume $f = 0$ outside of $(x - \delta, x + \delta)$. (What happens outside of a neighborhood of δ does not matter).

Using the symmetry of the kernel. We have

$$S_N f(x)(x) = \int_{|t| \leq 1/2} f(x-t) D_N(t) dt = \int_0^{1/2} (f(x+t) + f(x-t)) D_N(t) dt$$

Let $g(t) = f(x+t) + f(x-t)$, and we would want to show that $\int_0^{1/2} g(t) D_N(t) dt \to \frac{1}{2} g(0)$.

We can also assume g is monotonic. And we want to show that $\int_{|t| \leq 1/2} g(t) D_N(t) dt$ tends to $\frac{1}{2} g(0)$. Again we separate the integral into two parts and the second part is 0, and we use the second mean value theorem of definite integrals.

> **Lemma 1.2**
>
> *For f continuous and g monotonic, we have*
> $$\int_a^b fg = g(a+) \int_a^c f(t) dt + g(b-) \int_c^b f(t) dt$$

We would want to show that $g(\delta-) \int_\nu^\delta f(t) dt$ tends to 0.

$$\left| \int_\nu^\delta D_N(t) dt \right| = \left| \int_\nu^\delta \frac{\sin((2N+1)\pi t)}{\sin(\pi t)} \right|$$

$$\leq \left| \int \sin((2N+1)\pi t)(1/\sin(\pi t) - 1/\sin(t)) \right| + \left| \int \sin((2N+1)\pi t) \frac{1}{t} \right|$$

The first term is bounded by a constant, and the second term, by integration by parts, is also bounded.

Remark If we take δ_1 arbitrarily small, then we get that the integral is arbitrarily small.

Hence we can take δ to be arbitrarily small, to finish the proof.

□

1.3 Lecture 4

We will now prove continuity is too rough of a condition to ensure Fourier series converges to f. Namely, we will show the existence of continuous functions whose Fourier series diverges at a particular point.

> **Theorem 1.6**
>
> *There exists a continuous function whose ourier series diverges at a point.*
>
> ♡

Proof WLOG, we take this point to be $x = 0$. We shall use UBP. Let $X = C(\mathbb{T}), Y = \mathbb{C}$, and equip X with the $\|\cdot\|_\infty$ norm. Hence X is a Banach space.

Then we define T_N on X as follows:

$$T_N(f) = S_N f(0) = \int_{|t| \leq 1/2} f(t) D(t) dt$$

Hence it suffices to show $\|T_N\| = \infty$ as $N \to \infty$. We define a new quantity L_N as follows:

$$L_N = \int_{|t| \leq 1/2} |D_N(t)| dt$$

We would like to show that $\|T_N\| = L_N$, and then prove L_N tends to ∞ as $N \to \infty$. We first note that we have

$$|T_N(f)| \leq \int |f(t)||D_N(t)| dt \leq \|f\|_\infty L_N$$

Then for the reverse direction, we note that $|D_N(t)| = D_N(t) \cdot sgn(D_N)(t)$, and by noting that $D_N(t)$ has only finitely many zeros, we can modify $sgn(D_N)$ into a function f with $\|f\|_\infty$, and $f \in C(\mathbb{T})$ and also that $|T_N(f)| \geq L_N - \epsilon$. Then we get

$$\|T_N\| \geq L_N - \epsilon$$

Hence $\|T_N\| = L_N$. We next show, via computation, that $L_N \to \infty$ as $N \to \infty$.

> **Lemma 1.3**
>
> *Define $L_N = \int_{-1/2}^{1/2} |D_N(t)| dt$, then we have*
>
> $$L_N = \frac{4}{\pi^2} \log(N) + O(1)$$
>
> ♡

Proof

$$
\begin{aligned}
L_N &= 2 \int_0^{1/2} \left| \frac{\sin((2N+1)\pi t)}{\sin \pi t} \right| dt \\
&= 2 \int_0^{1/2} \left| \frac{\sin((2N+1)\pi t)}{\pi t} \right| dt + O(1) \\
&= 2 \int_0^{N+1/2} \left| \frac{\sin(\pi t)}{\pi t} \right| dt + O(1) \\
&= \frac{2}{\pi} \sum_{k=0}^{N-1} \int_k^{k+1} \left| \frac{\sin(\pi t)}{\pi t} \right| dt + O(1) \\
&= \frac{2}{\pi} \sum_{k=0}^{N-1} \int_k^{k+1} \left| \frac{\sin(\pi t)}{k+1} \right| dt + O(1) \\
&= \frac{2}{\pi} \log(N) \cdot \frac{2}{\pi} + O(1)
\end{aligned}
$$

□

Now we've shown $\|T_N\| \to \infty$ as $N \to \infty$, then by UBP, there exists $f \in C(\mathbb{T})$ such that $|T_N(f)| \to \infty$, the Fourier series diverges.

$\square$

Now we would like to address the next two questions on convergence in L^p norm. Namely, do we necessarily have if $f \in L^p, 1 \leq p < \infty$, then

1. Do we have

$$\lim_{N \to \infty} \|S_N(f) - f\|_{L^p} = 0$$

2. Do we have

$$\lim_{N \to \infty} S_N(f)(x) = f(x), \text{ for almost every } x$$

We will address Q1.

Lemma 1.4

Let $1 \leq p < \infty$, for all $f \in L^p$, we have $\lim_{N \to \infty} \|S_N(f) - f\|_{L^p} = 0$ if and only if there exists a constant C_p such that

$$\|S_N(f)\|_{L^p} \leq C_p \|f\|_{L^p}$$

$\heartsuit$

Proof ($\Rightarrow$) This follows from UBP. If $\|S_N\|$ is unbounded, then by UBP, there exists an f such that $\|S_N(f)\|_{L^p} \to \infty$. Hence would not converge to f in the L^p sense.

($\Leftarrow$) We note that trigonometric polynomials are dense in L^p, hence we can find g a trig polynomial such that $\|f - g\|_{L^p} < \epsilon$. Hence we have,

$$\|S_N(f) - f\|_{L^p} \leq \|S_N(f) - S_N(g)\|_{L^p} + \|S_N(g) - g\|_{L^p} + \|f - g\|_{L^p}$$

We note the first term is

$$\|S_N(f - g)\|_{L^p} \lesssim \|f - g\|_{L^p} < \epsilon$$

Hence we get convergence in the L^p norm.

$\square$

We note that when $1 < p < \infty$, we indeed have

$$\|S_N(f)\|_{L^p} \leq C_p \|f\|_{L^p}$$

but for $p = 1$, we have $\|S_N(f)\|_{L^1} = L_N$ as above, hence there is no convergence for $S_N(f)$ to f in the L^p sense. We will reformulate this theorem as follows.

Theorem 1.7 (Convergence of $S_N(f)$ for $1 < p < \infty$)

If $1 < p < infty$, for all $f \in L^p$, we have

$$\lim_{N \to \infty} \|S_N(f) - f\|_{L^p} = 0$$

$\heartsuit$

1.4 Lecture 5

We started late today.

Recall last night, we mentioned in order to ensure $S_N(f)$ converges to f in the L^p sense, we need to have uniform boundedness. specifically for $p = 2$, the functions $e^{2\pi i n x}$ form an orthonormal basis (by trig polynomials dense in L^p), hence we get

$$\|f\|_{L^2}^2 = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2$$

Last time we asked for $1 \leq p < \infty$, and $f \in L^p$, whether we have

$$S_N(f)(x) \to f(x) \text{ almost everywhere}$$

We won't discuss in detail the answer to this question. For $p = 1$, we know that this is not true, by Kolmogov constructing an integrable function whose Fourier series diverges almost everywhere. And we know that the answer is yes for all $p > 1$.

We now begin section 1.5 of the book.

> **Definition 1.2 (Cesaro sum)**
>
> *Define the N-th Cesaro sum of Fourier series as follows:*
> $$\sigma_N(f) = \frac{1}{N+1} \sum_{k=0}^{N} S_k(f)$$
> ♣

One can show, via the following,

$$\sigma_N(f)(x) = \frac{1}{N+1} \sum_{[} k = 0]^N S_k(f)(x)$$

$$= \int_0^1 f(t) \left(\frac{1}{N+1} \sum_{k=0}^{N} D_k(x-t) \right) dt$$

$$= \int_0^1 f(t) F_N(x-t) dt$$

where $F_n(t) = \frac{1}{N+1} \sum_{k=0}^{N} D_k(t)$, we call this the Fejer kernel, and one can show that

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

We now note a few important and nice properties of $F_N(t)$. We have

1. $F_N \geq 0$
2. $\int_0^1 F_N(t) dt = \|F_N\|_{L^1} = 1$
3. $\lim_{N \to \infty} \int_{\delta \leq |t| \leq 1/2} F_N(t) dt = 0$, for all $\delta > 0$.

Note that the third property follows from the fact that

$$F_N(t) \leq \frac{1}{(N+1)(\sin(\pi t)^2)}$$

Hence it's nicely behaved when staying away from 0.

> **Theorem 1.8**
>
> *For the following two cases,*
> 1. $\forall f \in L^p$, where $1 \leq p < \infty$
> 2. $f \in C(\mathbb{T})$, for $p = \infty$.
>
> *we get convergence in L^p norm for the Cesaro sum:*
> $$\lim_{N \to \infty} \|\sigma_N(f) - f\|_{L^p} = 0$$
> ♡

Proof [sketch] For $f \in L^p$, where $1 \leq p < \infty$, we have

$$\int_{|t|<1/2} \|f(x-t) - f(x)\|_{L^p} F_N(t) dt$$

Hence separating into $\int_{|t|<\delta}$, and $\int_{\delta \leq |t| < 1/2}$, the second one tends to 0, where the first one also tends to zero as we choose δ to be arbitrarily small and by $f \in L^p$. Hence the entire integral tends to 0. Similar proof for $f \in C(\mathbb{T})$, and $p = \infty$. □

1.5 Lecture 6

From the above theorem, we are above to conclude the following statements regarding denseness of trig polynomials and the Fourier coefficients uniquely determine f, if f is integrable.

> **Corollary 1.1**
>
> *For $1 \leq p < \infty$, trigonometric polynomials are dense in L^p.*

Proof Note that $\sigma_N(f)$ is a finite average of S_k's, which are trigonometric polynomials.

> **Corollary 1.2**
>
> *For $f \in L^1$, if $\widehat{f}(k) = 0$, for all k, then f is identically 0.*

We now introduce a different summability method, by treating Fourier series as the formal limit on the unit circle.
Finish Poisson kernel!

1.6 Lecture 7

We now define FT on the most intuitive class of functions.

> **Definition 1.3 (FT on L^1)**
>
> *For $f \in L^1(\mathbb{R}^n)$, we define its Fourier transform to be the following transformation:*
> $$\widehat{f}(\xi) = \int f(x) e^{-2\pi i x \cdot \xi} dx$$
> *where $x \cdot \xi = x_1\xi_1 + \ldots x_n\xi_n$.*

FT has nice following properties that are easily verified.

> **Proposition 1.2**
>
> 1. *FT is linear.*
> $$\mathcal{F}(\alpha f + \beta g) = \alpha \hat{f} + \beta \hat{g}$$
>
> 2. *Clearly $\|f\|_\infty \leq \|f\|_1$, and by DCT, $\hat{f}$ is continuous.*
> 3. *Riemann-Lebesgue states that*
> $$\lim_{|\xi| \to \infty} \hat{f}(\xi) = 0$$
>
> 4. *Convolution in the physical space is pointwise multiplication in the Fourier space. $\widehat{f * g} = \hat{f}\hat{g}$*
> 5. *Translation in the physical space is phase shift in the Fourier space; phase shift in the physical space is translation in the Fourier space.*
> $$\widehat{f(x+h)}(\xi) = e^{2\pi i h \cdot \xi}\hat{f}(\xi)$$
> $$\widehat{e^{2\pi i h \cdot x}f}(\xi) = \hat{f}(\xi - h)$$
>
> 6. *If ρ is an orthogonal transformation, then*
> $$\widehat{f(\rho)}(\xi) = \hat{f}(\rho\xi)$$
>
> 7. *FT under scaling. If $g(x) = \lambda^{-n} f(\lambda^{-1}x)$, then $\hat{g}(\xi) = \hat{f}(\lambda\xi)$*
> 8. *FT with derivative. $\widehat{\partial_{x_j}f} = 2\pi i \xi_j \hat{f}$.*
> 9. *FT with multiplication by x_j. $\widehat{x_j f} = 2\pi i \partial_{\xi_j} \hat{f}$.*

Note that for finite dimensional measure spaces $\mathbb{T}$, by Holder's inequality $L^p(\mathbb{T})$ embeds in $L^1(\mathbb{T})$ for $p > 1$. However, the same embedding does not hold for $L^p(\mathbb{R}^n)$, and $L^1(\mathbb{R}^n)$. We've only defined the Fourier transform on $L^1(\mathbb{R}^n)$, but not all other $p > 1$ yet. To define FT on these spaces, we start by defining it on a nicer class, the Schwartz functions.

Definition 1.4 (Schwartz Space)

f is in the Schwartz space, $\mathcal{S}(\mathbb{R}^n)$ if $f \in C^\infty$, and if f and all its derivatives decrease rapidly at infinity. Rigorously, it means for all $\alpha, \beta \in \mathbb{N}^n$, we have

$$\sup_x |x^\alpha D^\beta f(x)| < \infty$$

This means at infinity, the functions decay faster than any powers of x. Note that

$$\mathcal{D} \subset \mathcal{S}$$

However, there exists elements such as $e^{-|x|^2}$ the Gaussian, which do not have compact support, yet the Gaussian is a Schwartz function.

Definition 1.5 (Tempered distributions)

$\mathcal{S}'$ is the space of linear functions on $\mathcal{S}$. $T \in \mathcal{S}'$ if for every sequence $\{\phi_n\} \subset \mathcal{S}$, such that $\phi_n \to \phi$, then we have

$$\lim_{n \to \infty} T(\phi_n) = T(\phi)$$

Remark A sequence of functions $\{\phi_n\}$ converges to $\phi \in \mathcal{S}$, if for all α, β, we have

$$\sup_x |x^\alpha D^\beta (\phi_n(x) - \phi(x))| < \epsilon$$

Theorem 1.9

The Fourier transform on $\mathcal{S}$ is continuous, and $\mathcal{S}$ is closed under taking Fourier transforms.
For any $f, g \in \mathcal{S}$, we have

$$\int \hat{f} g = \int f \hat{g}$$

Moreover, for any $f \in \mathcal{S}$, we have the Fourier inversion formula:

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

Proof For $f, g \in \mathcal{S}$, we have fg are integrable over $\mathbb{R}^n \times \mathbb{R}^n$. Hence

$$\int \hat{f} g = \int f \hat{g}$$

follows from Fubini's theorem. In the following lemma, we show if $f = e^{-\pi|x|^2}$, then $\hat{f} = f$. Given this, we have that

$$f(0) \int \hat{g} = g(0) \int \hat{f}$$

If we take $g(x) = e^{-\pi|x|^2}$, thus we have

$$f(0) = \int \hat{f}(\xi) d\xi$$

Then a translation in the physical space is a phase shift in the Fourier space, hence we have

$$f(x) = \int e^{2\pi i x \cdot \xi} \hat{f}(\xi) d\xi$$

Lemma 1.5 (Gaussian is invariant under FT)

If $f = e^{-\pi|x|^2}$, then we have

$$\hat{f}(\xi) = e^{-\pi|\xi|^2}$$

Proof We note that it suffices to show this for the $n = 1$ case. And $f = ^{-\pi x^2}$ solves the differential equation below:

$$\begin{cases} u' + 2\pi x u = 0 \\ u(0) = 1 \end{cases}$$

To show $\hat{f} = f$, it suffices to show $\hat{f}$ solves the same equation and by uniqueness of solution.

We take the Fourier transform of both sides, we get

$$(u')\widehat{} + (2\pi xu)\widehat{}$$

Hence we get

$$2\pi i\xi\hat{u} + \partial_\xi u = 0$$

And also

$$\hat{u}(0) = \int u(x)dx = \int e^{-\pi x^2} = 1$$

Hence by uniqueness of ode, we get $\hat{u} = u$.

$\square$

$\square$

Once we have the Fourier inversion formula for $f \in \mathcal{S}$, we note that

$$(\hat{f})\widehat{}(x) = \int \hat{f}(\xi)e^{-2\pi ix\cdot\xi}d\xi = f(-x)$$

> **Corollary 1.3**
>
> *The Fourier transform has period 4.*
>
> ♡

Next, we define Fourier transform on the space of tempered distributions, $\mathcal{S}'$.

> **Definition 1.6**
>
> *If $T \in \mathcal{S}'$, then we define its Fourier transform as follows:*
> $$\hat{T}(f) = T(\hat{f}), f \in \mathcal{S}$$
>
> ♣

We do the check that $\hat{T} \in \mathcal{S}'$ is well defined. We check $\hat{T}$ is linear:

$$\hat{T}(f + g) = T(\hat{f} + \hat{g}) = T(\hat{f}) + T(\hat{g}) = \hat{T}(f) + \hat{T}(g)$$

Similarly for scalar multiplication.

> **Theorem 1.10**
>
> *FT is a linear, bounded bijection from $\mathcal{S}' \to \mathcal{S}'$. Moreover, it is invertible and its inverse is bounded as well.*
>
> ♡

Proof It is bounded if and only if it is continuous, hence we check for $T_n \to T$ in $\mathcal{S}'$, we have, for any $f \in \mathcal{S}$,

$$\hat{T}_n(f) = T_n(\hat{f}) \to T(\hat{f}) = \hat{T}(f)$$

Hence $\hat{T}_n \to \hat{T}$. And the invertability follows from FT has period 4.

$\square$

1.7 Lecture 8

Now we've defined FT on $\mathcal{S}$, by density of $\mathcal{S}$ in L^p, we have a well-defined FT on L^p. We start by proving the Plancherel identity.

> **Theorem 1.11**
>
> *FT on L^2 is an isometry, i.e.*
> $$\|f\|_{L^2} = \|\hat{f}\|_{L^2}, f \in L^2$$
>
> ♡

Proof For any $f, h \in \mathcal{S}$, we let $\hat{g} = \overline{h}$, then we have

$$g = (\overline{h})\check{} = \int h(-x)e^{2\pi ix\cdot\xi}d\xi = \int h(y)e^{-2\pi iy\cdot\xi}d\xi = \hat{\overline{h}}$$

Hence by the duality formula, we have

$$\int f\overline{h} = \int f\hat{g} = \int \hat{f}g = \int \hat{f}\hat{\overline{h}}$$

Setting $f = h$, we get

$$\|f\|_{L^2} = \|\hat{f}\|_{L^2}$$

$\square$

Next we stated Riesz-Thorin interpolation theorem and used it to show Hausdorff Young, and Young's inequality. We leave the proof of Riesz-Thorin as a homework question.

Theorem 1.12 (Hausdorff Young)

For $1 \leq p \leq 2$, if $f \in L^p$, then $\hat{f} \in L^{p'}$, and

$$\|\hat{f}\|_{L^{p'}} \leq \|f\|_{L^p}$$

$\heartsuit$

Proof This follows from noting the inequality holds for $p = 1, p' = \infty$, and $p = p' = 2$.

$\square$

Theorem 1.13 (Young's inequality)

For $f \in L^p, g \in L^q$, if we let $\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}$, then we get

$$\|f * g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q}$$

$\heartsuit$

Proof We note that

$$\|f * g\|_{L^p} = \left(\int \left| \int f(y)g(x-y)dy \right|^p dx \right)^{1/p} \leq \|f\|_{L^p}\|g\|_{L^1}$$

And by Holder, we have

$$\|f * g\|_{L^\infty} \leq \|f\|_{L^p}\|g\|_{L^{p'}}$$

And then we apply Riesz-Thorin.

$\square$

1.7.1 Convergence and partial summability methods

Recall what we did with the Fourier series: given $S_N(f) = \sum_{n=-N}^{N} \hat{f}(n)e^{2\pi inx}$, what can we say about

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

And in both pointwise and L^p sense. We now introduce the parallel concept with Fourier transform. We ask the question: when do we have

$$\lim_{R \to \infty} \int_{B_R} \hat{f}(\xi)e^{2\pi ix \cdot \xi} = f(x)$$

And we ask when this will converge to the original function.

Definition 1.7 (Paritial Sum operator)

Like how we want the partial sum of Fourier series to capture how f behaves, we define the partial sum operator for Fourier transform as follows:

$$(S_R(f))\hat{} = \chi_{B_R}\hat{f}$$

$\clubsuit$

Note that in the Fourier space, $S_R(f) \to \hat{f}$, and we would like to ask the question, does this $S_R(f)$ approximate f in the physical space? In other words, this becomes the question whether we have

$$\lim_{R \to \infty} S_R(f) = f$$

Note that we have

$$S_R(f) = \check{X_{B_R}} * f := D_R * f$$

where $D_R(x) := \int_{B_R} e^{2\pi ix \cdot \xi}d\xi$.

> **Proposition 1.3**
>
> *(Like with the Fourier series), we have*
> $$\lim_{R \to \infty} \|S_R(f) - f\|_{L^p} = 0$$
> *if and only if, there exists $C > 0$ such that*
> $$\|S_R(f)\|_{L^p} \le C\|f\|_{L^p}$$
> ♠

Remark We have $\|S_R\| < \infty$ for $n = 1$, but for $n > 1$, this is only true for $p = 2$. We will discuss this in the upcoming lectures.

> **Definition 1.8 (Dirichelt Kernel)**
>
> *We've defined the Dirichlet series as $D_N(x) = \sum_{n=-N}^{N} e^{2\pi i n x}$, and likewise, in $n = 1$, we define the Dirichlet kernel as*
> $$D_R(x) = \int_{-R}^{R} e^{2\pi i x \cdot \xi} d\xi$$
> ♣

Note that
$$\int_{-R}^{R} e^{2\pi i x \cdot \xi} d\xi = \frac{e^{2\pi i x \xi}}{2\pi i x}\Big|_{-R}^{R} = \frac{1}{2\pi x}\sin(2\pi R x) - \sin(-2\pi R x) = \frac{\sin(2\pi R x)}{\pi x}$$

> **Lemma 1.6**
>
> *(We might use this in the future). We have*
> $$D_R \notin L^1$$
> *However, for any $p > 1$, we have*
> $$D_R \in L^p$$
> ♡

Proof We have $\int_{-R}^{R} \sin(2\pi R x)/(\pi x)$ not integrable, however for $p > 1$, we have
$$\int |D_R|^p \lesssim \int_{-R}^{R} \frac{1}{|x|^p} < \infty$$

□

1.8 Lecture 9

Recall we defined Dirichlet kernel for $n = 1$, we first note that when $n > 1$, we have D_R integrates over a sphere with respect to a surface measure. Now we introduce partial summability methods just like with the Fourier series.
$$\sigma_R(f) = \frac{1}{R} \int_0^R S_r(f) dr$$
We have
$$\sigma_R(f)(x) = F_R * f(x)$$
where $F_R(x) = \frac{1}{R} \int_0^R D_r(x) dr$. We thus have
$$F_R(x) = \frac{1}{R\pi x}(1 - \cos(2\pi x R)/2\pi x) = \frac{\sin(\pi R x)}{R(\pi x)^2}$$
We have
1. $F_R \ge 0$
2. $\|F_R\|_{L^1} = 1$.

Hence it's again, a redistribution of mass, and as $R \to \infty$, F_R acts more and more like the dirac mass, the convolution identity. For $1 \le p < \infty$, we have
$$\lim_{R \to \infty} \|\sigma_R(f) - f\|_{L^p} = 0$$

Chapter 2 Hardy-Littlewood Maximal function

We will talk about partial sum convergence, and better behaved summability methods for $S_R f$, and we disucss both cases in the sense of L^p and a.e. pointwise convergence.

2.0.1 Chapter 2.1

Definition 2.1 (Approximation of the identity)

Let $\phi \in L^1(\mathbb{R}^n)$, and $\int \phi = 1$, and we define

$$\phi_t = t^{-n}\phi\left(\frac{x}{t}\right)$$

And we call the family $\{\phi_t\}_{t>0}$ approxiation of the identity.

♣

We shall justify the name.

Proposition 2.1

Let ϕ_t be defined above, then for any $f \in \mathcal{S}$, we have

$$\lim_{t\to 0}\phi_t * f(x) = f(x)$$

♠

Proof It suffices to show that $\phi_t \to \delta$ in $\mathcal{S}'$, since it would imply

$$\lim_{t\to 0}\phi_t * f(x) = \delta * f(x) = f(x)$$

And for any $f \in \mathcal{S}$, we have

$$\phi_t(f) = \int t^{-n}\phi\left(\frac{x}{t}\right)f(x)dx = \int \phi(y)f(ty)dy$$

Hence as $t \to 0$, we have

$$\lim_{t\to 0}\phi_t(f) = \int \phi(y)f(0)dy = f(0)$$

Hence $\phi_t \to \delta$, as $t \to 0$ in $\mathcal{S}'$.

$\square$

Remark Note that for Cesaro partial sum, we have the Fejer kernel as $F_R = \frac{1}{R}\int_0^R D_r(x)dr$, and in fact $F_R = \phi_{1/R}$.

Next we see that the family of ϕ_t allows for convergence in L^p norm as well.

Theorem 2.1

Let $\{\phi_t\}_{t>0}$ be an approximation of the identity, then for any $f \in L^p$, $1 \le p < \infty$, we have

$$\lim_{t\to 0}\|\phi_t * f - f\|_{L^p} = 0$$

♡

Proof We have

$$\phi_t * f = \int \phi(y)[f(x-ty) - f(x)]dy$$

Note that there exists $\delta > 0$, for $|ty| < \delta$, we have

$$\|f(x-ty) - f(x)\|_{L_x^p} < \frac{\epsilon}{2\|\phi\|_{L^1}}$$

Hence we separate the integral to $|y| < \frac{\delta}{t}$, and $|y| > \frac{\delta}{t}$. We thus have

$$
\begin{aligned}
\|\phi_t * f - f\|_{L^p} &\leq \int |\phi(y)| \|f(x - ty) - f(x)\|_{L^p_x} dy \\
&= \int_{|y|<\delta/t} |\phi(y)| \|f(x - ty) - f(x)\|_{L^p_x} dy + \int_{|y|>\delta/t} |\phi(y)| \|f(x - ty) - f(x)\|_{L^p_x} dy \\
&< \frac{\epsilon}{2\|\phi\|_{L^1}} \int_{|y|<\delta/t} |\phi(y)| dy + 2\|f\|_{L^p} \int_{|y|>\delta/t} |\phi(y)| dy \\
&< \epsilon
\end{aligned}
$$

$\square$

This convergence in L^p implies that there exists a sequence t_n such that as $t_n \to 0$, for almost every x, we have

$$\lim \phi_{t_n} * f(x) \to f(x)$$

2.1 Lecture 10

Last time we introduced the approxiation to identity funcitons, and

Exercise 2.1 If $\{\phi_t\}$ are approxiation of identities and $f * \phi_t$ converges a.e. for all $f \in L^p, 1 \leq p < \infty$, then it converges to f pointwise almost everywhere.

Remark Note that you can pick $\{\phi_t\}$ compactly supported and smooth, and would result in a Fourier transform that preserves where f originally lies, i.e. if $f \in L^p$, we have $\phi_t * f \in L^p$ as well. Even you can't take derivatives of f, $\phi * f$ is smooth for $\phi \in \mathcal{S}$, since convolution picks up the better derivatives.

Now we go back to the question of almost everywhere pointwise convergence.

Fix f, we could approximate f with a sequence by $\{f_n\} \subset \mathcal{S}$, for $f \in C_\infty(X)$, we do get almost everywhere convergence by remark from last time.

Now we look at

$$\phi_t(f - f_n)$$

we would want this as an "error" for all large n, and all small t on some $E \subset \mathbb{R}^n$, such that $m(\mathbb{R}^n \setminus E) < \epsilon$. This guarantees almost everywhere convergence, $\phi_t(f_n - f)$ could behave wildly on $\mathbb{R}^n \setminus E$, and it would not have mattered.

Note this is a two parameter family in t and n, taking advantage of what ϕ_t have in common, for all t.

1. $\int \phi_t = 1$.
2. They all seem to have some sort of bump near 0, and could be thought of as an average.

2.1.1 Weal type inequalities and a.e. convergence

> **Theorem 2.2**
>
> *Let $\{T_t\}$ be a family linear of sublinear operators on $L^p(X, \mu)$, $t \in A \subset \mathbb{R}$, and define*
>
> $$T^* f(x) = \sup_t |T_t f(x)|$$
>
> *this is defined as the maximal function. If T^* is of weak type-(p, q), then the set*
>
> $$\{f \in L^p(X, \mu) : \lim_{t \to t_0} T_t(f)(x) = f(x), a.e.\}$$
>
> *is closed in $L^p(X, \mu)$.*
>
> $\heartsuit$

Remark If $T : L^p \to L^q$, then if T is bounded, we call it strong-type (p, q), and strong type implies that T is of weak type (p, q). In most situations, we will define $p = q$.

Note *Note that we will use the word "sublinear" in the definition in the future, where we wanted to show the sublinear operator $\sup |\phi_t * f(x)|$ converges to $\int \phi f(x)$ almost everywhere, given that $\sup |\phi_t * f(x)|$ is weak (1,1).*

Let (X, μ), (Y, ν) be measure spaces, and T be an operator on $L^p(X, \mu)$ to measurable functions on L^q. All elements are ≥ 1.

> **Definition 2.2 (weakly bounded)**
>
> *For $q < \infty$, T is weak type (p, q) if there exists $C > 0$,*
> $$\nu\left(\{y \in Y : |T(f)(y)| > \lambda\}\right) \leq \left(\frac{C\|f\|_{L^p}}{\lambda}\right)^q$$
>
> ♣

Example 2.1 For $p = q = 1$, we take $X = Y = \mathbb{R}$. We allow $T(f)$ to be the function $1/x$, which is not bounded of strong type.

Exercise 2.2 Weak type (p, ∞) is equivalent to saying bounded from L^p to L^∞.

Proof For T of weak type (p, ∞), we have
$$\lambda\left(\nu\left(\{y \in Y : |T(f)(y)| > \lambda\}\right)\right)^{\frac{1}{p}} \leq C\|f\|_{L^p}$$
Hence we have $p \to \infty$, $\left(\nu\left(\{y \in Y : |T(f)(y)| > \lambda\}\right)\right)^{\frac{1}{p}} \to 1$ for any $\nu\left(\{y \in Y : |T(f)(y)| > \lambda\}\right) \neq 0$, hence if we let $\lambda \to \infty$, we get a contradiction. Hence we could allow allow $\nu\left(\{y \in Y : |T(f)(y)| > \lambda\}\right) < \epsilon$, for all $\epsilon > 0$.

□

Exercise 2.3 Strong type (p, q) implies the weak type (p, q).

Proof We denote the set $\{y \in Y : |T(f)(y)| > \lambda\}$ as E_λ, then we have
$$\nu(E_\lambda) = \int_{E_\lambda} 1 d\nu \leq \int_{E_\lambda} \left(\frac{|T(f)y|}{\lambda}\right)^q d\nu \leq \left(\frac{C\|f\|_{L^p}}{\lambda}\right)^q$$
in the last inequality, we used the fact that T is bounded from L^p to L^q.

□

Proof [of theorem] We would like to show the set is closed, hence we take a sequence of $\{f_n\}$ such that $\|f_n - f\|_{L^p} \to 0$, and $T_t(f_n)(x) \to f_n(x)$ for almost every x.

WLOG assume $t_0 = 0$. Assume $f_n \to f$ in L^p, and that
$$T_t(f_n)(x) \to f_n(x) \text{ for almost every } x$$
as $t \to 0$ a.e. for every f_n.

The idea of "oscillation." The goal is to show $T_t(f) \to f$ as $t \to 0$ almost everywhere.

Define
$$W_\lambda = \{x : \limsup_{t \to 0} |(T_t(f) - f)(x)| > \lambda\}$$
for each $\lambda > 0$. We are done if we show
$$\mu(W_\lambda) = 0$$

Note that
$$W_\lambda \subset \{x : \limsup_{t \to 0} |(T_t(f) - T_t(f_n))(x)| > \lambda/3\} \cup \{x : |(f - f_n)(x)| > \lambda/3\}$$
for any fixed n. Note that if we denote the first set as A, the second as B, then
$$\mu(B) \lesssim \left(\frac{\|f - f_n\|_{L^p}}{\lambda}\right)^p$$

For A, note that
$$A \subset \{x : T^*(f - f_n)(x) \geq \lambda/3\}$$

By weak type (p, q) of T^*, hence we have
$$\mu(B) \lesssim \left(\frac{\|f - f_n\|^p}{\lambda}\right)^q$$

We can make $\mu(A), \mu(B) < \epsilon$, by choosing n large enough. Hence we have
$$\mu(W_\lambda) < 2\epsilon$$

□

> **Definition 2.3 (Sublinear operator)**
>
> *T is called sublinear, we have*
> $$T(f+g) \leq T(f) + T(g), T(\lambda f) = \lambda T(f)$$
> ♣

2.2 Lecture 11

We will begin with one comment. (This is in 8.3 of the book)

> **Definition 2.4 (Lorentz space)**
>
> *For $1 \leq p < \infty$, we take $f \in L^{p,\infty}$ if and only if*
> $$\lambda^p |\{|f| \geq \lambda\}| < C(f), \text{ for all } \lambda$$
> ♣

Note that this connects to the weak type definition that we introduced, we have T is weak (p,p), if and only if

$$\|Tf\|_{p,\infty} \lesssim \|f\|_{L^p}$$

Note that $\|\cdot\|_{p,\infty}$ is a quasinorm, however, for $p \neq 1$, we have $\|\cdot\|_{p,\infty}$ generates a normable sapce.

> **Definition 2.5 (dyadic number)**
>
> *For $k \in \mathbb{Z}$, we have $\lambda = 2^k$ is a dyadic number.*
> ♣

We could decompose $\mathbb{R}$ into dyadic intervals.

$$\mathbb{R}^+ = ... \bigsqcup (1,2] \bigsqcup (2,4]...$$

This type of decomposition is mild, i.e. if we try to estimate the L^p norm of a function, then it is not too different from knowing how f behaves within one λ interval.

> **Definition 2.6 (mid-level sets)**
>
> *For dyadic λ, define*
> $$b_\lambda(f) = |\{\lambda : \lambda < |f(x)| \leq 2\lambda\}|$$
> ♣

✎ **Exercise 2.4**
  1. $f \in L^p$ if and only if $\sum_{\lambda, dyadic} \lambda^p b_\lambda(f) < \infty$.
  2. $f \in L^p$, if and oly if each $\lambda^p b_\lambda(f)$ is uniformly bounded.

Remark Note that the only way for $\sum_\lambda^p b_\lambda(f)$ to blow up is to have many $b_\lambda(f)$ to be 1 for n many lambda's, i.e. blow up in a log fashion.

We recall that for $f \in \mathcal{S}$, we have

$$\phi_t * f \to f \text{ pointwise a.e.}$$

Hence if $\limsup_{t \to 0} |\phi_t * f(x)|$ is weak type (p,p), we know a.e. convergence for $p \in [1, \infty)$. This is indeed true, and we will prove it in the upcoming lectures.

✎ **Exercise 2.5** Show that this holds for weak(∞, ∞).

2.2.1 Chapter 2.3 Marcinkiewicz interpolation theorem

This interpolation thoerem allows us to reduce the case to $p = 1$ (given that we know it holds for $p = \infty$). for the convergence of $|\phi_t * f|$.

Remark We cannot apply Riesz Thorin, since $\limsup_{t \to 0} |\phi_t * f(x)|$ is not linear.

> **Theorem 2.3 (Marcinkiewicz interpolation theorem)**
>
> *Let $(X, \mu), (Y, \nu)$ be measure spaces, and $1 \leq p_0 < p_1 \leq \infty$. Let T be a sublinear operator form $L^p(X, \mu) + L^{p_1}(X, \nu)$ to measurable functions on Y.*
>
> *If T is weak type (p_0, p_0) and weak (p_1, p_1), then T is strong type (p, p), for $p_0 < p < p_1$.*
>
> ♡

Remark Marcinkiewicz is applicable to sublinear operators, but only from $L^p \to L^p$. Riesz-Thorin only allows for linear operators, but could go from $L^p \to L^q$. The complex interpolation that stems from Riesz-Thorin also can be applied to a family of operators.

Remark Tomas-Stein constructed a family of operators, and then we applied Riesz-Thorin.

We first introduce the definition of distribution function of f.

> **Definition 2.7 (distribution function of f)**
>
> *Let $f : X \to \mathbb{C}$ be measurable, define*
> $$a_f(\lambda) = \mu(\{x \in X : |f(x)| > \lambda\})$$
> *to be the distribution function of f, and note that it is a decreasing function of λ. Sets of the form $\{x \in X : |f(x)| > \lambda\}$ are sometimes called the super-level sets.*
>
> ♣

> **Proposition 2.2**
>
> *Let $\phi : [0, \infty) \to [0, \infty)$ be differentiable, increasing with $\phi(0) = 0$, then we have*
> $$\int_X \phi(|f(x)|)d\mu = \int_0^\infty \phi'(\lambda)a_f(\lambda)d\lambda$$
>
> ♠

Proof We note that for the LHS, we have
$$\int_X \phi(|f(x)|)d\mu = \int_X \int_0^{|f(x)|} \phi'(\lambda)d\lambda d\mu = \int_0^\infty \phi'(\lambda)\int_{\{x:|f(x)|>\lambda\}} d\mu d\lambda = \int_0^\infty \phi'(\lambda)a_f(\lambda)d\lambda$$
$\square$

> **Corollary 2.1 (L^p norm in terms of $a_f(\lambda)$)**
>
> *Following is an alternative way to express the L^p norm in terms of the distribution function of f.*
> $$\int |f(x)|^p dx = \int_0^\infty p\lambda^{p-1}a_f(\lambda)d\lambda$$
>
> ♡

✎ **Exercise 2.6** What is the LHS in terms of $b_f(\lambda)$.

Proof [of the interpolation theorem]

Let $f = f_0 + f_1$, where $f_0 = f\chi_{x:|f(x)|>c\lambda}$, and $f_1 = f\chi_{x:|f(x)|\leq c\lambda}$, for some $c > 0$ to be determined later.

We have a simple observation that
$$a_{Tf}(\lambda) \leq a_{Tf_0}\left(\frac{\lambda}{2}\right) + a_{Tf_1}\left(\frac{\lambda}{2}\right)$$

We will prove the $p_1 < \infty$ case.

✎ **Exercise 2.7** Prove it for $p_1 = \infty$ case.

Proof For $p_1 = \infty$, we have T as a bounded operator from L^∞ to L^∞, hence we denote $\|T\|_{L^\infty \to L^\infty} = A_1$.

Note that we have

$$\int_X |Tf|^p d\nu = \int_0^\infty p\lambda^{p-1} a_{Tf}(\lambda) d\lambda$$

$$\leq p \int_0^\infty \lambda^{p-1} a_{Tf_0}\left(\frac{\lambda}{2}\right) d\lambda + p \int_0^\infty \lambda^{p-1} a_{Tf_1}\left(\frac{\lambda}{2}\right) d\lambda$$

$$= p \int_0^\infty \lambda^{p-1} \left(\frac{A_0}{\lambda/2}\|f_0\|_{L^{p_0}}\right)^{p_0} d\nu + p \int_0^\infty \lambda^{p-1}\left(\frac{A_1}{\lambda/2}\|f_1\|_{L^{p_1}}\right)^{p_1} d\nu$$

$$\leq p \int_0^\infty \lambda^{p-1} A_0^{p_0} 2^{p_0} \lambda^{-p_0} \int_{\{x: |f(x)|>c\lambda\}} |f(x)|^{p_0} d\mu + \text{ similar term}$$

$$= p \int |f(x)|^{p_0} 2^{p_0} A_0^{p_0} \int_0^{|f(x)|/c} \lambda^{p-p_0-1} d\lambda$$

Optimizing c, we have

$$\|Tf\|_{L^p} 2p^{1/p} \left(\frac{1}{p-p_0} + \frac{1}{p_1-p}\right)^{\frac{1}{p}} A_0^{1-\theta} A_1^\theta \|f\|_{L^p}$$

where $A_0 = \|T\|_{L^{p_0} \to L^{p_0}}, A_1 = \|T\|_{L^{p_1} \to L^{p_1}}$.

Note that for the first term, we have

$$\|f_0\|_{p_0}^{p_0} = \int |f_0|^{p_0} d\mu \leq (c\lambda)^{p_0-p} \int |f_0|^p d\mu = (c\lambda)^{p_0-p}\|f\|_{L^p}^p$$

$\square$

2.3 Lecture 12

We now begin our discussion of Hardy-Littlewood maximal function.

> **Definition 2.8 (Hardy-Littlewood maximal function)**
>
> *Let $f \in L_{loc}^1(\mathbb{R}^n)$, then we define the Hardy-Littlewood maximal function as follows:*
> $$Mf(x) = \frac{1}{|B_r|} \int_{B_r} |f(x-y)| dy$$
> *where B_r is the ball of radius of r around 0.*
> *Alternatively, we have the following equivalent definition of the Hardy-Littlewood maximal function:*
> $$Mf(x) = \frac{1}{|B_r(x)} \int_{B_r(x)} |f(y)| dy$$
> *where $B_r(x)$ is the ball of radius r centered at x.*
>
> ♣

Note that this is equivalent to $M'f$ where the balls are replaced with cubes, and we could non-center the tubes to define $M''(f)$.

> **Theorem 2.4**
>
> *M is weak (1,1) and strong (p,p) for $1 < p \leq \infty$.*
>
> ♡

Proof We have the following:

$$\|Mf\|_\infty \leq \|f\|_\infty$$

Hence M is bounded from $M: L^\infty \to L^\infty$.

Hence by Marcinkiewicz interpolation theorem, it suffices to prove that M is weak (1,1). We investigate $a_{Mf}(\lambda)$.

> **Lemma 2.1**
>
> *Let $\{I_\alpha\}$ be a collection of open intervals of $\mathbb{R}$, and let K be a compact subset such that $K \subset \bigcup_\alpha I_\alpha$. There exists a finite subcover such that $\{I_j\}_{j=1}^n$ such that for each $x \in \mathbb{R}$, it is not contained in more than two I_j's.*
>
> ♡

✍ **Exercise 2.8** complete the proof of the lemma. Hint: for each x, consider all the intervals that x is contained in, and only keep the left and right most ones.

Now for we prove that M is weak (1,1) for $n = 1$, we let

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

We would like to show that $\mu(E_\lambda) \lesssim \|f\|_{L^1}$. It suffices to show this for any K compact contained in E_λ.

Note that for each $x \in E_\lambda$, there exists an interval I_x (this is by the alternative definition of the Hardy-Littlewood maximal function) such that

$$\int_{I_x} |f| \leq \lambda |I_x|$$

Hence by lemma above, for any compact set K contained in I_x, there exists a finite subcollection of $\{I_j\}$, such that

$$\sum_j \chi_{I_j} \leq 2$$

Thus we have the following:

$$|K| \leq \sum_j |I_j| \leq \sum_j \frac{1}{\lambda} \int_{I_j} |f| dy = \frac{1}{\lambda} \int_{\mathbb{R}} |f(y)| \sum_j \chi_{I_j} dy \leq \frac{1}{\lambda} \|f\|_{L^1}$$

□

> **Proposition 2.3 (Hardy-Littlewood implies a.e. for some summability methods)**
>
> Let ϕ be positive, radial and decreasing, and $\phi \in L^1$, then
>
> $$\sup_{t>0} |\phi_t * f(x)| \leq \|\phi\|_{L^1} Mf(x)$$
>
> ♠

Proof One could show this using approximation, i.e. approximate ϕ by step funcitons such that in Figure 1. We let $\phi = \sum_j a_j \chi_{B_j}(x)$, with $a_j \geq 0$. It suffices to show the proposition holds for ϕ simples functions like above.

We have

$$\sup_{t>0} \|\phi_t * f(x)\| \leq \sum_j \sup_{t>0} |a_j \chi_{B_j} * f(x)| \leq \sum_j a_j |B_j| \frac{1}{|B_j|} \int_{B_j} f(x-y) \leq \sum_j a_j |B_j| Mf(x) = \|\phi\|_{L^1} Mf(x)$$

Then we approximate ϕ using this in the general case.

□

☞ **Note** *Under appropriate conditions, this allows us to say the sublinear operator $\sup |\phi_t * f(x)|$ is weak (1,1).*

> **Corollary 2.2**
>
> *If $|\phi| \leq \psi$, ψ such that it is positive, radial, and decreasing, then we have $\sup_t |\phi_t * f|$ is weak (1,1), and strong (p,p), for $1 < p \leq \infty$.*
>
> ♡

Hence we resolved a pointwise convergence issue that we raised previously.

> **Corollary 2.3**
>
> *If $|\phi| \leq \psi$, and ψ is such that is positive, radial and decreasing, then in particular for Cesaro, Abel-Poisson, and Guass-Weierstrass kernels, we have if $f \in L^p$, $1 \leq p < \infty$, or if $f \in C_0$, then we have*
>
> $$\lim_{t\to 0} \phi_t * f = \left(\int \phi(y) dy\right) \cdot f, \text{ almost everywhere}$$
>
> ♡

2.3.1 Dyadic maximal function

We define a r-cube in $\mathbb{R}^n$ by $[0,1)^n$, and translating it.

Definition 2.9

We define $Q_0 = \{$ lattice cubes in $\mathbb{R}^n\}$, where they are unit cubes such that the corners are lattices. We define the dyadic lattice cubes as follows:

$$Q_k = \{2^{-k} - \text{ unit cubes in} \mathbb{R}^n, k \in \mathbb{Z}\}$$

$\bigcup_{k \in \mathbb{Z}} Q_k$ is dyadic cubes.

♣

2.4 Lecture 13

We continue with dyadic maximal functions that we introduced last time.

There are some nice properties that we would like to note about dyadic cubes.

1. $\forall x \in \mathbb{R}^n$, and for all $k \in \mathbb{Z}$, there exists a unique $Q \in \mathcal{Q}_k$ such that $x \in Q$.
2. Any two dyadic cubes either are disjoint, or one is contained in the other. In other words, if two dyadic cubes intersect, then one is contained in the other.
3. For each $j < k$, all $Q \in \mathcal{Q}_k$ is contained in a unique $Q \in \mathcal{Q}_j$, and contains 2^n cubes in $\mathcal{Q}_{k+1}$.

Definition 2.10 (Dyadic maximal function)

For $f \in L^1_{loc}(\mathbb{R}^n)$, define the conditional expectation of f as follows:

$$E_k f(x) = \sum_{Q \in \mathcal{Q}_k} \left(\frac{1}{|Q|} \int_Q f \right) \chi_Q(x)$$

And define the dyadic maximal function M_d as follows:

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

♣

Note *We notice that the conditional expectation of f satisfy the following nice property: let Ω be the collection of all dyadic cubes in $\mathcal{Q}_k$, then we have*

$$\int_\Omega E_k f = \int_\Omega f$$

This can be seen as follows:

$$\int_\Omega \sum_{Q \in \mathcal{Q}_k} \frac{1}{|Q|} \int_Q f(y) dy \chi_Q(x) = \sum_{Q \in \mathcal{Q}_k} \int_Q f(y) dy = \int_{\Omega = \mathcal{Q}_k} f(y) dy \quad (2.1)$$

we will use this to prove dyadic maximal function is weak (1,1).

Alternatively, we could define the dyadic maximal function in a more intuitive way.

Definition 2.11 (Dyadic maximal function-take 2)

Let Q be dyadic cubes of different sizes, then the following is the dyadic maximal function:

$$M_d f(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| dy$$

where we take the supremum of all Q such that $x \in Q$.

♣

Theorem 2.5 (dyadic maximal function is weak (1,1))

We have the two following for the dyadic maximal function.

1. *The dyadic maximal function is weak $(1, 1)$*
2. *If $f \in L^1_{loc}$, we have*

$$\lim_{k \to \infty} E_k f(x) = f(x), \text{ almost everywhere}$$

This should be intuitive, since as $k \to \infty$, our dyadic cube gets smaller and smaller, should recover how $f(x)$ behaves. This is the same as, in our alterantive definition, as $l(Q) \to 0$.

♡

Proof Consider $f \in L^1$, and $\lambda > 0$, WLOG and $f \geq 0$. For each x such that $M_d f(x) > \lambda$, we can find $Q_x \in \mathcal{Q}_k$ for some k, and we have

$$\int_{Q_x} f dx > \lambda |Q_x|$$

And we note that this k is $>$ some constant, depending on f, since the L^1 norm of the function is bounded. (We should have $|Q_x| = 2^{-kn} < \infty$, hence k greater than some constant).

Proof sketch: we would like to show that for each x such that $M_d f(x) > \lambda$, we can put this x in some dyadic Q such that it satisfies $\int_Q f(y) dy > \lambda |Q|$, hence the volume of Q is bounded, and such all Q are all disjoint (hence we control how much they overlap) when we take the integral, is bounded by $\|f\|_{L^1}$.

Now we use the property if two dyadic cubes intersect, then one is contained in another. Take all such Q_x to form a collection of S, and take a subcollection of S as follows:

Put all $Q \in S \cap Q_{k+1}$ in S' And all $Q \in S \cap O_j$ such that Q not contained in every $\tilde{Q}$ in $S' \cap (Q_{k_0+1} \cup ... \cup Q_{j-1})$ in S' for $j > k_0 + 1$.

We are left with a subcollection of S' such that each x such that $M_d f(x) > \lambda$ is contained in some $Q \in S'$. (contained in a larger one) And the Q's in S' are disjoint and satisfy

$$\int_Q f dx > \lambda |Q|$$

Hence we have

$$|\{x : M_d f(x) > \lambda\}| \leq |\bigsqcup_{Q \in S'} Q| \leq \sum_{Q \in S'} |Q| \leq \frac{1}{\lambda} \int_{\bigsqcup Q} f dx \leq \frac{1}{\lambda} \|f\|_{L^1}$$

$\square$

Remark We can make the chocie such that on the parent of Q, the average of f is $\leq \lambda$. This ensures that

$$\lambda < \frac{1}{|Q|} \int_Q f \leq 2^n \lambda$$

Recall that weak (1,1) implies the almost everywhere convergence of good functions $f \in L^1$. This is theorem 2.2. To go from L^1 to L^1_{loc}, we note that this is a local property, as $k \to \infty$, and it is still contained in a ball (i.e. cover $\mathbb{R}^n$ with balls).

The book gives the alterantive proof on dyadic maximal function being weak (1,1) using the same ideas but more straight forward "disjointness" method. We will also insert it here.

Proof [dyadic maximal function is weak (1,1) take 2] If we examine the set $\{x : M_d f(x) > \lambda\}$, we would like to write it as a disjoint union. Hence we have the following:

$$\{x : M_d f(x) > \lambda\} \subset \bigsqcup_k \Omega_k$$

where $\Omega_k = \{x : E_k f(x) > \lambda \text{ and } E_j f(x) \leq \lambda, \forall j < k\}$. This ensures that all Ω_k's are disjoint. And we have $\int_{\Omega_k} E_k f(x) \geq \int_{\Omega_k} \lambda = |\Omega_k|\lambda$, hence we have the following:

$$|\{x : M_d f(x) > \lambda\}| = \sum_k |\Omega_k|$$

$$\leq \sum_k \frac{1}{\lambda} \int_{\Omega_k} E_k f(y) dy$$

$$= \sum_k \frac{1}{\lambda} \int_{\Omega_k} f(y) dy$$

$$\leq \frac{1}{\lambda} \|f\|_{L^1}$$

where we noted that we could write Ω_k as a union of disjoint dyadic cubes, and using $\int_{\Omega_k} E_k f(x) dx = \int_{\Omega_k} f(x) dx$.

$\square$

> **Theorem 2.6 (Calderon-Zygmund decomposition)**
>
> *Given $f \in L^1$ and $f \geq 0$, and some $\lambda > 0$, there exists $\{Q_j\}$ dyadic cubes such that*
>   1. *$f(x) \leq \lambda$, almost everywhere for $x \notin \bigsqcup_j Q_j$*
>   2. *$\left| \bigsqcup_j Q_j \right| \leq \frac{1}{\lambda} \|f\|_{L^1}$*
>   3. *For all j, also have*
> $$\lambda < \frac{1}{|Q_j|} \int_{Q_j} f \leq 2^n \lambda$$

Remark Note that (3) is good since this means average of f over any Q_j is bounded by λ and a constant times λ (great if you don't care about constants).

Proof (1) follows from Theorem 2.10(2), and the observation that if $x \notin \bigsqcup_j Q_j$, then

$$\int_Q |f| \leq \lambda |Q|$$

for all dyadic Q such that $x \in Q$. And (2) is the dyadic maximal function being weak (1,1), and (3) is noting that by definition of Ω_k, we have

$$\int_{Q_j} \leq \int_{Q_{j-1}} f \leq \lambda |Q_{j-1}|$$

And $|Q_{j-1}|/|Q_j| = 2^n$. Hence we are done.

$\square$

2.5 Lecture 14

We now prove that the Hardy-Littlewood maximal funtion is weak (1,1) for higher dimensions. (We proved it for $n = 1$.)

Proof sketch: We will use the dyadic function being weak (1,1) to relate the Hardy-Littlewood maximal function $M'f(x)$ super level-set with that of $\{x : M_d f(x) > \lambda\}$.

> **Lemma 2.2**
>
> *If $f \geq 0$, then*
> $$|\{x \in \mathbb{R}^n : M'f(x) > 4^n \lambda\} \leq 2^n |\{x \in \mathbb{R}^n : M_d f(x) > \lambda\}$$

Proof Note for x such that $M'f(x)$ is large, then we should have $M_d f(x)$ should also be large (in a cube near it), and for this we utilize how dyadic cubes are aligned. (x must be close to other dyadic cubes such that the average value is large.)

We dissect a level set into dyadic cubes.

$$\{x \in \mathbb{R}^n : M_d f(x) > \lambda\} = \bigsqcup_j Q_j$$

where Q_j are dyadic cubes. Now it suffices to show that

$$\{x \in \mathbb{R}^n : M'f(x) > 4^n \lambda\} \subset \cup_j (2Q_j)$$

Since this would imply that

$$|\{x \in \mathbb{R}^n : M'f(x) > 4^n \lambda\}| \leq \sum_j |2Q_j| = \sum 2^n |Q_j| = 2^n |\{x : M_d f(x) > \lambda\}$$

Note that we do not quite get the level set is inside of $\bigsqcup_j Q_j$, but instead a slightly "swollen " union of $\cup_j (2Q_j)$. For every x such that $x \notin \cup_j (2Q_j)$. Let Q be any cube centered at x. Then Q could be "disected" into different regions (see figure 1), where each region is contained in a dyadic cube Q_j. Assuming the edge length l of Q is

$$l(Q) \in [2^{k-1}, 2^k)$$

Consider all the dyadic cubes with length 2^k that intersect Q, and one could prove the number of dyadic cubes Q_k that intersect Q is $\leq 2^n$

Note *This gives rise to the 2^n factor.*

We denote the dyadic cubes of length 2^k that intersect Q as $R_1, ..., R_m$, then we note that $R_1, ..., R_m$ are not contained in any Q_j for any j, because otherwise $x \in \cup_j 2Q_j$. Hence we have

$$\frac{1}{|R_j|} \int_{R_j} f(y)dy \leq \lambda \tag{2.2}$$

This gives that

$$\frac{1}{|Q|} \int_Q f(y)dy \leq \frac{1}{|Q|} 2^n \sum_j \int_{Q \cap R_j} f(y)dy \leq \frac{1}{|Q|} \sum_j \int_{R_j} f(y)dy \leq 2^n \frac{|R_j|}{|Q|} \lambda \leq 4^n \lambda \tag{2.3}$$

This shows that we have $M'f(x) \leq 4^n \lambda$ for $x \notin \cup_j (2Q_j)$.

$\square$

> **Corollary 2.4 (Lebesgue differentiation theorem)**
>
> *For $f \in L^1_{loc}(\mathbb{R}^n)$, we have*
> $$\lim_{r \to 0^+} \frac{1}{|B_r|} \int_{B_r} f(x-y)dy = f(x), a.e.$$
> $\heartsuit$

Proof We first note that this is a local property, hence we can replace L^1 with L^1_{loc}. We have that the Hardy-Littlewood function is of type weak (1,1). In fact, we have the stronger limit:

$$\lim_{r \to 0^+} \frac{1}{|B_r|} \int_{B_r} |f(x-y) - f(x)|dy = 0, a.e.$$

This follows from $\frac{1}{|B_r|} \int_{B_r} |f(x-y) - f(y)| \leq Mf(x) + |f(x)|$.

> **Definition 2.12 (Lebesgue point)**
>
> *If x is such that*
> $$\lim_{r \to 0^+} \frac{1}{|B_r|} \int_{B_r} |f(x-y) - f(x)|dy = 0$$
> $\clubsuit$

Note *Hence the Lebesgue differentiation theorem can also be stated as, for $f \in L^1_{loc}$, almost every x is a Lebesgue point for f.*

> **Proposition 2.4**
>
> *If x is a Lebesgue point of f, and $\{B_j\}$ is a decreasing sequence of balls such that $x \in B_j$ for all j, and $\cap_j B_j = x$, then we have*
> $$\lim_{j \to \infty} \int_{B_j} f dx = f$$
>
> *The same would hold if we replace B_j with cubes.*
> $\spadesuit$

Remark prove this We replace B_j with something smaller, with a fixed constant of 2^n.

Problem 2.1 How far are M, M' from strong (1,1)?

We now address this by noting that weak (1,1) is best we can do, especially when we exmaine the behavior at infinity.

> **Proposition 2.5**
>
> *If $f \in L^1, f \neq 0$, then $Mf \notin L^1$.*
> $\spadesuit$

Proof Without loss of generality, $f \neq 0$ on B, the unit ball. Then we have

$$Mf(x) \geq C \frac{1}{|x|^n} \int_B |f|dx$$

The third term $\int_B |f|dx$ is a constant, but the term $\frac{1}{|x|^n}$ just fails to be integrable. Hence $Mf(x)$ is not integrable.

$\square$

Even though Mf is not in L^1, however, we coul obtain some bound over any bounded set B in $\mathbb{R}^n$.

Theorem 2.7

If B is bounded and measurable in $\mathbb{R}^n$, then we have the following bound

$$\int_B Mf(x)dx \le 2|B| + \int_{\mathbb{R}^n} f \log^+ |f| dx$$

where $\log^+ t = \max\{\log t, 0\}$

Proof Note that

$$\int_B Mf(x)dx \le 2\int_0^\infty |\{x \in B : Mf(x) > 2\lambda\}d\lambda$$

$$\le 2|B| + 2\int_1^\infty a_f(2\lambda)d\lambda$$

Let $f = f_1 + f_2$, where $f_1 = f \cdot \chi_{x:|f|>\lambda}$ Then we have

$$\int_1^\infty a_f(2\lambda)d\lambda \le \int_1^\infty a_{f_1}(\lambda)d\lambda \le \int_1^\infty \frac{1}{\lambda} \int_{\{x:|f_1(x)|>\lambda\}} |f|dxd\lambda$$

Then we apply Fubini to get

$$2|B| + C \int_{\mathbb{R}^n} \int_1^{|f(x)|} \frac{1}{\lambda}d\lambda|f|dx \tag{2.4}$$

$\square$

Chapter 3 Hilbert Transform

3.1 Lecture 15

We now begin our discussion of chapter 3.

3.1.1 Conjugate Poisson Kernel

We fix a poisson kernel, and form a holomorphic function by integrating

---
Definition 3.1 (Harmonic extension)

Given $f \in \mathcal{S}(\mathbb{R})$, we define its harmonic extension to $\mathbb{R}_+^2$ as $= \{(x, y) : y > 0\}$, i.e.the upper half plane as

$$u(x, t) = P_t * f(x) \tag{3.1}$$

where P_t is the Poisson kernel.

♣

Our $P_t = \frac{1}{\pi} \frac{t}{t^2 + x^2}$, and if $z = x + it$, then

$$u(z) = \int_0^\infty \hat{f}(\xi) e^{2\pi i z \cdot \xi} + \int_\infty^0 \hat{f}(\xi) e^{2\pi i \overline{z} \cdot \xi} d\xi \tag{3.2}$$

Note *You can view this harmonic extension as a sum of a harmonic function and an anti-harmonic function.*

Note that if we define

$$iv(z) = \int_0^\infty \hat{f}(\xi) e^{2\pi i z \cdot \xi} - \int_{-\infty}^0 \hat{f}(\xi) e^{2\pi i \overline{z} \cdot \xi} d\xi$$

Exercise 3.1 If f is real, then one can show that u, v are real.

Note that by construction, v is also harmonic, and $u + iv$ is hence analytic. This gives that v could be a good candidate for the harmonic conjugate for u.

Definition 3.2 (Harmonic conjugate)

We define $v(z)$ as follows:

$$v(z) = \int_{\mathbb{R}} -i sgn(\xi) e^{2\pi t |\xi|} \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi$$

Here $v(x, t) = Q_t * f(x)$, where

$$\hat{Q}_t(\xi) = -i sgn(\xi) e^{-2\pi t |\xi|}$$

♣

Exercise 3.2 Show that

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

We define Q_t to be the conjugate Poisson kernel, where Q_t is harmonic in $\mathbb{R}_+^2$, conjugate to P_t, and we also have

$$P_t(x) + iQ_x(x) = \frac{1}{\pi} \cdot \frac{i}{z}$$

which is analytic on $\mathbb{R}_+^2$, and is the approximation of identity as $t \to 0$.

$$u(x, t) \to u(x), \text{ almost everywhere}$$

Note Q_t *is not an approximation to identity, and* $Q_t \notin L^1$.

Formally, $Q_t \to \frac{1}{\pi x}$ which is not locally integrble, but is defined as a tempered distribution by looking at its limiting values. (principle values), **Remark** This gives rise to the study of the Hilbert transform.

3.1.2 Principal value of $\frac{1}{x}$

> **Definition 3.3**
>
> *The principal value of $\frac{1}{x}$ is a tempered distribution, as is define as follows: for $\phi \in \mathcal{S}$,*
> $$p.v. \left(\frac{1}{x} \right)(\phi) = \lim_{\epsilon \to 0^+} \int_{|x| > \epsilon} \frac{\phi(x)}{x}$$
> *The difference of ϕ around 0 will cancel out x as $\epsilon \to 0$.* ♣

There is an alternative, which replaces the ϵ around the origin with the value $\phi(0)$, which is as follows: for $\phi \in \mathcal{S}$, we have

$$p.v. \left(\frac{1}{x} \right)(\phi) = \int_{|x| < 1} \frac{\phi(x) - \phi(0)}{x} dx + \int_{|x| > 1} \frac{\phi(x)}{x} dx$$

We could express $\phi(x) - \phi(0)$ as some constant C times x, hence the left integral is well-defined. And p.v. is therefore a well-defined tempered distribtuion because it continuously depends on ϕ.

> **Proposition 3.1**
>
> Q_t converges to $\frac{1}{\pi} p.v. \frac{1}{x}$ in $\mathcal{S}'$. ♠

Proof Let $\psi_\epsilon = \frac{1}{x} \psi_{|x| > \epsilon}$, and we have

$$\lim_{\epsilon \to 0} \psi_\epsilon = p.v. \frac{1}{x}$$

Hence now we show that $\lim_{\epsilon \to 0} \psi_\epsilon - \pi Q_t$ is 0. This interpreted as a tempered distribtuion is as follows: for all $\phi \in \mathcal{S}'$, we have

$$\int (\pi Q_t - \psi_\epsilon) \phi \to 0$$

Hence we get

$$
\begin{aligned}
(\pi Q_t - \psi_\epsilon, \phi) &= \int_{\mathbb{R}} \frac{\phi(x)}{t^2 + x^2} dx - \int_{|x| > \epsilon} \frac{\phi(x)}{x} dx \\
&= \int_{|x| < t} \frac{x \phi(x)}{t^2 + x^2} dx - \int_{|x| > t} \frac{t^2 \phi(x)}{x(t^2 + x^2)} dx \\
&= \int_{|x| < 1} \frac{x \phi(tx)}{1 + x^2} dx - \int_{|x| > 1} \frac{\phi(tx)}{x(1 + x^2)} dx
\end{aligned}
$$

3.2 Lecture 16

> **Proposition 3.2**
>
> Q_t converges to $\frac{1}{\pi} p.v. \frac{1}{x}$ for $t \to 0^+$. ♠

Proof We showed last time that

$$(\pi Q_t - \psi_\epsilon, \phi) = \int_{|x| < 1} \frac{x \phi(tx)}{1 + x^2} dx - \int_{|x| > 1} \frac{\phi(tx)}{x(1 + x^2)} dx$$

The limit is 0 by dominated convergence theorem and that $\frac{x}{1 + x^2}, \frac{1}{x(1 + x^2)}$ are odd.

□

By the weak continuity of FT, we have

$$\left(\frac{1}{\pi} p.v. \frac{1}{x} \right)^{\widehat{}} (\xi) = -i sgn(\xi)$$

Now we formally define the Hilbert transform.

Definition 3.4 (Hilbert transform)

The Hilbert transform H is defined as follows: given $f \in \mathcal{S}$,

$$Hf = \lim_{t \to 0} Q_t * f$$

$$Hf = \frac{1}{\pi} p.v. \frac{1}{x} * f$$

or equivalently

$$(Hf)\hat{}(\xi) = -i sgn(\xi) \hat{f}(\xi) \in L^2$$

Exercise 3.3 Prove the equivalence the above three definitions as an exercise.

Remark Recall that $v(x_i t) = Q_t * f = P_t * Hf$. This establishes a direct usage of the Hilbert transform.

Definition 3.5

For $f \in L^2$, we define the Fourier transform of Hf as follows;

$$\widehat{Hf} = -i sgn(\xi) \cdot \hat{f}$$

Proposition 3.3

H is an L^2-isometry, and we have $H(Hf) = -f$. Moreover, we have

$$\int (Hf) \cdot g = -f \cdot Hg$$

Proof We will only prove the last inequality.

$$\int Hf \cdot g = \langle Hf, \overline{g} \rangle = \langle -isgn(\xi) \cdot \hat{f}, \hat{\tilde{f}} \rangle = \langle \hat{f}, isgn(\xi) \cdot \tilde{f} \rangle = \langle f, -\overline{Hg} \rangle$$

$\square$

Remark You could find some applications in the section 6.3 in chapter 3.

We now begin the theorems of Riesz and Kolmogorov.

3.2.1 3.3

Note for any theorems regarding the Fourier transform of tempered distributions, it suffcies to prove it for the Schwartz functions.

Theorem 3.1

For $f \in \mathcal{S}(\mathbb{R})$, we have the following two inequalities.
  1. *(Kolmogorov) H is weak (1,1):*

$$|\{x \in \mathbb{R} : |Hf(x)| > \lambda\}| \lesssim \frac{1}{\lambda} \|f\|_{L^1}$$

  2. *(Riesz) H is strong (p, p), and $1 < p < \infty$, we have*

$$\|Hf\|_{L^p} \lesssim_p \|f\|_{L^p}$$

Note *H is not strong type (1,1), however, it is not far from it either. If we modify H somehow, we could show it as a bounded operator from $L^1 \to L^1$.*

Proof We will first show that H is weak (1,1). We would like to apply Calderon-Zygmund decomposition. We first assume $f \geq 0$, for all $\lambda > 0$, we have the following C-Z decomposition results: there exists disjoint intervals I_j such that

$$f(x) \leq \lambda, \text{ for a.e. } x, \text{if } x \notin \bigsqcup_j I_j := \Omega$$

And We have

$$|\Omega| \leq \frac{1}{\lambda} \|f\|_{L^1}$$

Moreover, we have

$$\lambda \leq \int_{I_j} f \leq 2\lambda$$

Let $f = g + b$, where

$$g(x) = \begin{cases} f(x), x \notin \Omega \\ \frac{1}{|I_j|} \int_{I_j} f, x \in I_j, b(x) = \sum_j b_j(x) \end{cases}$$

Intuitively, Ω is the disjoint union of all the bad intervals. For the Hilbert transform of b is not quite defined, hence we for now take it to be the $f - g$ and take the Hilbert transform.

Note that we have

$$b_j = (f(x) - \frac{1}{|I_j|} \int_{I_j} f) \chi_{I_j}(x)$$

And we have $Hb = Hf - Hg$, and $g \in L^2$ since we have $0 \geq g \leq 2\lambda$.

Then we split the contribution of g, h:

$$|\{x : |Hf(x) > \lambda\}| \leq |\{x : |Hg(x) > \frac{\lambda}{2}\}| + |\{x : |Hb(x) > \frac{\lambda}{2}\}|$$

We use I to denote the first term in the sum. Key: we use the L^2 boundedness of the Hilbert transform given $g \in L^2$.

$$I \lesssim \frac{1}{\lambda^2} \|Hg\|_{L^2}^2 = \frac{1}{\lambda^2} \|g\|_{L^2}^2 = \frac{1}{\lambda^2} \int g 2\lambda dx \leq \frac{2}{\lambda} \|g\|_{L^1} \leq \frac{2}{\lambda} \|f\|_{L^1}$$

We will deal with the contribution of b. We can protect the points by enlarging the intervals by twice the length. Define

$$\Omega^* = \cup_j 2I_j$$

We only look at $Hb(x)$ for $x \notin \Omega^*$ since

$$|\Omega|^* \leq 2|\Omega| \leq \frac{2}{\lambda} \|f\|_{L^1}$$

Note that we have

$$|Hb(x)| \leq \sum_j |Hb_j(x)|$$

Suffices to know that

$$\sum_j \int_{\mathbb{R} \setminus 2I_j} |Hb_j(x)| dx \leq c\|f\|_{L^1}$$

3.3 Lecture 17

We continue with the proof that the Hilbert transform is weak (1,1). To deal with $\frac{1}{x}$ at ∞, we use $\int_{I_j} b_j = 0$ and use the smoothness of $\frac{1}{x}$

It suffces to show that

$$\int_{\mathbb{R} \setminus \Omega^*} |Hb(x)| \lesssim \|b_j\|_{L^1}$$

Since we can sum over all j, note that

$$\sum_j \|b_j\|_{L^1} \leq \sum_j (\|f\|_{L^1(I_j)} + \|g\|_{L^1(I_j)}) = 2\|f\|_{L^1}$$

Note that there exists formula for the Hilbert transform:

$$Hb_j(x) = \frac{1}{\pi} \int_{I_j} \frac{b_j(y)}{x - y} dy, x \notin 2I_j$$

For this one, we note that we can fix y and use the smoothness of $\frac{1}{x}$.

Let c_j be the center of I_j, then we have

$$
\begin{aligned}
\int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx &= \frac{1}{\pi}\int_{\mathbb{R}\setminus 2I_j} \left| \int_{I_j} \frac{b_j}{x-y}dydx \right. \\
&= \frac{1}{\pi}\int_{\mathbb{R}\setminus 2I_j} \left| \int_{I_j} \frac{b_j}{x-y} - \frac{c_j}{x-y} \right| dydx \\
&\lesssim \int_{\mathbb{R}\setminus 2I_j} \int_I |b_j(y)| \cdot \frac{|y-c_j|}{|x-y||x-c_j|}dydx \\
&\lesssim \int_{\mathbb{R}\setminus 2I_j} \int_I |b_j(y)| \frac{|I_j|}{|x-c_j|^2} \\
&= \int_I |b_j(y)| \left(\int_{\mathbb{R}\setminus 2I_j} \frac{|I_j|}{|x-c_j|^2}dx \right) dy \\
&\lesssim \int_I |b_j(y)|dy = \|b_j\|_{L^1}
\end{aligned}
$$

(using the fact that $\int_{I_j} b_j = 0$.) Hence we are done!. □

🏵 **Note** *The same proof works for $f \in L^1 \cap L^2$.*

One can use the fact that the Hilbert transform is weak (1,1) to defeine the Hilbert transform for all $f \in L^1$. We take a sequence $f_n \xrightarrow{L^1} f, f_n \in \mathcal{S}$, and we then have $\{Hf_n\}$ is a cauchy sequence in measure, thus converges to a funciton Hf.

Now we show that H is strong (p,p), and for that we duality. We already know that H is strong (p,p) by interpolation theorems for $1 < p < 2$, since we have $p = 2$ is bounded.

Now we extend to ∞ using duality. For $p > 2$, for $\|g\|_{p'} = 1$,

$$
\begin{aligned}
\|Hf\|_{L^p} &= \sup_{g \in \mathcal{S}} \left| \int_{\mathbb{R}} Hf \cdot g \right| \\
&= \sup_{g \in \mathcal{S}} \left| \int_{\mathbb{R}} f \cdot Hg \right| \\
&\leq C_{p'} \|f\|_{L^p} \|g\|_{L^{p'}}
\end{aligned}
$$

where the last line, we used Holder's inequality, and using the fact that H is bounded for $1 < p < 2$.

🏵 **Note** *If you know an operator is bounded for $1 < p < 2$, and the middle line holds where you can exchange the transformation onto the other one, you immediately get that the operator is bounded for all others using duality.*

Proposition 3.4

The Hilbert transform is not strong (1,1). ♠

Proof Take $\chi_{[0,1]}$, and note that

$$
H\chi_{[0,1]}(x) = \frac{1}{\pi} \log \left| \frac{x}{x-1} \right|
$$

Proposition 3.5

If $\phi \in \mathcal{S}$, then $H\phi \in L^1$ if and only if $\int \phi = 0$. ♠

3.3.1 Truncated integrals and pointwise convergence

We define

$$
H_\epsilon f(x) = \frac{1}{\pi} \int_{|y|>\epsilon} \frac{f(x-y)}{y}dy
$$

This is well defined if $f \in L^p$, for $1 \leq p < \infty$, and since the kernel in $L^q, q > 1$.

3.4 Lecture 18

3.4.1 Truncated integrals

If we define

$$H_\epsilon f(x) = \int_{|y|>\epsilon} \frac{f(x-y)}{y} dy$$

can be directly defined in L^p, for $1 \le p < \infty$, since the kernel is in $L^{1+\epsilon}$.

Note that we clearly have a convolution, and we would like to find out what it is multiplied by on the Fourier side.

$$\frac{1}{y}\widehat{\chi_{|y|<\epsilon}}(\xi) = \lim_{N\to\infty}\int_{\epsilon<|y|<N}$$

$$= -i\lim_{N\to\infty}\int_{\epsilon<|y|<N}\frac{\sin(2\pi\xi)}{y}dy$$

$$= -isgn(\xi)\int_{N\to\infty}\int_{2\pi\epsilon|xi|}^{2\pi N|\xi|}\frac{\sin(t)}{t}dt$$

We note that this is uniformly bounded, and hence is strong (2,2).

Note *We can show in the same way as the weak (1,1) holds uniformly.*

This implies that for $f \in L^p$, we have $H_\epsilon f \to Hf$ as $\epsilon \to 0^+$, in L^p for $1 < p < \infty$, and in measure if $p = 1$.

Exercise 3.4 Approximate f by $\{f_n\} \subset \mathcal{S}$ and use uniformity.

For almost everywhere convergence, we'll prove that the following.

Theorem 3.2

Given $f \in L^p$, $1 \le p < \infty$, we have

$$Hf_\epsilon(x) \to Hf(x), a.e.\ x \in \mathbb{R}$$

as $\epsilon \to 0^+$. ♡

By theorem 2.2, it suffices to show that it holds for

$$H^*f(x) = \sup_{\epsilon.0}|H_\epsilon f(x)|$$

is weak (p,p), and for $f \in \mathcal{S}$. For this we show the following:

Theorem 3.3

H^* *is strong (p,p) and weak $(1,1)$, for $1 < p < \infty$.* ♡

Proof For $1 < p < \infty$, the strong (p,p) property follows from

---

Lemma 3.1 (Cotlar's inequality)

If $f \in \mathcal{S}$, we have

$$H^*f(x) \le M(Hf)(x) + CMf(x)$$ ♡

This gives H^* strong (p,p) for all $1 < p < \infty$. However, it fails to give weak (1,1) bound, since we don't know $M(Hf)$ behaves. (The proof of the Cotlar's inequality is by comparing $H_\epsilon f$ to $Q_\epsilon * f$).

We will prove this inequality with H^* replaced by H_ϵ with an absolute constant on the RHS. Fix $\phi \in \mathcal{S}(\mathbb{R})$, non-negative, and even, decaying with $supp(\phi) \subset [-1/2, 1/2]$ and $\int \phi = 1$. Then if we let $\phi_\epsilon(x) = \frac{1}{\epsilon}\phi\left(\frac{x}{\epsilon}\right)$

$$\frac{1}{y}\chi_{|y|>\epsilon} = (\phi_\epsilon * p.v.\frac{1}{x}) + \left(\frac{1}{y}\chi_{|y|>\epsilon} - \phi_\epsilon * p.v.\frac{1}{x}\right)$$

Convolutoin with the first term is bounded by M(Hf)(x), and we bound the kernel by an integrable, decaying function as follows:

It suffices to do this for $\epsilon = 1$, and for $|y| < 1$, wehave

$$\left|\phi_1 * p.v. \frac{1}{x}(y)\right| = \left|\lim_{\delta \to 0} \int_{|x|>\delta} \frac{\phi(y-x)}{x} dx\right| \leq \|\phi\|_{C^1}$$

And for $|y| \geq 1$, we have

$$\left|\frac{1}{y} - \int_{|x|<1/2} \frac{\phi(x)}{y-x} dx\right| = \left|\int_{|x|<1/2} \phi(x)\left(\frac{1}{y} - \frac{1}{y-x}\right)\right| \leq \frac{C}{y^2}$$

Thus the contribution from the 2nd term is $\leq CMf(x)$.

$\square$

Now we show that lemma 3.5 implies theorem 3.4 for strong (p,p). Now we give the proof of weak $(1,1)$ of theorem 3.4, WLOG, assume $f \geq 0$, and fix $\lambda > 0$, C-Z decomposition at λ gives $f = g + \sum_j b_j$.

Now for exercise, since H^* is strong (2,2), and g can be handled in the old way.

Note *We will achieve this also in almost the old way with one additional help of the maximal function, And note that the average of b_j might no longer tend to 0. It suffices to show that*

$$|\{x \notin \Omega^* : H^*b(x) > \lambda\}| \leq \frac{C}{\lambda}\|b\|_{L^1}$$