Lecture Notes in Harmonic Analysis

Lectures by Dr. Charles Moore

Throughout this document, □ signifies end proof, and ▲ signifies end of example.

Table of Contents

Notes by Jakob Streipel. Last updated December 1, 2017.

Lecture 1 Introduction to Fourier Analysis

Harmonic analysis is a broad field involving a great deal of subjects concerning the art of decomposing functions into constituent parts. These might be Fourier coefficients, breaking them down into exponential parts, wavelet theory, tools to deal with partial differential equations, or Sobolev spaces.

This course will deal with the following:

- Fourier analysis,
- Harmonic functions,
- Singular integrals, and
- Maximal functions.

1.1 Fourier Analysis

Definition 1.1.1 (Inner product space)**.** Let V be a finite dimensional vector space over $\mathbb{C}$. Then V is called an ***inner product space*** if there is a mapping $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ which satisfies the following for all vectors $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ and scalars $\alpha \in \mathbb{C}$:

(i) Conjugate symmetry, meaning that $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$;

(ii) Linearity in the first argument, i.e. $\langle \alpha\boldsymbol{x}, \boldsymbol{y} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ and $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$;

(iii) Positive-definiteness, meaning that $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$.

All inner product spaces automatically induce a ***norm***, namely $\|\boldsymbol{v}\| = \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{1/2}$.

Moreover normed spaces are automatically metric spaces by $d(\boldsymbol{v}, \boldsymbol{w}) = \|\boldsymbol{v} - \boldsymbol{w}\|$. Ergo it has the following properties:

Definition 1.1.2 (Metric)**.** A function $d(\cdot, \cdot) \colon V \times V \to \mathbb{R}$ is called a ***metric*** if

(i) $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$;

(ii) symmetry, i.e. $d(x, y) = d(y, x)$; and

(iii) $d(x, z) \leq d(x, y) + d(y, z)$

for all $x, y, x \in V$.

Since we have an inner product, we are able to define all manner of other interesting concepts.

Definition 1.1.3 (Orthogonal, orthonormal)**.** A basis $B = \{ \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \}$ is called ***orthogonal*** if the basis elements are pairwise orthogonal, i.e. $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$ for all $i \neq j$. Moreover the basis is called ***orthonormal*** if in addition $\|\boldsymbol{v}_i\| = 1$ for all i.

Since a basis spans the space, we can write $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$. Now if the basis in addition is orthonormal we have the illuminating property that

$$\begin{aligned}\langle \boldsymbol{v}, \boldsymbol{v}_1\rangle &= \langle c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n, \boldsymbol{v}_1\rangle \\ &= c_1\langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle + c_2\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle + \ldots + c_n\langle \boldsymbol{v}_1, \boldsymbol{v}_n\rangle \\ &= c_1\|\boldsymbol{v}_1\|^2 + 0 + 0 + \ldots + 0 = c_1,\end{aligned}$$

and similarly $\langle \boldsymbol{v}, \boldsymbol{v}_i\rangle = c_i$ for all i.

Therefore

$$\boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{v}_1\rangle\boldsymbol{v}_1 + \langle \boldsymbol{v}, \boldsymbol{v}_2\rangle\boldsymbol{v}_2 + \ldots + \langle \boldsymbol{v}, \boldsymbol{v}_n\rangle\boldsymbol{v}_n,$$

and $\langle vvv, \boldsymbol{v}_i\rangle$ are called the ***Fourier coefficients*** of $\boldsymbol{v}$.

We can do the same thing on a slightly more interesting space than ordinary Euclidean space, namely L^1:

Definition 1.1.4. Suppose f is a function on $[-\pi, \pi)$ with

$$\int_{-\pi} \pi|f(x)|\,dx < \infty$$

meaning that $f \in L^1[-\pi, \pi)$. Then for $n \in \mathbb{Z}$ we define

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx,$$

called the nth ***Fourier coefficient*** of f.

On $[-\pi, \pi)$, consider the set $\{\, e^{inx} \,\}_{n\in\mathbb{Z}}$. Define

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx.$$

Exercise 1.1.5. Show that this is indeed an inner product.

Solution. First conjugate symmetry, and clearly

$$\begin{aligned}\langle f, g\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{\overline{f(x)\overline{g(x)}}}\,dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{\overline{f(x)}g(x)}\,dx = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)\overline{f(x)}\,dx} = \overline{\langle g, f\rangle}.\end{aligned}$$

Linearity follows by integration being linear on its own, and note for positive-definiteness that

$$\langle f, f\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{f(x)}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx \geq 0$$

since for any complex number $z = a + ib$,

$$z\overline{z} = (a + ib)(a - ib) = a^2 - i^2b^2 = a^2 + b^2 = |z|^2.$$

For the final condition, note that in fact they are implied by the first three for all inner products. To see this, take $f = 0$ in linearity, i.e.

$$\langle 0, 0\rangle = \langle g, 0\rangle + \langle -g, 0\rangle = \langle g, 0\rangle - \langle g, 0\rangle = 0$$

and from positive-definiteness we have that $\langle x, x\rangle = 0$ implies $x = 0$. ♦

Now suppose that $n \neq m$ are integers. Then

$$\begin{aligned}\langle e^{inx}, e^{imx}\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{e^{imx}}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-imx}\,dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)x}\,dx = \frac{1}{2\pi}\frac{e^{i(n-m)x}}{i(n-m)}\bigg|_{x=-\pi}^{\pi} \\ &= \frac{1}{2\pi}\Big(\frac{e^{i(n-m)\pi}}{i(n-m)} - \frac{e^{-i(n-m)\pi}}{i(n-m)}\Big) = 0,\end{aligned}$$

and

$$\|e^{inx}\| = \langle e^{inx}, e^{inx}\rangle = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{e^{inx}}\,dx\Big)^{1/2} = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{0}\Big)^{1/2} = 1.$$

Therefore $\{\, e^{inx} \,\}_{n\in\mathbb{Z}}$ is an orthonormal set.

Given a function f, we defined $\hat{f}(n) = \langle f, e^{inx}\rangle$. We would like for this to mean that

$$f = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$$

like

$$\boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{v}_1\rangle\boldsymbol{v}_1 + \langle \boldsymbol{v}, \boldsymbol{v}_2\rangle\boldsymbol{v}_2 + \ldots + \langle \boldsymbol{v}, \boldsymbol{v}_n\rangle\boldsymbol{v}_n.$$

But is it a basis? In what sense does this infinite sum converge?

1.2 In more general settings...

Given f on $[-\pi,\pi)$ and $\int_{-\pi}^{\pi}|f(x)|\,dx < \infty$ we defined $\hat{f}(n) = \langle f, e^{inx}\rangle$.

Then $\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$ is called the ***Fourier series*** of f, often written

$$f \sim \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}.$$

Now it is extremely important to note that this means *only* that $\hat{f}(n)$ is defined as the integral we discussed previously, and *nothing more*.

In general, if μ is a measure on $[-\pi,\pi)$, then $\hat{\mu}(n) = \int_{-\pi}^{\pi} e^{inx}\,d\mu(x)$, and indeed if f is a function than $f(x)\,dx$ is an example of a measure.

We can do it more generally in higher dimensions. If $f\colon \mathbb{R}^n \to \mathbb{C}$, and if $\int_{\mathbb{R}^n}|f(\boldsymbol{x})|\,d\boldsymbol{x} < \infty$, we define

$$\hat{f}(x) = \int_{\mathbb{R}^n} e^{i\boldsymbol{x}\cdot\boldsymbol{t}}f(\boldsymbol{t})\,d\boldsymbol{t},$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$, and $\boldsymbol{x}\cdot\boldsymbol{t} = x_1t_1 + x_2t_2 + \ldots + x_nt_n$. This $\hat{f}\colon \mathbb{R}^n \to \mathbb{C}$ is called the ***Fourier transform***.

Remark 1.2.1. This is of course not unlike, say, the Laplace transform,

$$\mathcal{L}\{f\}(s) = \int_0^{\infty} f(t)e^{-st}\,dt.$$

Indeed we could consider n-dimensional tori too, if we like, where'd we'd be doing the exact same thing, just over a different integrating domain (and with a different scalar in front).

More generally we may do this over any locally compact Abelian group G (which is what it sounds like, with the caveat that the group structure and the topological structure be connected in that the group operations of addition and negation are both continuous).

It is a fact that on G there exists a measure μ called the Haar measure such that it is translation invariant, i.e. $\mu(E + x) = \mu(E)$, for evert Borel set $E \subseteq G$ and $x \in G$.

If $\gamma\colon G \to \mathbb{C}$ is a mapping such that $\gamma(x + y) = \gamma(x)\gamma(y)$ for every $x, y \in G$, e.g. $\gamma(x) = e^{inx}$, then γ is called a (multiplicative) character on G.

In $L^2[-\pi, \pi)$, e^{inx} are characters. Moreover the set of all characters of a group is called the dual group, often denoted Γ.

Now we can define

$$\hat{f}(\gamma) = \int_G f(x)\overline{\gamma(x)}\,dx.$$

(These things are discussed in Rudin's *Fourier analysis on LCA groups.*)

Lecture 2 More Fourier Analysis

Recall from last time that we write

$$f \sim \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx}$$

to mean that

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt,$$

for functions $f \in F^1([-\pi, \pi))$, which means that

$$\int_{-\pi}^{\pi} |f(t)|\,dt.$$

Similarly if f is defined on $\mathbb{R}^d$ we write

$$\hat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} e^{-i\boldsymbol{\xi}\cdot\boldsymbol{t}}\,d\boldsymbol{t}.$$

2.1 Elementary Facts from Fourier Analysis

Proposition 2.1.1. *With these definitions,*

(i) $|\hat{f}(n)| \leq \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} |f(t)|\,dt$, *and*

(ii) $|\hat{f}(\boldsymbol{\xi})| \leq \displaystyle\int_{\mathbb{R}^d} |f(\boldsymbol{t})|\,d\boldsymbol{t}$.

Proof. We prove the first one; the second one is almost identical:

$$\begin{aligned}|\hat{f}(n)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt\right| \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t)||e^{-int}|\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t)|\,dt.\end{aligned}$$ □

We can do better, in fact.

Proposition 2.1.2 (Riemann-Lebesgue lemma). *(i) If $f \in L^1([-\pi,\pi))$ then $|\hat{f}(n)| \to 0$ as $|n| \to \infty$, and*

(ii) If $f \in L^1(\mathbb{R}^d)$ then $\hat{f}(\boldsymbol{\xi})$ is continuous and $|\hat{f}(\boldsymbol{\xi})| \to 0$ as $|\boldsymbol{\xi}| \to \infty$.

Note that whilst (*ii*) might seem stronger, since (*a*) is a sequence it is automatically continuous in the discrete topology.

Proof. We prove the first one (since again, save for the continuity, the second one is similar), but we prove it only for indicator functions. This is of course fine since they are dense in the space of L^1 functions.

For an indicator function

$$\chi_{(a,b)}(x) = \begin{cases} 1, & \text{if } x \in (a,b) \\ 0, & \text{otherwise}\end{cases}$$

we have

$$\begin{aligned}\hat{\chi}_{(a,b)}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\chi_{(a,b)}e^{-inx}\,dx = \frac{1}{2\pi}\int_a^b e^{-inx}\,dx \\ &= \frac{1}{2\pi}\frac{e^{-inx}}{-in}\bigg|_a^b = \frac{1}{2\pi}\Big(\frac{e^{-inb}-e^{-ina}}{-in}\Big) \to 0\end{aligned}$$

as $n \to \infty$ since e^{ix} is bounded by 1.

Now since characteristic functions are dense in L^1, this means that there exists a function g that is a linear combination of characteristic functions such that

$$\int_{-\pi}^{\pi}|f(x)-g(x)|\,dx < \varepsilon$$

for any $\varepsilon > 0$. Now since $\hat{g}(n) \to 0$ as $|n| \to \infty$ there exists an M such that for all $|n| > M$ we have $|\hat{g}(n)| < \varepsilon$. Take such an n, then

$$\begin{aligned}|\hat{f}(n)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt\right| \\ &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt + \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-int}\,dt - \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-int}\,dt\right| \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t)-g(t)||e^{-int}|\,dt + |\hat{g}(n)| \leq \frac{\varepsilon}{2\pi} + \varepsilon\end{aligned}$$

which can of course be made arbitrarily small.

For the continuity in the second case, let $\boldsymbol{\xi} \in \mathbb{R}^d$ and morever let $\boldsymbol{\xi}_j \in \mathbb{R}^d$ such that $\boldsymbol{\xi}_j \to \boldsymbol{\xi}$. Then

$$\begin{aligned} |\hat{f}(\boldsymbol{\xi}_j) - \hat{f}(\boldsymbol{\xi})| &= \left| \int_{\mathbb{R}^d} f(\boldsymbol{t}) e^{-i\boldsymbol{\xi}_j \cdot \boldsymbol{t}} \, d\boldsymbol{t} - \int_{\mathbb{R}^d} f(\boldsymbol{t}) e^{-i\boldsymbol{\xi} \cdot \boldsymbol{t}} \, d\boldsymbol{t} \right| \\ &\leq \int_{\mathbb{R}^d} |f(\boldsymbol{t})| |e^{-i\boldsymbol{x} i_j \cdot \boldsymbol{t}} - e^{-i\boldsymbol{\xi} \cdot \boldsymbol{t}} \, d\boldsymbol{t}| \to 0 \end{aligned}$$

as $j \to \infty$. □

Remark 2.1.3. Note how without comment we took limits from outside of the integral to the inside of the integral just then. This is of course allowed since the integrand is finite, by previous discussion.

This is contrary to classical examples such as $f_n(x) = n\chi_{(0,1/n)}(x)$. Since this converges pointwise to 0 as n goes to infinity, we have

$$\int_0^1 \lim_{n\to\infty} f_n(x) \, dx = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \int_0^1 f_n(x) \, dx = 1.$$

Remark 2.1.4. The Riemann-Lebesgue lemma is not true for measures. Take for instance the Dirac measure δ, for which we have

$$\hat{\delta}(n) = \int_1 2\pi \int_{-\pi}^{\pi} e^{-int} \, d\delta(t) = 1$$

for all n, so it certainly does not go to 0.

Remark 2.1.5. We have now shown that the Fourier coefficients of an L^1 function f is a two-sided decaying sequence. One might ask, then, whether if for each decaying two-sided sequence we might find an L^1 function who has precisely that sequence as its Fourier coefficients.

The answer to this question is no, by Bochner's theorem, but more on this later.

Proposition 2.1.6. *Let $f \in L^1([-\pi, \pi))$. Extend f to all of $\mathbb{R}$ as a 2π periodic function, i.e. $f(2\pi n + x) = x$ for all $n \in \mathbb{Z}$. Then*

(*i*) *Let $y \in \mathbb{R}$ and define $g(x) = f(x - y)$. Then $\hat{g}(n) = \hat{f}(n)e^{-iyn}$;*

(*ii*) *If $m \in \mathbb{Z}$, $g(x) = f(x)e^{imx}$, then $\hat{g}(n) = \hat{f}(n - m)$;*

Proof. (*i*) We compute, making a change of variable along the way:

$$\begin{aligned} \hat{g}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) e^{-int} \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t - y) e^{-int} \, dt \\ &= \frac{1}{2\pi} \int_{-\pi-y}^{\pi-y} f(s) e^{-in(s+y)} \, ds = \frac{e^{-iny}}{2\pi} \int_{-\pi-y}^{\pi-y} f(s) e^{-ins} \, ds \\ &= e^{-iny} \hat{f}(n). \end{aligned}$$

(*ii*) is similar. □

Something similar is true for products of functions—almost.

Definition 2.1.7 (Convolution)**.** If $f \in L^p(\mathbb{R}^d)$, $g \in L^1(\mathbb{R}^d)$, then we let

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

called the ***convolution*** of f and g.

Not is turns out that $\widehat{fg} \neq \hat{f}\hat{g}$ in general, but $\widehat{f * g} = \hat{f} \cdot \hat{g}$.

Lecture 3 Convolving Functions

3.1 Properties of Convolution

Recall that for $f, g \in L^1([-\pi, \pi))$, we define

$$f * g(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-y)g(y)\,dy$$

and for $f, g \in L^1(\mathbb{R}^d)$ we define

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

Proposition 3.1.1 (Young's inequality)**.** *If $f \in L^p$ and $g \in L^1$, then $\|f*g\|_p \leq \|f\|_p \|g\|_1$.*

We will need the following lemma:

Lemma 3.1.2 (Holder's inequality)**.** *Let $1 \leq p, q \leq \infty$ be such that $1/p + 1/q = 1$. Then*

$$\int_{\mathbb{R}^d} |h(\boldsymbol{x})k(\boldsymbol{x})|\,d\boldsymbol{x} \leq \Big(\int_{\mathbb{R}^d} |h(\boldsymbol{x})|^p\,d\boldsymbol{x}\Big)^{1/p} \Big(\int_{\mathbb{R}^d} |k(\boldsymbol{x})|^q\,d\boldsymbol{x}\Big)^{1/q}.$$

Proof. We prove the case of $\mathbb{R}^d$; the other one is similar.

$$\begin{aligned}
\|f * g\|_p^p &= \int_{\mathbb{R}^d} |f * g(\boldsymbol{x})|^p\,d\boldsymbol{x} = \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} f(\boldsymbol{x}-\boldsymbol{y})g(\boldsymbol{y})\,d\boldsymbol{y} \right|^p d\boldsymbol{x} \\
&\leq \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} |f(\boldsymbol{x}-\boldsymbol{y})||g(\boldsymbol{y})|^{1/p}|g(\boldsymbol{y})|^{1/q}\,d\boldsymbol{y} \right|^p d\boldsymbol{x} \\
&\leq \int_{\mathbb{R}^d} \left| \Big(\int_{\mathbb{R}^d} |f(\boldsymbol{x}-\boldsymbol{y})|^p |g(\boldsymbol{y})|\,d\boldsymbol{y}\Big)^{1/p} \Big(\int_{\mathbb{R}^d} |g(\boldsymbol{y})|\,d\boldsymbol{y}\Big)^{1/q} \right|^p d\boldsymbol{x} \\
&= \|g\|_1^{p/q} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(\boldsymbol{x}-\boldsymbol{y})|^p |g(\boldsymbol{y})|\,d\boldsymbol{y}\,d\boldsymbol{x} \\
&= \|g\|_1^{p/q} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(\boldsymbol{x}-\boldsymbol{y})|^p |g(\boldsymbol{y})|\,d\boldsymbol{x}\,d\boldsymbol{y} \\
&= \|g\|_1^{p/q} \int_{\mathbb{R}^d} |g(\boldsymbol{y})| \int_{\mathbb{R}^d} |f(\boldsymbol{x}-\boldsymbol{y})|^p\,d\boldsymbol{x}\,d\boldsymbol{y} \\
&= \|g\|_1^{p/q} \|f\|_p^p \|g\|_1 = \|f\|_p^p \|g\|_1^{p/q+1} = \|f\|_p^p \|g\|_1^p.
\end{aligned}$$

□

Proposition 3.1.3. *(i) If $f, g \in L^1([-\pi,\pi))$ then*

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

(ii) If $f, g \in L^1(\mathbb{R}^d)$ then

$$\widehat{f * g}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi})\hat{g}(\boldsymbol{\xi}).$$

Proof. We prove (i):

$$\begin{aligned}
\widehat{f * g}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f * g(t) e^{-int}\, dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)g(s)\, ds\, e^{-int}\, dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)g(s)e^{-int}\, ds\, dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)g(s)e^{-int}\, dt\, ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)e^{-int}\, dt\, ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\int_{-\pi-s}^{\pi-s} f(v)e^{-in(s+v)}\, dv\, ds \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)e^{-ins}\frac{1}{2\pi}\int_{-\pi-s}^{\pi-s} f(v)e^{-inv}\, dv\, ds = \hat{g}(n)\hat{f}(n). \qquad \square
\end{aligned}$$

Proposition 3.1.4. *(i) If $f \in L^1(\mathbb{R}^d)$ and if $x_k f(\boldsymbol{x}) \in L^1(\mathbb{R}^d)$ (where $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$), then $\hat{f}(\boldsymbol{\xi})$ is differentiable almost everywhere with respect to ξ_k and*

$$\frac{\partial}{\partial \xi_k}\hat{f}(\boldsymbol{\xi}) = \widehat{-ix_k f}(\boldsymbol{\xi}).$$

(ii) If $f, \partial f/\partial x_k \in L^1(\mathbb{R}^d)$, then

$$\widehat{\frac{\partial f}{\partial x_k}}(\boldsymbol{\xi}) = i\xi_k \hat{f}(\boldsymbol{\xi}).$$

Proof. Set $\boldsymbol{h} = (0, \ldots, 0, h, 0, \ldots, 0)$, with h being in the kth position. Consider

$$\begin{aligned}
\frac{\hat{f}(\boldsymbol{\xi} - \boldsymbol{h}) - \hat{f}(\boldsymbol{\xi})}{h} &= \frac{1}{h}\int_{\mathbb{R}^d} f(\boldsymbol{t})(e^{-i(\boldsymbol{\xi}+\boldsymbol{h})\cdot \boldsymbol{t}} - e^{-i\boldsymbol{\xi}\cdot\boldsymbol{t}})\, d\boldsymbol{t} \\
&= \int_{\mathbb{R}^d} f(\boldsymbol{t})e^{-i\boldsymbol{\xi}\cdot\boldsymbol{t}}\frac{e^{-i\boldsymbol{h}\cdot\boldsymbol{t}} - 1}{h}\, d\boldsymbol{t}.
\end{aligned}$$

Thus taking limits we get the partial derivative in $\boldsymbol{\xi}$, whereby

$$\begin{aligned}
\frac{\partial \hat{f}}{\partial \xi_k}(\boldsymbol{\xi}) &= \int_{\mathbb{R}^d} f(\boldsymbol{t})e^{-i\boldsymbol{\xi}\cdot\boldsymbol{t}}\lim_{h\to 0}\frac{e^{-iht_k} - 1}{h}\, d\boldsymbol{t} \\
&= \int_{\mathbb{R}^d} f(t)e^{-i\boldsymbol{\xi}\cdot\boldsymbol{t}}(-it_k)\, d\boldsymbol{t} = \widehat{-it_k f}(\boldsymbol{\xi}). \qquad \square
\end{aligned}$$

Proposition 3.1.5. *If* $f, df/dx \in L^1([-\pi,\pi))$ *then*

$$\widehat{\frac{df}{dx}}(n) = in\hat{f}(n).$$

Proof. By simple computation using integration by parts,

$$\begin{aligned}\widehat{\frac{df}{dx}}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{df}{dx}(x)e^{-inx}\,dx \\ &= \frac{1}{2\pi}f(x)e^{-inx}\Big|_{-\pi}^{\pi} - \frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)(-in)e^{-inx}\,dx \\ &= in\int_{-\pi}^{\pi}f(x)e^{-inx}\,dx = in\hat{f}(n).\end{aligned}$$

□

Lecture 4 An Application

4.1 Photographing a Star

We will describe an application of the Fourier transform. Imagine taking a two-dimensional photograph of a start from Earth. This star looks like a disc from our perspective, and due to there being atmosphere between us and the star the picture will be blurry. A question one might ask oneself then is: despite the blurriness, would it be possible to determine the radius of the star?

To describe how this is accomplished using the Fourier transform, we first set the problem up in terms of functions on the plane.

The true image of the star can be described as

$$f(\boldsymbol{x}) = \lambda\chi_B\Big(\frac{\boldsymbol{x}-\boldsymbol{y}}{\varepsilon}\Big)$$

where $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{y} = (y_1, y_2)$ is the center of the disc (which is the image of the star), ε is the radius of the disc, and λ is the brightness or luminosity of the star. Moreover χ_B is the characteristic function of the unit disc, from which it follows that the above is the disc of radius ε centred on $\boldsymbol{y}$ with brightness λ.

The blurry photograph of the star can be described by

$$f * k(\boldsymbol{x}) = \int_{\mathbb{R}^2} f(\boldsymbol{y})k(\boldsymbol{x}-\boldsymbol{y})\,d\boldsymbol{y}$$

where $k(\boldsymbol{y})$ is some sort of smooth function with a bump around the origin and 0 elsewhere. Such a function k is called a ***mollifier***, which comes from the fact that the convolution above will produce an image of f, with the caveat that sharp edges have been smoothed out.

Now imagine us taking n photos, getting $f*k_1, f*k_2, \ldots, f*k_n$. We superimpose these, yielding $f*k_1 + f*k_2 + \ldots + f*k_n$, and then take the Fourier transform, producing

$$\hat{f}\hat{k}_1 + \hat{f}\hat{k}_2 + \ldots + \hat{f}\hat{k}_n = \hat{f}(\hat{k}_1 + \hat{k}_2 + \ldots + \hat{k}_n).$$

Now $\hat{f}$ has zeros, as do $\hat{k}_1, \ldots, \hat{k}_n$. Let us assume that the zeros of Fourier transform are the zeros of $\hat{f}$ (i.e. that the zeros of the Fourier transforms of the mollifier don't happen to coincide with each other and the zeros of $\hat{f}$). Then, using the substitutions $\boldsymbol{s} = \boldsymbol{x} - \boldsymbol{y}$ and $\boldsymbol{u} = \boldsymbol{y}/\varepsilon$ along the way,

$$\begin{aligned}
\hat{f}(\boldsymbol{\xi}) &= \int_{\mathbb{R}^2} e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}} \lambda \chi_B\Big(\frac{\boldsymbol{x}-\boldsymbol{y}}{\varepsilon}\Big)\, d\boldsymbol{x} \\
&= \lambda \int_{\mathbb{R}^2} e^{-i(\boldsymbol{s}+\boldsymbol{y})\cdot\boldsymbol{\xi}} \chi_B(\boldsymbol{s}/\varepsilon)\, d\boldsymbol{s} \\
&= \lambda e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}} \int_{\mathbb{R}^2} e^{-i\boldsymbol{s}\cdot\boldsymbol{\xi}} \chi_B(\boldsymbol{s}/\varepsilon)\, d\boldsymbol{y} \\
&= \varepsilon^2 \lambda e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}} \int_{\mathbb{R}^2} e^{-i\varepsilon\boldsymbol{u}\cdot\boldsymbol{\xi}} \chi_B(\boldsymbol{u})\, d\boldsymbol{u} \\
&= \varepsilon^2 \lambda e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}} \int_{\mathbb{R}^2} e^{-i\boldsymbol{u}\cdot(\varepsilon\boldsymbol{\xi})} \chi_B(\boldsymbol{u})\, d\boldsymbol{u} \\
&= \varepsilon^2 \lambda e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}} \widehat{\chi_B}(\varepsilon\boldsymbol{\xi})
\end{aligned}$$

It happens to be a fact that the Fourier transform of the unit ball is

$$\widehat{\chi_B}(\boldsymbol{\xi}) = \frac{2\pi J_1(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|},$$

where J_1 is the Bessel function of the first kind. Therefore

$$\hat{f}(\boldsymbol{\xi}) = \varepsilon\lambda e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}} 2\pi \frac{J_1(\varepsilon|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|},$$

so under our assumption the zeros of the superimposed image of the star should be the zeros of $\hat{f}$, which are then exactly characterised by the zeros of the Bessel function, which are well known. Knowing then that the distance between the observed zeros would be the distance between the zeros of the Bessel function scaled by the radius ε, we can recover this radius.

4.2 Results in L^2

A question we have asked ourselves previously is when a function $f \in L^1([-\pi, \pi))$ is equal to its own Fourier series. In other words, when is

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx}?$$

Recall first that $f \in L^2([-\pi, \pi))$ means that

$$\int_{-\pi}^{\pi} |f(t)|^2\, dt < \infty.$$

Proposition 4.2.1. *Suppose $e_1, e_2, \dots, e_N$ is an orthonormal set in an inner product space. Then*

$$\left\| \sum_{n=1}^{N} a_n e_n \right\|^2 = \sum_{n=1}^{N} |a_n|^2$$

for all $a_n \in \mathbb{C}$.

Proof. That this is true is quite obvious: expanding the left-hand sum we get mostly mixed $\langle e_n, e_m \rangle$, which by orthogonality are zero, and the remaining $\langle e_n, e_n \rangle$ are 1 by normality. □

Proposition 4.2.2. *Let f be a function in an inner product space, and $\{ e_1, e_2, \dots, e_N \}$ an orthonormal set. Then*

$$\left\| \sum_{n=1}^{N} \langle f, e_n \rangle - f \right\|^2 = \|f\|^2 - \sum_{n=1}^{N} |\langle f, e_n \rangle|^2.$$

Proof. By computation:

$$\begin{aligned}
\left\| \sum_{n=1}^{N} \langle f, e_n \rangle - f \right\|^2 &= \left\langle \sum_{n=1}^{N} \langle f, e_n \rangle e_n - f, \sum_{n=1}^{N} \langle f, e_n \rangle e_n - f \right\rangle \\
&= \left\langle \sum_{n=1}^{N} \langle f, e_n \rangle e_n, \sum_{n=1}^{N} \langle f, e_n \rangle e_n \right\rangle - \left\langle \sum_{n=1}^{N} \langle f, e_n \rangle e_n, f \right\rangle \\
&\quad - \left\langle f, \sum_{n=1}^{N} \langle f, e_n \rangle e_n \right\rangle + \langle f, f \rangle \\
&= \sum_{n=1}^{N} |\langle f, e_n \rangle|^2 - 2 \sum_{n=1}^{N} \langle f, e_n \rangle \overline{\langle f, e_n \rangle} + \|f\|^2 \\
&= \sum_{n=1}^{N} |\langle f, e_n \rangle|^2 - 2 \sum_{n=1}^{N} |\langle f, e_n \rangle|^2 + \|f\| \\
&= \|f\| - \sum_{n=1}^{N} |\langle f, e_n \rangle|^2.
\end{aligned}$$ □

Corollary 4.2.3 (Bessel's inequality)**.** *In the same setting,*

$$\sum_{n=1}^{N} |\langle f, e_n \rangle|^2 \leq \|f\|^2.$$

Proof. This is immediate byu the previous proposition since norms are nonnegative. □

Example 4.2.4. For any N,

$$\sum_{n=-N}^{N} |\hat{f}(n)|^2 \leq \|f\|$$

whereby

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 \leq \|f\|_2^2.$$

▲

Lecture 5 Hilbert Spaces

5.1 Fourier Series on L^2

Recall that $L^2([-\pi, \pi))$ is the set of all functions f satisfying

$$\|f\|_2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2 \, dt\right)^{1/2} < \infty$$

the norm of which is induced by the inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)} \, dt,$$

and the space, being normed, is automatically metric, by

$$d(f, g) = \|f - g\|_2.$$

Definition 5.1.1 (Hilbert space)**.** If H is an inner product space so that the resulting metric is complete, then H is called a ***Hilbert space***.

Recall moreover that a space being complete means that every Cauchy sequence converges in the space.

Definition 5.1.2 (Complete set)**.** An orthonormal set $\{ e_\alpha \} \subset H$, with H a Hilbert space, is called ***complete*** or ***maximal*** if $\langle f, e_\alpha\rangle = 0$ for every α implies that $f = 0$.

Example 5.1.3. Let $H = \mathbb{R}^3$, with $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$, as well as $f = (f_1, f_2, f_3)$. Then $\langle f, e_i\rangle = 0$ for $i = 1, 2, 3$ implies that $f = 0$. ▲

Proposition 5.1.4. *Let $\{ e_\alpha \}$ be an orthonormal set in a Hilbert space H. Then*

(*i*) $\|f\|_2^2 = \sum_\alpha |\langle f, e_\alpha\rangle|^2$ *(Bessel's equality).*

(*ii*) $f = \sum_\alpha \langle f, e_\alpha\rangle e_\alpha$ *in H, by which we mean that*

$$\left\| f - \sum_{\alpha\in S} \langle f, e_\alpha\rangle e_\alpha \right\|$$

can be made arbitrarily small by choosing an appropriate finite set S of indices.

(*iii*) *If in addition $g \in H$, then*

$$\langle f, g\rangle = \sum_\alpha \hat{f}(\alpha)\overline{\hat{g}(\alpha)}$$

where $\hat{f}(\alpha) = \langle f, e_\alpha\rangle$, and again we mean convergence in the above sense.

Example 5.1.5. Let $H = L^2([-\pi, \pi))$ and $e_n = e^{inx}$, then functions are in fact the limits of their Fourier series. ▲

Remark 5.1.6. We don't actually know yet that $\{ e^{inx} \} \subset L^2([-\pi, \pi))$ is complete, i.e. that $\langle f, e^{inx} \rangle = 0$ for all $n \in \mathbb{Z}$ implies that $f = 0$. This is intuitively true: the scalar products somehow measure how much f oscillates in the given frequency, so if f is constant (i.e. doesn't oscillate at all), then all scalar products except for $n = 0$ would be 0, and finally for this last one to be 0 the constant would indeed have to be 0.

Proof. (ii) We know that

$$\sum_\alpha |\langle f, e_\alpha \rangle|^2 \leq \|f\|_2^2,$$

i.e. Bessel's inequality. Suppose there exists some $\varepsilon > 0$ such that

$$\left\| f - \sum_{\alpha \in S} a_\alpha e_\alpha \right\| > \varepsilon$$

for every finite set S and $a_\alpha \in \mathbb{C}$. Let

$$M = \overline{\left\{ \sum_{\alpha \in S} a_\alpha e_\alpha \;\middle|\; S \text{ finite and } a_\alpha \in \mathbb{C} \right\}}.$$

An elementary fact from linear algebra is that we can write $f = f_M + f_{M^\perp}$, where $f_M \in M$ and $\langle f_{M^\perp}, m \rangle = 0$ for all $m \in M$.

Since $f \notin M$ (since f is a positive distance away from all things before taking closure, it remains outside afterward; just put a ball of radius $\varepsilon/2$ around it), then $f_{M^\perp} \neq 0$.

But $\langle f_{M^\perp}, e_\alpha \rangle = 0$ for every α since $e_\alpha \in M$, so by completeness $f_{M^\perp} = 0$, which is a contradiction. Therefore we cannot bound the distance between f and its Fourier series by any positive value.

A simple computation shows that

$$\left\| f - \sum_{\alpha \in S} \langle f, e_\alpha \rangle e_\alpha \right\| \leq \left\| f - \sum_{\alpha \in S} a_\alpha e_\alpha \right\|,$$

i.e. this sort of quantity is minimised by the Fourier coefficients. Moreover from last lecture

$$\left\| f - \sum_{\alpha \in S} \langle f, e_\alpha \rangle e_\alpha \right\|^2 = \|f\| - \sum_{\alpha \in S} |\langle f, e_\alpha \rangle|^2,$$

and by (ii) we can make this arbitrarily small, and so we have (i).

Furthermore

$$\langle f, g \rangle - \left\langle \sum_{\alpha \in S} \langle f, e_\alpha \rangle e_\alpha, g \right\rangle = \langle f, g \rangle - \sum_{\alpha \in S} \langle f, e_\alpha \rangle \overline{\langle g, e_\alpha \rangle} = \langle f, g \rangle - \sum_{\alpha \in S} \hat{f}(\alpha) \overline{\hat{g}(\alpha)}$$

but

$$\left| \langle f, g \rangle - \left\langle \sum_{\alpha \in S} \langle f, e_\alpha \rangle e_\alpha, g \right\rangle \right| = \left| \left\langle f - \sum_{\alpha \in S} \langle f, e_\alpha \rangle e_\alpha, g \right\rangle \right|$$

which by Cauchy-Schwartz inequality is bounded by

$$\left\| f - \sum_{\alpha \in S} \langle f, e_\alpha \rangle e_\alpha \right\| \cdot \|g\|$$

and this we can make arbitrarily small by (*ii*), and so (*iii*) follows. □

Theorem 5.1.7 (Riesz-Fischer). (*i*) *Given* $f \in H$ *and* $\{ e_\alpha \}$ *an orthonormal set indexed by* $\alpha \in A$*, then* $\{ \hat{f}(\alpha) \}_{\alpha \in A} \in \ell_2(A)$*.*

(*ii*) *Conversely, given a sequence* $\{ a_\alpha \} \in \ell^2(A)$*, then*

$$\sum_{\alpha \in A} a_\alpha e_\alpha$$

defines an element of H*.*

Recall before we proceed that $\ell^2(A)$ is the set of square summable sequences indexed by A, where we mean that the supremum of all finite square sums is finite.

Example 5.1.8. For instance,

$$\ell^2(\mathbb{N}) = \left\{ \{ a_n \} \;\middle|\; \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

▲

Proof. (*i*) By Bessel's inequality,

$$\sum_{\alpha \in A} |\langle f, e_\alpha \rangle|^2 \leq \|f\|^2$$

and so the sequence of Fourier coefficients is square summable and therefore is in $\ell^2(A)$.

(*ii*) Choose finite sets $B_n \subset A$ such that

$$\sum_{\alpha \in A} |a_\alpha|^2 - \sum_{\alpha \in B_n} |a_\alpha|^2 \leq \frac{1}{2^n}.$$

We can assume $B_{n+1} \supset B_n$. Then if $n > m$, since we are operating on an orthonormal set,

$$\left\| \sum_{\alpha \in B_n} a_\alpha e_\alpha - \sum_{\alpha \in B_m} a_\alpha e_\alpha \right\|^2 = \sum_{\alpha \in B_n \setminus B_m} |a_\alpha|^2 \leq \frac{2}{2^m}$$

which can be made arbitrarily small, whereby these sums indexed by n form a Cauchy sequence in H, and since H is a Hilbert space it must converge to an element in H. □

Example 5.1.9. Consider a positive function h on $[-1,1]$ such that

$$\int_{-1}^{1} h(t)\,dt = 1.$$

Suppose further that

$$\int_{-1}^{1} t^n h(t)\,dt \leq \infty$$

for all $n = 0, 1, 2, \ldots$.

Then the set $1, x, x^2, \ldots$ isn't orthonormal, but we can orthogonalise by Gram-Schmidt, i.e. let $v_0 = 1$,

$$v_1 = x - \frac{\langle x, v_0 \rangle}{\|v_0\|^2} v_0$$

and

$$v_2 = x^2 - \frac{\langle x^2, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle x^2, v_0 \rangle}{\|v_0\|^2} v_0$$

and so on. Finally normalise by $u_i = v_i / \|v_i\|$, and we have an orthonormal set of functions (which will be complete for some choices of h).

For instance $h(x) = 1/2$ yields the Legendre polynomials, $h(x) = (1-x)^\alpha (1+x)^\beta$ produces Jacobi functions, having Gaussian h gives Hermite functions, and so on.

These special functions arise as solutions to some ordinary differential equations, as orthonormal bases for some L^2 spaces, or as certain recurrence relations. ▲

Lecture 6 More on Hilbert Spaces

6.1 Haar Functions

Example 6.1.1. Consider the following set of functions on $[0,1]$ with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\,dt.$$

Let $h_0(x) = 1$, and let $h(x) = \chi_{[0,1/2)}(x) - \chi_{[1/2,1)}(x)$.

Now we recursively define

$$h_{j,k}(x) = 2^{j/2} h(2^j x - k)$$

for all $j = 0, 1, 2, \ldots$ and for each j we have $k \in \{0, 1, 2, \ldots, 2^j - 1\}$.

It is then straightforward to verify that

$$\|h_{j,k}\| = \int_0^1 |h_{j,k}(x)|^2\,dx = 1$$

and that

$$\langle h_{j,k}, h_{l,m} \rangle = \int_0^1 h_{j,k}(x) h_{l,m}(x)\,dx$$

is 0 if $(j,k) \neq (l,m)$. This is clear since if $j \neq l$, then either their supports are disjoint, in which we get 0, or they aren't, but then they average to 0 on the intersection of their supports. Similarly if $j = l$ but $k \neq m$, then the supports are disjoint, and again we have 0.

Hence the Haar functions form an orthonormal set. Moreover it is complete, and any function on $[0,1]$ can be written as a linear sum of Haar functions. ▲

6.2 Fourier Transform on L^2

We will consider functions in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

Lemma 6.2.1. *Let $f(x) = e^{-a|x|^2}$, with $a > 0$. Then*

$$\hat{f}(\xi) = \left(\frac{\pi}{a}\right)^{n/2} e^{-|\xi|^2/(4a)}.$$

In other words, the Fourier transform of a Gaussian is (up to some constant) again Gaussian.

Proof. It is pretty much by straightforward—but long—computation:

$$\begin{aligned}
\hat{f}(\xi) &= \int_{\mathbb{R}^d} e^{-ix\cdot\xi} e^{-a|x|^2}\,dx = \int_{\mathbb{R}^d} e^{-i(x_1\xi_1+\ldots+x_d\xi_d)-a(x_1^2+\ldots x_d^2)}\,dx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}} e^{-ix_1\xi_1-ax_1^2}\cdot e^{-ix_2\xi_2-ax_2^2}\cdot\ldots\cdot e^{-ix_d\xi_d-ax_d^2}\,dx_1\,dx_2\,\ldots\,dx_d \\
&= \left(\int_{\mathbb{R}} e^{-ix_1\xi_1-ax_1^2}\,dx_1\right)\cdot\ldots\cdot\left(\int_{\mathbb{R}} e^{-ix_d\xi_d-ax_d^2}\,dx_d\right) \\
&= \prod_{j=1}^{d}\int_{\mathbb{R}} e^{-ix_j\xi_j-ax_j^2}\,dx_j,
\end{aligned}$$

meaning that it is sufficient to evaluate this integral in one variable.

So we compute:

$$\begin{aligned}
\int_{\mathbb{R}} e^{-ix\xi-ax^2}\,dx = \int_{\mathbb{R}} e^{-a(x^2-\xi/aix)}\,dx &= e^{-\xi^2/(4a)}\int_{\mathbb{R}} e^{-a(x^2\xi/aix-\xi^2/(4a^2))}\,dx \\
&= e^{-\xi^2/(4a)}\int_{\mathbb{R}} e^{-a(x+i\xi/(2a))^2}\,dx.
\end{aligned}$$

We evaluate this by means of a contour integral of $g(z) = e^{-az^2}$, around the curve C which is the rectangle with corners in N, $N + i\xi/(2a)$, $-N + i\xi/(2a)$, and $-N$. Then

$$\begin{aligned}
\oint_C g(z)\,dz = \int_{-N}^{N} e^{-ax^2}\,dx + i\int_0^{\xi/(2a)} e^{-a(N+is)^2}\,ds \\
- \int_{-N}^{N} e^{-a(x+i\xi/(2a))^2}\,dx - i\int_0^{\xi/(2a)} e^{-a(-N+is)^2}\,ds.
\end{aligned}$$

The integrals on the vertical parts both vanish as $N \to \infty$ since they're products of a negative exponential and bounded terms. Moreover the whole thing is 0 since we're integrating an entire function on a simple, closed curve, and therefore

$$\int_{\mathbb{R}} e^{-ax^2}\,dx = \int_{\mathbb{R}} e^{-a(x+i\xi/(2a))^2}\,dx,$$

and the left-hand side is well-known to be $\sqrt{\pi/a}$. Therefore

$$\hat{f}(\xi) = \prod_{j=1}^{d} \int_{\mathbb{R}} e^{-ix_j\xi_j - ax_j^2}\, dx_j = \prod_{j=1}^{d} \left(\frac{\pi}{a}\right)^{n/2} e^{-\xi_j^2/(4a)} = \left(\frac{\pi}{a}\right)^{n/2} e^{-|\xi|^2/(4a)}. \quad \square$$

Lecture 7 Inverse Fourier Transform

7.1 Undoing Fourier Transforms

It is helpful to know that we can 'move the hat' inside of integrals:

Lemma 7.1.1. *If $f, g \in L^1(\mathbb{R}^d)$, then*

$$\int_{\mathbb{R}^d} \hat{f}(y)g(y)\, dy = \int_{\mathbb{R}^d} f(y)\hat{g}(y)\, dy.$$

Proof. We prove it by straightforward computation:

$$\int_{\mathbb{R}^d} \hat{f}(y)g(y)\, dy = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-iy\cdot x} f(x)\, dx g(y)\, dy,$$

and by Fubini's theorem we can switch the order of integration since the integrand is bounded, so

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-ix\cdot y} f(x)g(y)\, dy\, dx &= \int_{\mathbb{R}^d} f(x) \int_{\mathbb{R}^d} e^{-ix\cdot y} g(y)\, dy\, dx \\ &= \int_{\mathbb{R}^d} f(x)\hat{g}(x)\, dx. \quad \square\end{aligned}$$

Definition 7.1.2 (Inverse Fourier transform)**.** For $f \in L^1(\mathbb{R}^d)$ we define

$$\check{f}(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot t} f(t)\, dt,$$

the so-called ***inverse Fourier transform*** of f.

Theorem 7.1.3 (Fourier inversion theorem)**.** *If $f, \hat{f} \in L^1(\mathbb{R}^d)$ and f bounded, then*

$$f(x) = \check{\hat{f}}(x)$$

almost everywhere.

Proof. It follows from computation and our two latest lemmas. Let $\phi(t) = e^{ix\cdot t - \varepsilon^2|t|^2}$, x fixed. Then

$$\hat{\phi}(\xi) = \int_{\mathbb{R}^d} e^{ix\cdot t - \varepsilon^2|t|^2} e^{-i\xi\cdot t}\, dt = \int_{\mathbb{R}^d} e^{-i(\xi - x)\cdot t - \varepsilon^2|t|^2}\, dt.$$

By the lemma from last lecture this is

$$\sqrt{\frac{\pi}{\varepsilon^2}}^{d} e^{-|\xi - x|^2/(4\varepsilon^2)}.$$

Now for convenience write

$$g(\xi) = \sqrt{\pi}^d e^{-|\xi|^2/4},$$

then $\hat{\phi}(\xi) = \varepsilon^{-d} g((x-\xi)/\varepsilon)$.

Therefore

$$\begin{aligned}\int_{\mathbb{R}^d} f(y)\frac{1}{\varepsilon^d} g\Big(\frac{x-y}{\varepsilon}\Big)\,dy &= \int_{\mathbb{R}^d} f(y)\hat{\phi}(y)\,dy \\ &= \int_{\mathbb{R}^d} \hat{f}(y)\phi(y)\,dy = \int_{\mathbb{R}^d} \hat{f}(y) e^{ix\cdot y - \varepsilon^2|y|^2}\,dy.\end{aligned}$$

Now let $\varepsilon \to 0$. The right-hand side tends to $\check{\hat{f}}(x)$, where we can pass the limit inside since it is bounded by an integrable function.

For the left-hand side, suppose f is continuous at x and consider

$$\left|\int_{\mathbb{R}^d} f(y)\frac{1}{\varepsilon^d} g\Big(\frac{x-y}{\varepsilon}\Big)\,dy - f(x)\right|$$

which, since

$$\int_{\mathbb{R}^d} \frac{1}{\varepsilon^d} g\Big(\frac{x-y}{\varepsilon}\Big)\,dy = 1,$$

is the same if we multiply $f(x)$ by the above integrand to combine the integrals.

Then it is the same as

$$\left|\int_{\mathbb{R}^d} |f(y)-f(x)|\frac{1}{\varepsilon^d} g\Big(\frac{x-y}{\varepsilon}\Big)\,dy\right|$$

and if we let $u = (x-y)/\varepsilon$ this becomes

$$\int_{\mathbb{R}^d} |f(x-\varepsilon u) - f(x)| g(u)\,du.$$

Letting ε tend to 0, then this tends to 0, meaning that f equals $\check{\hat{f}}$ in L^1, but we want almost everywhere.

But the above means that there exists sequences ε_j tending to 0 for which the above holds, which results in convergence almost everywhere. □

Theorem 7.1.4 (Plancheral's theorem)**.** *Suppose $f, g \in L^1 \cap L^2(\mathbb{R}^d)$, then*

(*i*) $\displaystyle\int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,d\xi,$

(*ii*) $\displaystyle\int_{\mathbb{R}^d} |f(x)|^2\,dx = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |\hat{f}(y)|^2\,dy.$

Proof. We compute the first part:

$$\begin{aligned}\int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx &= \int_{\mathbb{R}^d} \check{\hat{f}}(x)\overline{g(x)}\,dx \\ &= \int_{\mathbb{R}^d} \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{iy\cdot x}\hat{f}(y)\,dy\overline{g(x)}\,dx \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{f}(y)\overline{\int_{\mathbb{R}^d} e^{-ix\cdot y} g(x)\,dx\,dy} \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{f}(y)\overline{\hat{g}(y)}\,dy.\end{aligned}$$

The second part follows immediately by letting $g = f$. □

Corollary 7.1.5. *Let $B(0,r)$ be the ball of radius r centred on 0 in $\mathbb{R}^d$. Let $f \in L^1 \cap L^2(\mathbb{R}^d)$. Then*

$$\widecheck{\chi_{B(0,r)}\hat{f}} \to f$$

in L^2 as $r \to \infty$.

Compare this with how partial sums of Fourier series converge to the function in L^2.

One might ask similar questions for $L^p \cap L^2$. In one dimension this is true, but in dimensions 2 or greater it is not. This is due to Fefferman in the seventies, part of what earned him the Field's medal.

Lecture 8 Fejer Kernels

8.1 Fejer Kernels and Approximate Identities

One of our fondest hopes in this course is that the Fourier series of a function converges, in some reasonable way, to the function itself. Another way of asking if this happens is to study the partial sums

$$S_n f(x) = \sum_{k=-n}^{n} \hat{f}(k) e^{ikx}$$

and ask whether $S_n f(x) \to f(x)$ in some sense of convergence, be it in L^p norm, almost everywhere, uniformly, and so on.

What Fejer showed is that if we define

$$\sigma_n = \frac{S_0 + S_1 + \ldots + S_n}{n+1},$$

i.e. the arithmetic average of the first $n+1$ partial sums, then $\sigma_n f \to f$ in L^p and almost everywhere.

In general, suppose $\{ a_n \}_{n=0}^{\infty} \subset \mathbb{R}$, and $a_n \to L$, then $\sigma_n \to L$ as well. We say that a sequence for which σ_n converges is ***Cesaró summable***.

The converse is in general not true:

Example 8.1.1. Let $a_n = (-1)^n$. Clearly this does not converge to anything—it jumps between 1 and -1 indefinitely. However it does converge in the Cesaró sense, since $\sigma_0 = 1$, $\sigma_1 = 0$, $\sigma_2 = 1/3$, $\sigma_3 = 0$, $\sigma_4 = 1/5$, and so on. ▲

There are other curious ways to sum things:

Example 8.1.2. For a sequence $a_0, a_1, \ldots$, let $0 < r < 1$ and consider $S(r) = a_0 + a_1 r + a_2 r^2 + \ldots$. If

$$\lim_{r \to 1} S(r)$$

exists, we say that the sequence $a_0, a_1, \ldots$ is ***Abel summable***. ▲

Let $f \in L^1([-\pi, \pi))$, then

$$\begin{aligned} S_n f(x) &= \sum_{k=-n}^{n} \hat{f}(k) e^{ikx} = \sum_{k=-n}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt e^{ikx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{k=-n}^{n} e^{ik(x-t)}\, dt. \end{aligned}$$

We identify the inner sum

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx}$$

called the ***Dirichlet kernel***. Therefore

$$S_k f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) D_k(x-t)\, dt$$

and

$$\sigma_n f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \sum_{k=0}^{n} \frac{D_k(x-t)}{n+1}\, dt.$$

Now we once more identify the inner sum as a new piece of notation—this time it will turn out to be very useful, after some algebra:

$$K_n(x) = \sum_{k=0}^{n} \frac{D_k(x)}{n+1}$$

is the so-called ***Fejer kernel***, and using it we have

$$\sigma_n f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) K_n(x-t)\, dt.$$

We will now spend some rewriting the Fejer kernel in a more practical way:

$$K_n(x) = \frac{e^{i0x} + \sum_{k=-1}^{1} e^{ikx} + \sum_{k=-2}^{2} e^{ikx} + \ldots + \sum_{k=-n}^{n} e^{ikx}}{n+1}$$

and if we just count how many times each e^{ikx} appears for each k, we can clearly rewrite it as

$$K_n(x) = \sum_{l=-n}^{n} \frac{(n+1)-|l|}{n+1} e^{ils}.$$

Now as an aside, note that

$$\Big(\sum_{j=0}^{n} e^{i(j-n/2)x}\Big)^2 = \sum_{l=0}^{2n} \sum_{j+k=l} e^{i(l-n)x} = \sum_{l=0}^{2n} ((n+1)-|l-n|) e^{i(l-n)x}$$

and therefore

$$K_n(x) = \frac{1}{n+1}\Big(\sum_{j=0}^{n} e^{i(j-n/2)x}\Big)^2 = \frac{1}{n+1}\Big(e^{-in/2x} \sum_{j=0}^{n} e^{ijx}\Big)^2$$

and the sum in the last step is geometric, so this is the same as

$$\frac{1}{n+1}\Big(e^{-in/2x} \frac{1-(e^{ix})^{n+1}}{1-e^{ix}}\Big)^2 = \frac{1}{n+1}\Big(\frac{e^{-in/2x} i e^{i(n/2+1)x}}{1-e^{ix}}\Big)^2$$

and by multiplying and dividing by $e^{-ix/2}$ we get

$$\frac{1}{n+1}\Big(\frac{e^{-i(n+1)/2x} - e^{i(n+1)/2x}}{e^{-ix/2} - e^{ix/2}}\Big)^2 = \frac{1}{n+1}\Big(\frac{-2i\sin((n+1)/2x)}{-2i\sin(x/2)}\Big)^2$$

and simplifying this is just

$$K_n(x) = \frac{1}{n+1}\Big(\frac{\sin(\frac{n+1}{2}x)}{\sin(\frac{x}{2})}\Big)^2.$$

Lemma 8.1.3 (Properties of the Fejer kernel). *(i)* $K_n(x) \geq 0$ *for all* $x \in [-\pi, \pi)$.

(ii) *Fix a* $\delta > 0$. *Then* $K_n(x) \to -$ *as* $n \to \infty$ *uniformly on* $[-\pi, \pi) \setminus (-\delta, \delta)$.

(iii) $\displaystyle \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\,dx = 1.$

Proof. (i) is quite clear, since we have the square of a real number.

(ii) is reconciled by noting that the sin in the numerator is bounded by 1, and in the bottom we can take $x = \delta$ since $|x| > \delta$ means that

$$K_n(x) = \frac{1}{n+1}\Big(\frac{\sin(\frac{n+1}{2}x)}{\sin(\frac{x}{2})}\Big)^2 = \frac{1}{n+1}\frac{1}{\sin(\delta/2)^2}$$

which goes to 0 uniformly as $n \to \infty$, since the estimate is independent of x so long as $|x| > \delta$.

(iii) We just compute, recalling one of the earlier forms of the Fejer kernel:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=-n}^{n}\frac{n+1-|k|}{n+1}e^{ikx}\,dx \\ &= \sum_{k=-n}^{n}\frac{1}{2\pi}\frac{n+1-|k|}{n+1}\int_{-\pi}^{\pi} e^{ikx}\,dx\end{aligned}$$

and this last integral is 0 unless $k = 0$, and so the whole thing is equal to

$$\frac{1}{2\pi}\frac{n+1}{n+1}2\pi = 1.$$

□

Remark 8.1.4. Any sequence of functions satisfying (i)–(iii) is called an ***approximate identity***.

Theorem 8.1.5. *(i) For* $f \in L^p([-\pi, \pi))$, $1 \leq p < \infty$, *then* $\sigma_n f(x) \to f(x)$ *in the* $L^p([-\pi, \pi))$ *norm, i.e.*

$$\|\sigma_n f - f\|_p = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |\sigma_n f(x) - f(x)|^p\,dx\Big)^{1/p} \to 0$$

as $n \to \infty$.

(ii) If $f \in L^1$, *and* f *is continuous at* x, *then* $\sigma_n f(x) \to f(x)$.

(iii) If $f \in L^1([-\pi, \pi))$ *then* $\sigma_n f(x) \to f(x)$ *almost everywhere.*

Remark 8.1.6. The proof of this relies on the fact that $K_n(x)$ is an approximate identity, and no other special properties of $K_n(x)$. Therefore any other approximate identity has the same properties.

Remark 8.1.7. The answer to our original question—actual convergence of partial sums, not of the means—is a question that has been subject to much study. Carleson proved that $S_n f(x) \to f(x)$ almost everywhere in L^2, and Hunt later proved the same in L^p for $1 < p < \infty$. It was shown to be false by Kolmogorov for $p = 1$.

Lecture 9 Convergence of Cesaró Means

9.1 Convergence of Fourier Sums

We prove the theorem stated at the end of last lecture.

Theorem 9.1.1. *(i) For $f \in L^p([-\pi,\pi))$, $1 \le p < \infty$, then $\sigma_n f(x) \to f(x)$ in the $L^p([-\pi,\pi))$ norm, i.e.*

$$\|\sigma_n f - f\|_p = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |\sigma_n f(x) - f(x)|^p\,dx\Big)^{1/p} \to 0$$

as $n \to \infty$.

(ii) If $f \in L^1$, and f is continuous at x, then $\sigma_n f(x) \to f(x)$.

(iii) If $f \in L^1([-\pi,\pi))$ then $\sigma_n f(x) \to f(x)$ almost everywhere.

Proof. We start by proving (ii). Fix an x, then

$$\begin{aligned}\sigma_n f(x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)K_n(x-t)\,dt - f(x)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)K_n(t)\,dt - \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)K_n(t)\,dt\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x-t)-f(x))K_n(x)\,dt.\end{aligned}$$

Letting $\varepsilon > 0$, then there exists a $\delta > 0$ such that if $|t| < \delta$ then $|f(x-t)-f(x)| < \varepsilon$ by continuity.

By the second property of approximate identities, we can choose an n large enough so that $|K_n(t)| < \varepsilon/(\|f\|_1 + |f(x)|)$ for all $t \in [-\pi,\pi) \setminus (-\delta,\delta)$. Then

$$\begin{aligned}|\sigma_n f(x) - f(x)| &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x-t)-f(x)|K_n(t)\,dt\\ &= \frac{1}{2\pi}\int_{-\delta}^{\delta} |f(x-t)-f(x)|K_n(t)\,dt +\\ &\quad + \frac{1}{2\pi}\int_{\substack{[-\pi,\pi)\\ \setminus(-\delta,\delta)}} |f(x-t)-f(x)|K_n(t)\,dt\\ &\le \varepsilon\cdot 1 + \frac{\varepsilon(\|f\|_1 + |f(x)|)}{\|f\|_1 + |f(x)|} = 2\varepsilon.\end{aligned}$$

For (i), we compute and cleverly use Jensen's inequality at one point:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} |\sigma_n f(x) - f(x)|^p\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}(f(x-t)-f(x))K_n(t)\,dt\right|^p dx\\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x-t)-f(x)|^p K_n(t)\,dt\,dx.\end{aligned}$$

Now noting that $d\mu = K_n(t)/(2\pi)\,dt$ is a measure with total mass 1, and that $|\cdot|^p$ is convex, we have by Jensen's inequality that this equals

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x-t)-f(x)|^p\,dx}_{=h(t)}K_n(t)\,dt.$$

We therefore have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}h(t)K_n(t)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}h(0-t)K_n(t)\,dt$$

but this goes to $h(0) = 0$ as $n \to \infty$, by (ii).

(iii) We have

$$\begin{aligned}
|\sigma_n f(x)| &= \frac{1}{2\pi}\left|\int_{-\pi}^{\pi}K_n(t)f(x-t)\,dt\right| \le \frac{1}{2\pi}\int_{-\pi}^{\pi}K_n(t)|f(x-t)|\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_0^{K_n(t)}dr|f(x-t)|\,dt \\
&= \frac{1}{2\pi}\int_0^{\infty}\chi_{[0,K_n(t))}(r)\,dr|f(x-t)|\,dt \\
&= \int_0^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\chi_{[0,K_n(t))}(r)|f(x-t)|\,dt\,dr.
\end{aligned}$$

Now $\chi_{[0,K_n(t))}(r)$ is the same as $\chi_{[r,\infty)}(K_n(t))$, meaning that we have

$$\int_0^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\chi_{[r,\infty)}(K_n(t))|f(x-t)|\,dt\,dr.$$

Letting I_r be the subset of $[-\pi,\pi)$ where that characteristic function is 1, we have

$$\int_0^{\infty}\frac{|I_r|}{2\pi}\int_{I_r}\frac{1}{|I_r|}|f(x-t)|\,dt\,dr.$$

This inner integral is bounded by the maximal average of f, defined by

$$Mf(x) = \sup_{r>0}\frac{1}{B_r(x)}\int_{B_r(x)}|f(t)|\,dr,$$

which by a theorem of Hardy and Little wood is proven to be in L^p if f is.

So our expression is bounded by

$$Mf(x)\|K_n\|_1 = Mf(x).$$

We wish to conclude that $\sigma_n f(x) \le Mf(x)$. To do so, let

$$Tf(x) = \limsup_{n\to\infty}|\sigma_n f(x) - f(x)|.$$

If we can show that $Tf(x) = 0$ almost everywhere, we are done. Let $N \in \mathbb{N}$, and choose a continuous function g such that

$$\|f-g\|_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t)-g(t)|\,dt < \frac{1}{N}.$$

In fact (ii) essentially shows that $\sigma_n g \to g$ uniformly. Then

$$\begin{aligned} |\sigma_n f(x) - f(x)| &= |\sigma_n (f-g)(x) + \sigma_n g(x) - (f-g)(x) - g(x)| \\ &\leq |\sigma_n (f-g)(x)| + |\sigma_n g(x) - g(x)| + |(f-g)(x)| \\ &\leq M(f-g)(x) + |\sigma g(x) - g(x)| + |(f-g)(x)|. \end{aligned}$$

Taking lim sup, we get $Tf(x) \leq M(f-g)(x) + |f(x) - g(x)|$ and

$$\{\, x \mid Tf(x) > \varepsilon \,\} \subset \{\, x \mid M(f-g)(x) > \frac{\varepsilon}{2} \,\} \cup \{\, x \mid |f(x) - g(x)| > \frac{\varepsilon}{2} \,\}$$

and taking measures of these sets we have that

$$Tf(x) \leq \frac{C/N}{\varepsilon/2} + \frac{1/N}{\varepsilon/2}.$$

Now let $N \to \infty$, then $|\{\, x \mid Tf(x) > \varepsilon \,\}| = 0$. Take a countable sequence of ε_n going to 0, then $|\{\, x \mid Tf(x) > 0 \,\}| = 0$, and we're done. □

Lecture 10 Toward Convergance of Partial Sums

10.1 Dirichlet Kernels

Recall that

$$\begin{aligned} S_n f(x) &= \sum_{k=-n}^{n} \hat{f}(k) e^{ikx} = \sum_{k=-n}^{n} \Big(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt \Big) e^{ikx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{k=-n}^{n} e^{ik(x-t)}\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_n(x-t)\, dt \end{aligned}$$

wherre

$$D_n(s) = \sum_{k=-n}^{n} e^{iks}$$

is the ***Dirichlet Kernel***.

We can rewrite

$$e^{is/2} D_n(s) - e^{-is/2} D_n(s) = \sum_{k=-n}^{n} e^{i(k+1/2)s} - \sum_{k=-n}^{n} e^{i(k-1/2)s} = e^{i(n+1/2)s} - e^{-i(n-1/2)s}.$$

Therefore

$$D_n(s) = \frac{e^{i(n+1/2)s} - e^{-i(n+1/2)s}}{e^{is/2} - e^{-is/2}} = \frac{2i \sin\left(\frac{n+1}{2} s\right)}{2i \sin\left(\frac{s}{2}\right)} = \frac{\sin\left(\frac{n+1}{2} s\right)}{\sin\left(\frac{s}{2}\right)}.$$

So in all

$$S_n f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin\left(\frac{n+1}{2}(x-t)\right)}{\sin\left(\frac{x-t}{2}\right)}\, dt.$$

Remark 10.1.1. Note that the Dirichlet kernel D_n is *not* an approximate identity; certainly it changes sign, and also

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}|D_n(s)|\,dt &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin\left(\frac{n+1}{2}s\right)}{\sin\left(\frac{s}{2}\right)}\right|ds \geq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin\left(\frac{n+1}{2}s\right)}{s}\right|ds \\
&= \frac{1}{\pi}\int_{(-n+1)/2\pi}^{(n+1)/2\pi}\frac{|\sin(u)|}{|u|}\,du = \frac{2}{\pi}\int_{0}^{(n+1)/2\pi}\frac{|\sin(u)|}{u}\,du \\
&\geq \frac{2}{\pi}\sum_{k=0}^{(n-1)/2}\int_{k\pi}^{(k+1)\pi}\frac{|\sin(u)|}{k\pi} = \frac{2}{\pi^2}\sum_{k=0}^{(n-1)/2}\int_{k\pi}^{(k+1)\pi}|\sin(u)|\,du \\
&= \frac{2}{\pi^2}\sum_{k=0}^{(n-1)/2}\frac{1}{k} \approx \frac{2}{\pi^2}\log((n-1)/2),
\end{aligned}$$

meaning that the L^1 norm of the Dirichlet kernel diverges as $n \to \infty$.

There are now two principal things we wish to discuss. First: convergence of the partial sums for continuous functions.

10.2 Convergence for Continuous Functions

For $f \in C([-\pi, \pi))$, set

$$T_n f = S_n f(0) = \sum_{k=-n}^{n} e^{ik0}\hat{f}(k) = \sum_{k=-n}^{n}\hat{f}(k).$$

Note that

$$|\hat{f}(k)| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{ikt}\,dt\right| \leq \|f\|_\infty,$$

if f is bounded. Then also

$$|T_n f| \leq (2n+1)\|f\|_\infty.$$

In other words every T_n is a bounded linear functional from $C([-\pi, \pi))$ to $\mathbb{C}$, however the bound grows with n.

We'll now do something clever: for a fixed n, construct the function g such that

$$g(t) = \begin{cases} 1, & \text{if } D_n(t) \geq 0 \\ -1, & \text{if } D_n(t) < 0. \end{cases}$$

In other words g is a bunch of line segments at $y = 1$ and $y = -1$, jumping between the two. Certainly g is discontinuous at these jumps, but we can approximate it to any desired accuracy by $g_j \in C([-\pi, \pi))$ such that $\|g_j\|_\infty \leq 1$ and $g_j(t) \to g(t)$ pointwise, where we connect the line segments with steeper and steeper lines for each j.

Then

$$\begin{aligned}
\lim_{j\to\infty} T_n g_j = \lim_{j\to\infty} S_n g_j(0) &= \lim_{j\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} g_j(t)D_n(0-t)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_n(t)|\,dt \approx \log(n)|g_j|_\infty.
\end{aligned}$$

by dominated convergence.

Therefore

$$\|T_n\| = \sup_{g\in C([-\pi,\pi))} \frac{T_n g}{norm g_\infty} \geq \log(n).$$

Therefore we have a Banach space containing a family of bounded linear functionals T_n. Recall the following from functional analysis:

Theorem 10.2.1 (Principle of Bounded Convergence)**.** *Suppose X is a Banach space, T_α, $\alpha \in \Lambda$ is a family of bounded linear functionals on X, i.e. for all α*

$$\|T_\alpha\| = \sup_{x\in X} |T_\alpha x| \|x\| < \infty.$$

Then either

$$\sup_{\alpha\in\Lambda} |T_\alpha x| = \infty$$

for all x in a dense subset of X, or there exists an M such that$\|T_\alpha\| \leq M$ for all $\alpha \in \Lambda$.

In our case $\|T_n\| \geq \log(n)$ (maybe with some suitable constant) for all n, and so by the bounded convergence principle

$$\sup_n \|T_n f\| = \infty$$

for all f in a dense subset of $C([-\pi,\pi))$. In other words $S_n f(0)$ doesn't converge to $f(0)$ pointwise for all f in some dense subset, so even for nice, continuous functions we are very far indeed from having pointwise convergence of the Fourier series.

Next time we'll tackle the same problem with convergence in L^p instead.

Lecture 11 Convergence in L^p

Last time we established that there exists a dense subset of $C([-\pi,\pi))$ with $S_n f(x)$ diverging. Recall that

$$S_n f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(x-t) f(t)\, dt$$

with

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

11.1 Convergence in L^p

Theorem 11.1.1. *If $f \in L^p([-\pi,\pi))$, with $1 < p < \infty$, then*

$$\|S_n f - f\|_p \to 0$$

as $n \to \infty$.

In order to study this, we will make use of the ***conjugate series*** of f, which is defined as

$$\tilde{f} \sim \sum_{k=-\infty}^{\infty} -i\,\mathrm{sgn}(k)\hat{f}(k)e^{ikx}.$$

This turns out to be an interesting construction:

Theorem 11.1.2 (M. Riesz)**.** *Given $f \in L^p([-\pi,\pi))$, $1 < p < \infty$, $\tilde{f}$ defines a unique function in L^p, i.e. there exists a unique function $\tilde{f} \in L^p([-\pi,\pi))$ such that $\hat{\tilde{f}}(k) = i\,\mathrm{sgn}(k)\hat{f}(k)$ for every k.*

Furthermore there exists a constant C_p such that $\|\tilde{f}\|_p \le C_p\|f\|_p$.

Remark 11.1.3. Given a harmonic function u on the unit disc, and assuming u is somewhat well-behaved, then it has boundary values on the unit circle. It turns out then that

$$u(\theta) = \lim_{r\to 1^-} u(rx^{i\theta}) = f(\theta),$$

and with v being the conjugate function of u we have $v(\theta) = \tilde{f}(\theta)$. That is to say, this conjugate series does not come from nowhere!

For $f \in L^p([-\pi,\pi))$, define $P_-f = (f - i\tilde{f})/2$ and $P_+f = (f + i\tilde{f})/2$. Then

$$\|P_-f\|_p \le \frac{\|f\|_p}{2} + \frac{\|\tilde{f}\|_p}{2} \le \frac{\|f\|_p}{2} + \frac{C_p\|f\|_p}{2} = \frac{1+C_p}{2}\|f\|_p,$$

so $P_- \colon L^p([-\pi,\pi)) \to L^p([-\pi,\pi))$ is a bounded operator, since

$$\|P_-\| = \sup_{f\in L^p} \frac{normP_-f_p}{\|f\|_p} = \frac{1+C_p}{2}.$$

In the same way P_+ is a bounded operator. We call these P_- and P_+ because they are projections:

$$\begin{aligned} P_+f &\sim \sum_{k=-\infty}^{\infty} \frac{\hat{f}(k)}{2}e^{ikx} + i\sum_{k=-\infty}^{\infty} \frac{-i\,\mathrm{sgn}(k)\hat{f}(k)}{2}e^{ikx} \\ &= \sum_{k=-\infty}^{\infty} \frac{\hat{f}(k) + \mathrm{sgn}(k)\hat{f}(k)}{2}e^{ikx} = \sum_{k=0}^{\infty}\hat{f}(k)e^{ikx}. \end{aligned}$$

If we now apply P_+ again,

$$P_+ \circ P_+(f) = \sum_{k=0}^{\infty}\frac{\hat{f}(k)}{2}e^{ikx} + i\sum_{k=0}^{\infty}\frac{-i\,\mathrm{sgn}(k)\hat{f}(k)}{2}e^{ikx} = \sum_{k=0}^{\infty}\hat{f}(k)e^{ikx} = P_+f.$$

In other words, P_+ sets negative Fourier coefficients to 0, and P_- sets nonnegative ones to 0.

The reason this is interesting is this: S_nf is also the Fourier series of f with a bunch of coefficients set to 0;

$$\begin{aligned}
S_n f(x) = \sum_{k=-n}^{n} \hat{f}(k)e^{ikx} &= e^{ix(n+1)} \sum_{k=-n}^{n} \hat{f}(k)e^{ix(k-n-1)} \\
&= e^{ix(n+1)} P_-\Big(\sum_{k=-n}^{\infty} \hat{f}(k)e^{ix(k-n-1)} \Big) \\
&= e^{ix(n+1)} P_-\Big(e^{-ix(2n+2)} \sum_{k=-n}^{\infty} \hat{f}(k)e^{ix(k+n)} \Big) \\
&= e^{ix(n+1)} P_-\Big(e^{-ix(2n+2)} P_+\Big(\sum_{n=-\infty}^{\infty} \hat{f}(k)e^{ix(k+n)} \Big)\Big) \\
&= e^{ix(n+1)} P_-\Big(e^{-ix(2n+2)} P_+\Big(e^{-inx} \sum_{n=-\infty}^{\infty} \hat{f}(k)e^{ikx} \Big)\Big)
\end{aligned}$$

which means that $\|S_n f\|_p \le C_p^2 \|f\|_p$.

Now we're ready to prove that $S_n f \to f$ in L^p:

Proof. Let $\varepsilon > 0$. Pick a trigonometric polynomial q (the Fejer kernels, for instance) with $\|f - q\|_p < \varepsilon$. Then

$$\begin{aligned}
\|S_n f - f\|_p &\le \|S_n f - S_n q\|_p + \|S_n q - q\|_p + \|q - f\|_p \\
&= \|S_n(f - q)\|_p + \|S_n q - q\|_p + \|q - f\|_p \\
&\le C_p^2 \|f - q\|_p + \|S_n q - q\|_p + \|q - f\|_p
\end{aligned}$$

which is bounded by $C_p^2 \varepsilon + \varepsilon$ since the middle term is 0 for a sufficiently large n. □

11.2 Almost Everywhere Convergence

Kolmogorov proved in 1925 that there exists a function $f \in L^1([-\pi, \pi))$ such that $S_n f(x)$ diverges almost everywhere. He later showed that in fact there exists such an f that diverges on all x, not just on sets of positive measure.

Carleson proved in 1966 that if you instead work in L^2, then $S_n f(x) \to f(x)$ almost everywhere.

Hunt improved this in 1967 to work for $f \in L^p$ for $1 < p < \infty$.

We will prove this under the assumption of a lemma we will not prove:

Lemma 11.2.1. *Define*

$$Mf(x) = \sup_n |S_n f(x)| = \sup_n \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(x - t) f(t)\, dt \right|.$$

For $1 < p < \infty$, there exists a constant C_p such that $\|Mf\|_p \le C_p \|f\|_p$ for every $f \in L^p$.

Let us prove that $S_n f$ converges almost everywhere to f given the lemma:

Proof. Set

$$Tf(x) = \limsup_{n\to\infty} |S_n f(x) - f(x)|.$$

We want $Tf(x) = 0$ almost everywhere. Let $N \in \mathbb{N}$, and choose a trigonometric polynomial q such that $\|f - q\|_p < 1/N$. Then

$$|S_n f - f| \le |S_n f - S_n q| + |S_n q - q| + |q - f|.$$

Note that $S_n q \to q$ at every x since they're equal for large n. Taking lim sup we get

$$Tf \le M(f - q) + |f - q|$$

and so

$$|SetxTf > \varepsilon| = |\{ x \mid M(f-q) \} > \varepsilon/2| + |\{ x \mid |f-q| > \varepsilon/2 \}|$$

but by the lemma $\|M(f-q)\|_p \le C_p \|f - q\|_p$. By Chebyshev

$$\begin{aligned} |\{ x \mid Tf(x) > \varepsilon \}| &\le \frac{\|M(f-q)\|_p}{(\varepsilon/2)^p} + \frac{\|f-q\|_p}{(\varepsilon/2)^p} \\ &\le \frac{2^p}{\varepsilon^p}\Big(C_p \|f-q\|_p + \|f-q\|_p \Big) \\ &=\le \frac{2^p}{\varepsilon^p}(C_p + 1)/N \end{aligned}$$

which goes to 0 as $N \to \infty$. □

Lecture 12 Maximal Functions

12.1 Hardy-Littlewood Maximal Functions

Definition 12.1.1. Suppose f is a Lebesgue measurable function on $\mathbb{R}^d$. We say that $f \in L^1_{loc}(\mathbb{R}^d)$ if

$$\int_B |f(x)|\, dx < \infty$$

for every ball $B \subset \mathbb{R}^d$.

Note that we do not require

$$\int_{\mathbb{R}^d} |f(x)|\, dx < \infty,$$

though clearly L^1 functions have the local property.

Definition 12.1.2 (Maximal function)**.** Suppose $f \in L^1_{loc}(\mathbb{R}^d)$, we set

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy,$$

where by $B(x,r)$ we mean the ball centred on x with radius r.

Remark 12.1.3.

- Given a function $f \in L^1_{loc}(\mathbb{R}^d)$, it is not clear whether $Mf(x) < \infty$ at any x.

- Suppose $f \in L^p(\mathbb{R}^d)$, and let $B \subset R^d$ be a ball. Then
$$\begin{aligned}\int_B |f(y)|\,dy &= \int_{\mathbb{R}^d} \chi_B(y)|f(y)|\,dy \\ &\le \Big(\int_{\mathbb{R}^d} |\chi_B(y)|^q\,dy\Big)^{1/q}\Big(\int_{\mathbb{R}^d} |f(y)|^p\,dy\Big)^{1/p} \le |B|^{1/q}\|f\|_p.\end{aligned}$$
In other words $L^p \subset L^1_{loc}$, as hinted at above.
- If μ is a positive Borel measure, then we can define the maximal function of a measure analogously:
$$M\mu(x) = \sup_{r>0} \frac{1}{|B(x,r)|}\mu(B(x,r)) = \sup_{r>0} \frac{1}{|B(x,r)|}\int_{B(x,r)} d\mu.$$
This is a generalisation of the previous definition, since $d\mu = |f(y)|\,dy$ defines a measure.

Proposition 12.1.4. *If μ is a Borel measure, then $M\mu$ is a Borel measurable function.*

Proof. Let $\lambda > 0$, and let $E_\lambda = \{\, x \mid M\mu(x) > \lambda \,\}$. Take $x \in E_\lambda$. Then there exists some $r_0 > 0$ such that
$$\frac{\mu(B(x,r_0))}{|B(x,r_0)|} = t > \lambda.$$
Choose δ such that $(r_0+\delta)^n < r_0^n t/\lambda$, which is possible since $t/\lambda > 1$. Suppose $y \in B(x,\delta)$. Then $B(y, r_0+\delta) \supset B(x,r_0)$. This follows directly by the triangle inequality; take z to be in $B(x,r_0)$, then
$$d(z,y) \le d(x,y) + d(x,z) \le \delta + r_0.$$
Therefore
$$\begin{aligned}\mu(B(y,r_0+\delta)) &\ge \mu(B(x,r_0)) = t|B(x,r_0)| \\ &> \frac{(r_0+\delta)^n}{r_0^n}\lambda|B(x,r)| = \lambda|B(y,r_0+\delta)|.\end{aligned}$$
Since $(r_0+\delta)^n/r_0^n$ is the ratio of the volumes of the two balls. This means that $M\mu(y) > \lambda$, meaning that E_y is open, making it measurable, which in turn makes $M\mu$ a measurable function. □

Remark 12.1.5. One can also define $M\mu$ and Mf using cubes instead of balls, say $Q(x,r) = \{\, y \in \mathbb{R}^d \mid |x_i - y_i| < r,\ i = 1,2,\dots,d \,\}$. Then
$$M_Q\mu(x) = \sup_{r>0}\frac{\mu(Q(x,r))}{|Q(x,r)|},$$
using which
$$\frac{\mu(B(x,r))}{|B(x,r)|} \le \frac{\mu(Q(x,r))}{|Q(x,r)|}\frac{|Q(x,r)|}{|B(x,r)|} = \frac{\mu(Q(x,r))}{|Q(x,r)|}\frac{2^d}{c_d}$$
where $c_n = \pi^{n/2}\Gamma(n/2+1)$ is the volume of a d-dimensional unit sphere. In other words these different ways of defining the maximal function are the same up to multiplication by some constant depending on the dimension.

Remark 12.1.6. One can also define it in terms of surface integrals over the boundaries of balls, but this is less well-understood.

Theorem 12.1.7 (Hardy-Littlewood, 1930)**.** *If μ is a positive Borel measure on $\mathbb{R}^d$, then for every $\lambda > 0$,*

$$|\{\, x \in \mathbb{R}^d \mid M\mu(x) > \lambda \,\}| \geq \frac{3^d}{\lambda}\mu(\mathbb{R}^d).$$

Remark 12.1.8. If $f \in L^1_{loc}(\mathbb{R}^d)$ then

$$\mu(E) = \int_E |f(y)|\, dy,$$

meaning that

$$|\{\, x \in \mathbb{R}^d \mid Mf(x) > \lambda \,\}| \leq \frac{3^d}{\lambda}\int_{\mathbb{R}^d} |f(y)|\, dy = \frac{3^d}{\lambda}\|f\|_1.$$

Remark 12.1.9. Let δ_0 be the Dirac measure at 0, i.e. $\delta_0(E) = 1$ if $0 \in E$, and 0 otherwise. In $\mathbb{R}^1$, we then have

$$\frac{\delta_0(B(x, |x| + \varepsilon))}{|B(x, |x| + \varepsilon)|} = \frac{1}{2(|x| + \varepsilon)},$$

meaning that $M\delta_0(x) \geq 1/(2|x|)$. Rearranging we therefore have

$$\{\, x \mid M\delta_0(x) > \lambda \,\} \supset \left\{\, x \,\middle|\, |x| < \frac{1}{2\lambda} \,\right\},$$

which it we measure the sets yields

$$|\{\, x \mid M\delta_0(x) > \lambda \,\}| \geq \left|\left\{\, x \,\middle|\, |x| < \frac{1}{2\lambda} \,\right\}\right| = \frac{1}{\lambda}.$$

This serves to demonstrate that the bound in Hardy-Littlewood's theorem is about as good as it gets.

Remark 12.1.10. Suppose $f \in L^1(\mathbb{R}^d)$. Then the theorem says

$$|\{\, x \in \mathbb{R}^d \mid Mf(x) > \lambda \,\}| \leq \frac{3^d}{\lambda}\|f\|_1.$$

Suppose it were the case that we knew

$$\int_{\mathbb{R}^d} Mf(x)\, dx \leq C \int_{\mathbb{R}^d} |f(x)|\, dx.$$

Then

$$\begin{aligned} \lambda|\{\, x \mid Mf(x) > \lambda \,\}| &\leq \int_{\{\, x \mid Mf(x) > \lambda \,\}} Mf(x)\, dx \leq \int_{\mathbb{R}^d} Mf(x)\, dx \\ &\leq C \int_{\mathbb{R}^d} |f(x)|\, dx. \end{aligned}$$

It turns out, however, that what we assume above it never true.

We will state and prove a lemma that takes us most of the way toward the Hardy-Littlewood theorem:

Lemma 12.1.11 (Wiener's covering lemma). *Suppose W is a set in $\mathbb{R}^d$, and that*

$$W \subset \bigcup_{i=1}^{N} B(x_i, r_i),$$

i.e. W can be covered by a finite set of balls. Then there exists a set of indices $S \subset \{1, 2, \ldots, N\}$ such that

(*i*) *The balls $\{B(x_i, r_i) \mid i \in S\}$ are disjoint,*

(*ii*) $W \subset \bigcup_{i \in S} B(x_i, 3r_i)$, *and*

(*iii*) $|W| \leq 3^d \sum_{i \in S} |B(x_i, r_i)|$.

This is a so-called covering lemma, since it tells about covers. In particular it tells us that if we can cover a set in a finite set of balls, then we can pick a subset of that finite set of balls that is disjoint, and still covers the set if we blow each of the remaining balls up to thrice their original radius.

Proof. (*ii*) implies (*iii*) quite trivially:

$$|W| \leq \left| \bigcup_{i \in S} B(x_i, 3r_i) \right| \leq \sum_{i \in S} |B(x_i, 3r_i)| = 3^d \sum_{i \in S} |B(x_i, r_i)|.$$

By reordering, we may assume $r_1 \geq r_2 \geq \ldots \geq r_N$. Let $i_1 = 1$, and consider the biggest ball $B(x_1, r_1) = B(x_{i_1}, r_{i_1})$. Discard all balls that intersect this one (the idea being that the three-fold enlargement of this ball envelopes any ball that intersects it, so we don't need them anyway).

It no such balls exist, we simply stop. If not, let $B(x_{i_2}, r_{i_2})$ be the largest remaining ball in the list. Throw away any ball that intersects it; If no such ball exists, stop, otherwise continue in the same fashion.

This process must eventually terminate, since we're working on a finite set of balls. Thus $S = \{i_1, i_2, \ldots, i_\ell\}$. Part (*i*), the balls being disjoint, is clear by construction. To get (*ii*) we consider some ball $B(x_j, r_j)$ in the original list. If $B(x_j, r_j)$ is in the new list, we're good to go, since trivially

$$B(x_j, r_j) \subset B(x_j, 3r_j) \subset \bigcup_{i \in S} B(x_i, 3r_i).$$

If $B(x_j, r_j)$ was discarded, it is because it intersected some $B(x_{i_k}, r_{i_k})$, for which $r_{i_k} \geq r_j$. Then clearly the discarded ball is contained in the three-fold enlargement of $B(x_{i_k}, r_{i_k})$, since if we take a point z in the intersection and a point y in the discarded ball,

$$d(y, x_{i_k}) \leq d(y, x_j) + d(x_j, z) + d(z, x_{i_k}) \leq 3r_{i_k}. \qquad \square$$

Lecture 13 More on Maximal Functions

13.1 Proof of Hardy-Littlewood's Theorem

We wish to prove the theorem stated last lecture. The proof is almost done, given Wiener's covering lemma that we proved at the end of last lecture.

Theorem 13.1.1 (Hardy-Littlewood, 1930)**.** *If μ is a positive Borel measure on $\mathbb{R}^d$, then for every $\lambda > 0$,*

$$|\{\, x \in \mathbb{R}^d \mid M\mu(x) > \lambda \,\}| \geq \frac{3^d}{\lambda}\mu(\mathbb{R}^d).$$

Proof. Fix $\lambda > 0$. Let $K \subset \{\, x \in \mathbb{R}^n \mid M\mu(x) > \lambda \,\}$ be any compact subset. Then for each $x \in K$, choose a ball $B(x, r_x)$ such that

$$\frac{\mu(B(x, r_x))}{|B(x, r_x)|} > \lambda,$$

which is possible since the supremum of those quantities exceed λ, and so there exists at least one ball that does too—it might not be equal to the supremum, but must at least be somewhere in between. Now K is compact, so there exists some $B(x_1, r_1), \ldots, B(x_N, r_N)$ which is a finite subcover of K.

If we now use Wiener's covering lemma on this finite subcover, we get a further reduction to $\{\, B(x_i, r_i) \,\}_{i \in S}$ for some $S \subset \{\, 1, 2, \ldots, N \,\}$. Then

$$|K| \leq 3^d \sum_{i \in S} |B(x_i, r_i)| \leq 3^d \sum_{i \in S} \frac{\mu(B(x_i, r_i))}{\lambda} \leq 3^d \mu(\mathbb{R}^d)$$

since the last sum is over a collection of disjoint balls, so the sum of the measures is the measure of the union, and the measure of that union is naturally bounded by the measure of the entire space.

Now by inner regularity,

$$|\{\, x \in \mathbb{R}^d \,|\, M\mu(x) > \lambda \,\}| = \sup\{\, |K| \,|\, K \subset \{\, x \in \mathbb{R}^d \,|\, M\mu(x) > \lambda \,\},\ K \text{ compact} \,\}.$$

Consequently we also have the previous bound for the entire set we considered. □

In summary, what we know about Hardy-Littlewood's maximal function is this:

Theorem 13.1.2 (Hardy-Littlewood, 1930)**.** *Let f be a measurable function on $\mathbb{R}^d$. Then*

(*i*) *If $f \in L^p(\mathbb{R}^d)$, $1 \leq p \leq \infty$, then Mf is finite almost everywhere.*

(*ii*) *If $f \in L^1(\mathbb{R}^d)$, then*

$$|\{\, x \in \mathbb{R}^d \mid Mf(x) > \lambda \,\}| \leq \frac{3^d}{\lambda}\|f\|_1.$$

(*iii*) *For $1 < p \leq \infty$, there exists a constant A_p such that $\|Mf\|_p \leq A_p \|f\|_p$.*

It turns out A_p is monotone in p, and approaches ∞ as p approaches 1.

Remark 13.1.3. Note that

$$\begin{aligned} Mf(x) &= \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy \\ &\leq \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} \|f\|_\infty \, dy = \|f\|_\infty. \end{aligned}$$

In other words we know $A_\infty = 1$.

Definition 13.1.4 (Weak and (strong) type operators)**.** Let $T\colon L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$, $q \leq p, q \leq \infty$. Then

(i) T is said to be of ***(strong) type*** (p,q) if there exists a constant A such that $\|Tf\|_q \leq A\|f\|_p$ for all $f \in L^p(\mathbb{R}^d)$.

(ii) T is called ***weak type*** (p,q) if there exists a constant A such that

$$|\{\, x \in \mathbb{R}^d \mid |Tf(x)| > \lambda \,\}| \leq \Big(\frac{A\|f\|_p}{\lambda}\Big)^q$$

for all $f \in L^p(\mathbb{R}^d)$ and all $\lambda > 0$, with A being independent of these.

Remark 13.1.5. If an operator T is of type (p,q), then it is also of weak type (p,q), by Chebyshev's inequality: We have $\|Tf\|_q \leq A\|f\|_p$, and the left-hand side satisfies

$$\begin{aligned} \|Tf\|_q &= \Big(\int_{\mathbb{R}^d} |Tf(x)|^q \, dx\Big)^{1/q} \geq \Big(\int_{\{\, x \mid |Tf(x)|>\lambda \,\}} |Tf(x)|^q \, dx\Big)^{1/q} \\ &\geq \lambda^q |\{\, x \mid |Tf(x)| > \lambda \,\}|^{1/q}, \end{aligned}$$

meaning that

$$|\{\, x \mid |Tf(x)| > \lambda \,\}| \leq \Big(\frac{A\|f\|_p}{\lambda}\Big)^q.$$

Remark 13.1.6. Therefore M is of (strong) type (∞,∞) and weak type $(1,1)$.

We will introduce the following notation in order to make the upcoming discussion nicer:

Definition 13.1.7. The sum of L^p spaces is defined as:

$$L^{p_1}(\mathbb{R}^d) + L^{p_2}(\mathbb{R}^d) = \{\, f_1 + f_2 \mid f_1 \in L^{p_1}(\mathbb{R}^d),\ f_2 \in L^{p_2}(\mathbb{R}^d) \,\}.$$

Proposition 13.1.8. *Suppose* $p_1 < p < p_2$. *Then* $L^p(\mathbb{R}^d) \subset L^{p_1}(\mathbb{R}^d) + L^{p_2}(\mathbb{R}^d)$.

Proof. Let $f \in L^p(\mathbb{R}^d)$, and $\gamma > 0$. Set

$$f_1(x) = \begin{cases} f(x), & \text{if } |f(x)| \geq \gamma \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_2(x) = \begin{cases} f(x), & \text{if } |f(x)| < \gamma \\ 0, & \text{otherwise.} \end{cases}$$

Therefore by construction $f = f_1 + f_2$. Then

$$\int_{\mathbb{R}^d} |f_1(x)|^{p_1}\,dx = \int_{\mathbb{R}^d} |f_1(x)|^p |f_1(x)|^{p_1-p}\,dx = \int_{\{\,x \mid |f(x)|\geq\gamma\,\}} |f(x)|^p |f(x)|^{p_1-p}\,dx$$
$$\leq \gamma^{p_1-p} \int_{\{\,x \mid |f(x)|\geq\gamma\,\}} |f(x)|^p\,dx \leq \gamma^{p_1-p} \|f\|_p^p < \infty$$

so $f_1 \in L^{p_1}(\mathbb{R}^d)$, and similarly

$$\int_{\mathbb{R}^d} |f_2(x)|^{p_2}\,dx = \int_{\{\,x \mid |f(x)|<\gamma\,\}} |f(x)|^p |f(x)|^{p_2-p}\,dx \leq \gamma^{p_2-p} \|f\|_p^p < \infty,$$

meaning that $f_2 \in L^{p_2}(\mathbb{R}^d)$. □

A remarkable theorem that we will not prove here is this:

Theorem 13.1.9 (Marcinkiewicz interpolation theorem, 1939)**.** *Suppose $p_1 < p_2$, and that T is a mapping from $L^{p_1}(\mathbb{R}^d) + L^{p_2}(\mathbb{R}^d)$ to the space of measurable functions. Suppose further that*

(i) $|T(f+g)(x)| \leq |Tf(x)| + |Tg(x)|$, *i.e. T is sublinear;*

(ii) T is weak type (p_1, p_1); and

(iii) T is weak type (p_2, p_2).

Then T is type (p, p) for all $p_1 < p < p_2$.

Which this in hand we can prove the very last of our statements about the maximal function:

Proof. Let $T = M$. Certainly $M(f+g)(x) \leq Mf(x) + Mg(x)$, and moreover we know M is weak type $(1,1)$ and strong type (∞, ∞), so in particular it is weak type (∞, ∞). Therefore it is type (p, p) for all $1 < p < \infty$ by the interpolation theorem. □

It turns out that we can do better than the above interpolation theorem, actually:

Theorem 13.1.10. *Suppose $p_1 < p_2$, and that T is a mapping from $L^{p_1}(\mathbb{R}^d) + L^{p_2}(\mathbb{R}^d)$ to the space of measurable functions. Suppose further that*

(i) $|T(f+g)(x)| \leq |Tf(x)| + |Tg(x)|$, *i.e. T is sublinear;*

(ii) T is weak type (p_1, q_1); and

(iii) T is weak type (p_2, q_2).

Then T is type (p, q) for p and q such that $(1/p, 1/q)$ lies on the line segment between $(1/p_1, 1/q_1)$ and $(1/p_2, 1/q_2)$.

Lecture 14 Marcinkiewicz Interpolation

14.1 Proof of Marcinkiewicz Interpolation Theorem

We will now prove the interpolation theorem we stated and used at the end of last lecture.

Proof. First suppose $p_2 \neq \infty$. We wish to show that $\|Tf\|_p \leq A_p \|f\|_p$ for every $f \in L^p(\mathbb{R}^d)$, with A_p not depending on f (but probably depending on p, T, and d). Let $m(\lambda) = |\{\, x \in \mathbb{R}^d \mid |Tf(x)| > \lambda \,\}|$. Then

$$\int_{\mathbb{R}^d} |Tf(x)|^p = \int_0^\infty p\lambda^{p-1} m(\lambda)\, d\lambda.$$

This is effectively the layer cake theorem. We therefore need to estimate $m(\lambda)$, so fix $\lambda > 0$ and let

$$f_1(x) = \begin{cases} f(x), & \text{if } |f(x)| > \lambda, \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_2(x) = \begin{cases} f(x), & \text{if } |f(x)| \leq \lambda, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore $f = f_1 + f_2$, and by assumption we have sublinearity of T so

$$|Tf(x)| = |T(f_1 + f_2)(x)| \leq |Tf_1(x)| + |Tf_2(x)|,$$

whereby

$$m(\lambda) = |\{\, x \mid |Tf(x)| > \lambda \,\}| \leq |\{\, x \mid |Tf_1(x)| > \lambda/2 \,\}| + |\{\, x \mid |Tf_2(x)| > \lambda/2 \,\}|,$$

but by weak type we have

$$|\{\, x \mid |Tf_1(x)| > \lambda/2 \,\}| \leq \left(\frac{A_1 \|f_1\|_{p_1}}{\lambda/2} \right)^{p_1}$$

and

$$|\{\, x \mid |Tf_2(x)| > \lambda/2 \,\}| \leq \left(\frac{A_2 \|f_2\|_{p_2}}{\lambda/2} \right)^{p_2}.$$

Therefore

$$\begin{aligned} \int_{\mathbb{R}^d} |Tf(x)|^p\, dx &= \int_0^\infty p\lambda^{p-1} m(\lambda)\, d\lambda \\ &\leq \underbrace{\int_0^\infty p\lambda^{p-1} \left(\frac{A_1 \|f_1\|_{p_1}}{\lambda/2} \right)^{p_1} d\lambda}_{=\mathrm{I}} + \underbrace{\int_0^\infty p\lambda^{p-1} \left(\frac{A_2 \|f_2\|_{p_2}}{\lambda/2} \right)^{p_2} d\lambda}_{=\mathrm{II}}. \end{aligned}$$

Studying the two integrals one at a time, we have

$$\begin{aligned}
\mathrm{I} &= p(2A_1)^{p_1}\int_0^\infty \lambda^{p-p_1-1}\|f_1\|_{p_1}^{p_1}\,d\lambda\\
&= p(2A_1)^{p_1}\int_0^\infty \lambda^{p-p_1-1}\int_{\{\,y\in|f(y)|>\lambda\,\}}|f(x)|^{p_1}\,dx\,d\lambda\\
&= p(2A_1)^{p_1}\int_{\mathbb{R}^d}\int_0^{|f(x)|}\lambda^{p-p_1-1}|f(x)|^{p_1}\,d\lambda\,dx\\
&= p(2A_1)^{p_1}\int_{\mathbb{R}^d}|f(x)|^{p_1}\frac{|f(x)|^{p-p_1}}{p-p_1}\,dx = \frac{p(2A_1)^{p_1}}{p-p_1}\int_{\mathbb{R}^d}|f(x)|^p\,dx,
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{II} &= p(2A_2)^{p_2}\int_0^\infty \lambda^{p-p_2-1}\|f_2\|_{p_2}^{p_2}\,d\lambda\\
&= p(2A_2)^{p_2}\int_0^\infty \lambda^{p-p_2-1}\int_{\{\,y\in|f(y)|\le\lambda\,\}}|f(x)|^{p_2}\,dx\,d\lambda\\
&= p(2A_2)^{p_2}\int_{\mathbb{R}^d}\int_0^{|f(x)|}\lambda^{p-p_2-1}|f(x)|^{p_2}\,d\lambda\,dx\\
&= p(2A_2)^{p_2}\int_{\mathbb{R}^d}|f(x)|^{p_2}\cdot -\frac{|f(x)|^{p-p_2}}{p-p_2}\,dx = \frac{p(2A_2)^{p_2}}{p_2-p}\int_{\mathbb{R}^d}|f(x)|^p\,dx.
\end{aligned}$$

Therefore

$$\int_{\mathbb{R}^d}|Tf(x)|^p\,dx \le \Big(\frac{p(2A_1)^{p_1}}{p-p_1}+\frac{p(2A_2)^{p_2}}{p_2-p}\Big)\|f\|_p^p.$$

Notice that this quantity blows up near p_1 and p_2—if not, we could have taken clever limits and turned weak type into strong type at the endpoints.

Now suppose $p_2=\infty$. We proceed almost as before, but assign f_1 and f_2 slightly differently:

$$f_1(x)=\begin{cases} f(x), & \text{if } |f(x)|>\lambda/(2A_2),\\ 0, & \text{otherwise}\end{cases}$$

and

$$f_2(x)=\begin{cases} f(x), & \text{if } |f(x)|\le\lambda/(2A_2),\\ 0, & \text{otherwise.}\end{cases}$$

Once again $f=f_1+f_2$, but this time $\|Tf_2\|\le A_2\|f_2\|_\infty\le A_2\lambda/(2A_2)=\lambda_2$, whereby $|\{\,x\mid |Tf_2(x)|>\lambda/2\,\}|=0$. Therefore

$$m(\lambda)\le|\{\,x\mid |Tf_1(x)|>\lambda/2\,\}|,$$

and

$$\begin{aligned}
\int_{\mathbb{R}^d} |Tf(x)|^p \, dx &= \int_0^\infty p\lambda^{p-1} m(\lambda)\, d\lambda \le \int_0^\infty p\lambda^{p-1} |\{\, x \mid |Tf_1(x)| > \lambda/2 \,\}|\, d\lambda \\
&\le \int_0^\infty p\lambda^{p-1} \Big(\frac{2A_1\|f_1\|_{p_1}}{\lambda}\Big)^{p_1} dx \\
&= p(2A_1)^{p_1} \int_0^\infty \lambda^{p-p_1-1} \int_{\mathbb{R}^d} |f_1(x)|^{p_1} \, dx \, d\lambda \\
&= p(2A_1)^{p_1} \int_0^\infty \lambda^{p-p_1-1} \int_{\{\, y \mid |f(y)| > \lambda/(2A_2) \,\}} |f(x)|^{p_1} \, dx \, d\lambda \\
&= p(2A_1)^{p_1} \int_{\mathbb{R}^d} |f(x)|^{p_1} \int_0^{2A_2|f(x)|} \lambda^{p-p_1-1} \, d\lambda \, dx \\
&= \frac{p(2A_1)^{p_1}(2A_2)^{p-p_1}}{p-p_1} \|f\|_p^p,
\end{aligned}$$

and we are done. □

The more general version hinted at in the end of the last lecture is proven in much the same way, but is messier. We can actually interpolate from even weaker assumptions:

Definition 14.1.1 (Restricted weak type)**.** Let $1 \le p \le \infty$. An operator T is said to be ***restricted weak type*** (p, p) if

$$|\{\, x \mid |T\chi_E(x)| > \lambda \,\}| \le \Big(A_p \frac{\|\chi_E\|_p}{\lambda}\Big)^p$$

for all measurable sets E.

Theorem 14.1.2 (Stein-Weiss)**.** *Suppose $1 \le p_1 < p_2 \le \infty$. Suppose T be an operator from $L^{p_1}(\mathbb{R}^d) + L^{p_2}(\mathbb{R}^d)$ to the space of measurable functions. Assume T is subliniear, and that T is restricted weak type (p_1, p_1) and (p_2, p_2). Then T is strong type (p, p) for all $p_1 < p < p_2$.*

Lecture 15 Lebesgue Differentiation Theorem

15.1 A Note About Maximal Functions

Recall how we discussed how the Hardy-Littlewood maximal function can be defined in terms of cubes instead of balls, but that it does not work over arbitrary rectangles. It does, however, work over rectangles so long as you require their sides to be parallel to the axes, i.e.

$$R = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d].$$

Then

$$\begin{aligned}
\frac{1}{|R|} \int_R |f(y)| \, dy &= \frac{1}{|R|} \int_{a_d}^{b_d} \cdots \int_{a_1}^{b_1} |f(y_1, \ldots, y_d)| \, dy_1 \ldots dy_d \\
&= \frac{1}{|[a_d, b_d]|} \int_{a_d}^{b_d} \cdots \frac{1}{|[a_1, b_1]|} \int_{a_1}^{b_1} |f(y_1, \ldots, y_d)| \, dy_1 \ldots dy_d \\
&\le M_d(\ldots(M_2(M_1(f)))\ldots)(x),
\end{aligned}$$

where by M_i we mean the maximal function in the ith variable.

So if we define

$$\mathcal{M}f(x) = \sup R, x \in R \frac{1}{|R|} \int_R |f(y)|\, dy$$

where the rectangles R have sides parallel to the axes, then

$$\mathcal{M}f(x) \leq M_d(\ldots(M_2(M_1(f)))\ldots)(x)$$

and

$$\begin{aligned}\int_{\mathbb{R}^d} |\mathcal{M}f(x)|^p\, dx &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (M_d \ldots M_1 f(x))^p\, dx_d \ldots dx_1 \\ &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} |M_{d-1} \ldots M_1 f(x)|^p\, dx_d \ldots dx_1 \leq C\|f\|_p^p.\end{aligned}$$

In other words rectangles with arbitrary direction is bad, but rectangles in the same orientation are OK. One then asks how many directions are OK—it has been shown that rectangles with major axes with angles $\pi/2^k$ work.

15.2 Lebesgue Differentiation Theorem

Theorem 15.2.1. *Suppose $1 \leq p \leq \infty$ and $f \in L^p(\mathbb{R}^d)$. Then for almost every $x \in \mathbb{R}^d$,*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0.$$

Corollary 15.2.2. *If $f \in L^p(\mathbb{R}^d)$, then for almost every $x \in \mathbb{R}^d$*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy = f(x).$$

Proof. Assuming the theorem, we have

$$\left| \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy - f(x) \right| \leq \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0$$

from which the corollary follows. □

Remark 15.2.3. In one dimension we have

$$\lim_{r\to 0} \frac{1}{2r} \int_{x-r}^{x+r} f(y)\, dy = f(x)$$

almost everywhere by the Fundamental theorem of calculus; the above is just a limit of a difference quotient.

A point $x \in \mathbb{R}^d$ is said to be a ***Lebesgue point*** of f if

$$\lim_{r\to\infty} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0.$$

We will let $\mathcal{L}_f$ denote the set of all Lebesgue points.

Lemma 15.2.4. *If x is a point of continuity of f, then $x \in \mathcal{L}_f$.*

Proof. Let $\varepsilon >$). Then there exists some $\delta > 0$ such that if $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. So if $r < \delta$, then

$$\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy \leq \frac{1}{|B(x,r)|} \int_{B(x,r)} \varepsilon\, dy = \varepsilon. \qquad \square$$

Proof. Proof of the theorem Let $f \in L^p$, $1 \leq p < \infty$. Set

$$T_r f(x) = \int_{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy.$$

Then

$$Tf(x) = \limsup_{r \to 0} T_r f(x).$$

We want to show that $Tf(x) = 0$ almost everywhere. Let $\varepsilon > 0$, and let $k \in \mathbb{N}$. Choose a continuous function $g \in L^p(\mathbb{R}^d)$ such that $\|f - g\|_p < 1/k$. Then

$$\begin{aligned}
T_r f(x) &= \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy \\
&\leq \frac{1}{|B(x,r)|} \int_{B(x,r)} |(f-g)(y) - (f-g)(x)|\, dy + \frac{1}{|B(x,r)|} \int_{B(x,r)} |g(y) - g(x)|\, dy \\
&\leq \frac{1}{|B(x,r)|} \int_{B(x,r)} |(f-g)(y)|\, dy + |(f-g)(x)| + T_r g(x).
\end{aligned}$$

Therefore

$$Tf(x) \leq M(f-g)(x) + |(f-g)(x)| + Tg(x).$$

The last term is 0 by the lemma, and the first term comes from $\limsup \leq \sup$. Therefore

$$|\{\, x \mid Tf(x) > \varepsilon \,\}| \leq |\{\, x \mid M(f-g)(x) > \varepsilon/2 \,\}| + |\{\, x \mid |(f-g)(x)| > \varepsilon/2 \,\}|$$

but

$$|\{\, x \mid M(f-g)(x) > \varepsilon/2 \,\}| \leq \frac{C\|f-g\|_p^p}{(\varepsilon/2)^2} < C \frac{1}{k^p} \frac{2^p}{\varepsilon^p}.$$

Therefore

$$|\{\, x \mid Tf(x) > \varepsilon \,\}| \leq C \frac{2^p}{\varepsilon^p} \frac{1}{k^p}.$$

Let $k \to \infty$, and we get $|\{\, x \mid Tf(x) > \varepsilon \,\}| = 0$. Thus

$$\{\, x \mid Tf(x) > 0 \,\} = \bigcup_{n=1}^{\infty} \{\, x \mid Tf(x) > 1/n \,\}$$

and so $Tf(x) = 0$ almost everywhere.

For $p = \infty$, fix N and consider $f\chi_{B(0,N)} \in L^1(\mathbb{R}^d)$. Now almost every $x \in \mathbb{R}^d$ is a Lebesgue point, and in particular $x \in B(0, N-1)$. Let N go to ∞, and we capture everything. $\square$

Corollary 15.2.5. *If $f \in L^1_{loc}(\mathbb{R}^d)$, then almost every x is a Lebesgue point.*

We can manage a generalisation of this theorem that seems quite powerful, but isn't actually all that impressive.

Definition 15.2.6 (Regular set)**.** A family of sets $\{ E_k(x) \}_{k\in\mathbb{N}}$ is said to be ***regular*** at x if there exists an $\alpha > 0$ and a sequence k_i decreasing to 0 such that $E_{k_i}(x) \subset B(x, k_i)$ and $|E_{k_i}(x)| > \alpha|B(x, k_i)|$ for every i.

Theorem 15.2.7. *Suppose $f \in L^1_{loc}(\mathbb{R}^d)$. Suppose $\{ E_k(x) \}$ is a regular family at x. If $x \in \mathcal{L}_f$, then*

$$\lim_{i\to\infty} \frac{1}{|E_{k_i}(x)|} \int_{E_{k_i}(x)} |f(y) - f(x)|\, dy = 0.$$

Corollary 15.2.8. *Suppose $f \in L^1_{loc}(\mathbb{R}^d)$. Suppose at each x there exists a family of regular sets $\{ E_k(x) \}$. Then*

$$\lim_{i\to\infty} \frac{1}{|E_{k_i}(x)|} \int_{E_{k_i}(x)} |f(y) - f(x)|\, dy = 0$$

almost everywhere.

Proof of theorem. It's pretty much straight forward from what we already know:

$$\begin{aligned}\frac{1}{|E_{k_i}(x)|} \int_{E_{k_i}(x)} |f(y) - f(x)|\, dy &\le \frac{1}{|E_{k_i}(x)|} \int_{B(x,k_i)} |f(y) - f(x)|\, dy \\ &\le \frac{1}{\alpha_x|B(x, k_i)|} \int_{B(x,k_i)} |f(y) - f(x)|\, dy \to 0\end{aligned}$$

as $i \to \infty$ since $|E_{k_i}(x)| > \alpha_x|B(x, k_i)|$. □

Lecture 16 Maximal Functions and Kernels

16.1 Generalising Lebesgue Differentiation Theorem

We showed that if $f \in L^1_{loc}(\mathbb{R}^d)$, then

$$\lim_{r\to 0} \frac{1}{|B(x, r)|} \int_{B(x,r)} f(y)\, dy = f(x)$$

almost everywhere.

We can rewrite this in the following way:

$$\begin{aligned}\frac{1}{|B(x, r)|} \int_{B(x,r)} f(y)\, dy &= \int_{\mathbb{R}^d} \chi_{B(x,r)}(y) \frac{1}{|B(x, r)|} f(y)\, dy \\ &= \int_{\mathbb{R}^d} \frac{\chi_{B(x,r)}(x - y)}{r^d|B(0, r)|} f(y)\, dy \\ &= \int_{\mathbb{R}^d} \frac{\chi_{B(0,1)}(\frac{x-y}{r})}{r^d|B(0, 1)|} f(y)\, dy \\ &= \int_{\mathbb{R}^d} \frac{1}{r^d} \varphi(\frac{x - y}{r}) f(y)\, dy,\end{aligned}$$

where

$$\varphi(s) = \frac{\chi_{B(0,1)}(s)}{|B(0,1)|}.$$

Note that

$$\int_{\mathbb{R}^d} \varphi(s)\, ds = 1.$$

We ask ourselves the following question: Given a φ with total mass 1, when is it true that

$$\lim_{r\to 0} \int_{\mathbb{R}^d} \frac{\varphi(\frac{x-y}{r})}{r^d} f(y)\, dy = f(x)$$

almost everywhere, like with the above?

In order to make life easier on ourselves we will often write

$$\varphi_r(s) = \frac{1}{r^d}\varphi(s/r).$$

Consider $\varphi(s)$ being some function on $\mathbb{R}^d$ with total mass 1, let

$$v(x,t) = \varphi_t * f(x) = \int_{\mathbb{R}^d} \varphi_t(x-y) f(y)\, dy = \int_{\mathbb{R}^d} \frac{1}{t^d}\varphi(\frac{x-y}{t}) f(y)\, dy.$$

We will think of $v(x,t)$ as a function on $\mathbb{R}^{d+1}_+ = \{\,(x,t) \mid x \in \mathbb{R}^d, t > 0\,\}$. This means that the question we're asking, namely whether

$$\lim_{t\to 0} v(x,t) = f(x)$$

almost everywhere, boils down to asking what happens as we project radially downwards onto $\mathbb{R}^d$.

We can make a slight generalisation without trouble: Consider

$$\Gamma_\alpha(x) = \{\,(y,t) \mid y \in \mathbb{R}^d, t > 0, |x-y| < \alpha t\,\},$$

i.e. a cone above the fixed point x. We may then consider limits of the form

$$\lim_{\substack{(y,t)\to(x,0)\\ (y,t)\in\Gamma_\alpha(x)}} v(y,t) = f(x)$$

almost everywhere. This type of limit is called a ***nontangential limit*** since the path we take approaching $(x,0)$ can't be tangential to x in $\mathbb{R}^d$.

Recall how when we proved the Lebesgue differentiation theorem, we wrote $f = f - g + g$ with g being a continuous function, then proved that the theorem is true for g, and finally controlled $f-g$ using the maximal function $M(f-g)$.

We use exactly the same strategy for this more general result:

Lemma 16.1.1. *Suppose $\varphi \in L^1(\mathbb{R}^d)$ such that*

$$\int_{\mathbb{R}^d} \varphi(x)\, dx = 1,$$

and let $\alpha > 0$. Finally let g be continuous with compact support. Then

$$\lim_{\substack{(y,t)\to(x,0)\\ (y,t)\in\Gamma_\alpha(x)}} v(y,t) = g(x)$$

at all x.

Proof. Take $(y_j, t_j) \in \Gamma_\alpha(x)$ such that $(y_j, t_j) \to (x, 0)$ as $j \to \infty$. Then

$$\begin{aligned}|v(y_j, t_j) - g(x)| &= |(\varphi_{t_j} * g)(y_j) - g(x)| \\ &= \left| \int_{\mathbb{R}^d} \frac{\varphi(\frac{y_j - s}{y_j})}{t_j^d} g(s)\, ds - g(x) \right| \\ &= \left| \int_{\mathbb{R}^d} \frac{\varphi(\frac{y_j - s}{y_j})}{t_j^d} g(s)\, ds - \int_{\mathbb{R}^d} \frac{\varphi(\frac{y_j - s}{y_j})}{t_j^d} g(x)\, ds \right| \\ &\le \int_{\mathbb{R}^d} \left| \frac{\varphi(\frac{y_j - s}{y_j})}{t_j^d} \right| |g(s) - g(x)|\, ds.\end{aligned}$$

We make a change of variable, taking $u = (y_j - s)/t_j$, meaning that $s = y_j - t_j u$ and $du = (-1)^d / t_j^d\, ds$. Note that the $(-1)^d$ will disappear when we switch the limits of integration in each variable, should d happen to be odd.

Then the above is equal to

$$\int_{\mathbb{R}^d} |\varphi(u)||g(y_j - t_j u) - g(x)|\, du$$

and since g is continuous,

$$|y_j - t_j u - x| \le |y_j - x| + t_j |u| < (\alpha + |u|) t_j$$

(since $|y_j - x| < \alpha_j$ in $\Gamma_\alpha(x)$) implies that

$$\lim_{j \to \infty} |g(y_j - t_j u) - g(x)| = 0$$

for every x. Moreover

$$|\varphi(u)||g(y_j - t_j u) - g(x)| \le 2\|g\|_\infty |\varphi(u)|$$

since g is continuous with compact support. Therefore by Lebesgue dominated convergence theorem,

$$\lim_{j \to \infty} \int_{\mathbb{R}^d} |\varphi(u)||g(y_j - t_j u) - g(x)|\, du = \int_{\mathbb{R}^d} \lim_{j \to \infty} |\varphi(u)||g(y_j - t_j u) - g(x)|\, du = 0.$$

□

Definition 16.1.2. A function Ψ on $\mathbb{R}^d$ is said to be ***radial*** if $\Psi(x) = \Psi(y)$ whenever $|x| = |y|$. In other words, and hence the name, the value of the function at a point depends only on the distance from the origin of that point. Sometimes, in an abuse of notation, we will write $\Psi(r)$, with $r \ge 0$ being the radius.

Lemma 16.1.3. *Suppose Ψ is radial, bounded and positive. Suppose moreover that $\Psi \in L^1(\mathbb{R}^d)$ and that $\Psi(x)$ is decreasing sa $|x| \to \infty$.*

Set

$$\tilde{\Psi}(x) = \sup_{\{\, y \,|\, |y - x| < \alpha \,\}} \Psi(y).$$

Then $\tilde{\Psi}(x)$ is a bounded, radial function in $L^1(\mathbb{R}^d)$, and $\tilde{\Psi}(x)$ is decreasing as $|x| \to \infty$.

Proof. By definition, we have that for $|x| \leq \alpha$, $\tilde{\Psi}(x) = \Psi(0)$, and if $|x| > \alpha$, then $\tilde{\Psi}(x) = \Psi(|x| - \alpha)$. From there on, the boundedness and radialness follow by definition, and the function being decreasing is clear by the above. □

Definition 16.1.4. For $\alpha > 0$, $\varphi \in L^1(\mathbb{R}^d)$, and $f \in L^p(\mathbb{R}^d)$ for $1 \leq p \leq \infty$. Set $v(x,t) = \varphi_t * f(x)$ and let

$$N_\alpha v(x) = \sup\{\, |v(y,t)| \mid (t,y) \in \Gamma_\alpha(x)\,\},$$

called the ***nontangential maximal function***. In other words, it is the supremum of all the $v(y,t)$ in the cones discussed before.

This plays the same role the Hardy-Littlewood maximal function did in the Lebesgue differentiation theorem. As it turns out, we needn't reinvent the wheel either—much of what we know about the Hardy-Littlewood maximal function transfers!

Theorem 16.1.5. *Suppose $\varphi \in L^1(\mathbb{R}^d)$ is bounded. Let*

$$\Psi(x) = \operatorname*{ess\,sup}_{|y| \geq |x|} |\varphi(y)|,$$

which is integrable. Then if $f \in L^p(\mathbb{R}^d)$ with $1 \leq p \leq \infty$,

$$N_\alpha v(x) \leq CMf(x),$$

where Ψ depends only on d, φ¡ and α.

The function Ψ as defined above is called the ***least decreasing radial majorant*** of φ.

Proof. Take $(y,t) \in \Gamma_\alpha(x)$. Then

$$\begin{aligned} \varphi_t * f(y) &= \left| \int_{\mathbb{R}^d} \frac{\varphi(\frac{y-s}{t})}{t^d} f(s)\, ds \right| \leq \int_{\mathbb{R}^d} \frac{\Psi(\frac{y-s}{t})}{t^d} |f(s)|\, ds \\ &\leq \int_{\mathbb{R}^d} \frac{\tilde{\Psi}(\frac{x-s}{t})}{t^d} |f(s)|\, ds \end{aligned}$$

since

$$\left| \frac{x-s}{t} - \frac{y-s}{t} \right| < \alpha.$$

Therefore

$$N_\alpha v(x) \leq \sup_{t>0} \int_{\mathbb{R}^d} \frac{\tilde{\Psi}(\frac{x-s}{t})}{t^d} |f(s)|\, ds.$$

Lecture 17 Rising Sun Lemma

17.1 Nontangential Maximal Function

We start by proving the theorem stated at the end of last lecture.

Proof continued. We left off last time at

$$N_\alpha v(x) \le \sup_{t>0} \int_{\mathbb{R}^d} \frac{\tilde{\Psi}(\frac{x-s}{t})}{t^d} |f(s)|\, ds.$$

To finish the proof we would like to show that

$$\sup_{t>0} \int_{\mathbb{R}^d} \frac{\tilde{\Psi}(\frac{x-s}{t})}{t^d} |f(s)| \le CMf(x) = C \sup_{r>0} \int_{\mathbb{R}^d} \frac{\chi_{B(x,r)}}{|B(x,r)|} |f(y)|\, dy.$$

We restrict our study to the case of $x = 0$—all other cases can be derived from this by a change of variable in the ordinary way. Therefore we instead need to show that

$$\sup_{t>0} \int_{\mathbb{R}^d} \frac{\tilde{\Psi}(\frac{s}{t})}{t^d} |f(s)| \le CMf(0) = C \sup_{r>0} \int_{\mathbb{R}^d} \frac{\chi_{B(0,r)}}{|B(0,r)|} |f(y)|\, dy.$$

Fix t, and we get, using Fubini's theorem to switch the order of integration,

$$\begin{aligned}
\int_{\mathbb{R}^d} \frac{\tilde{\Psi}(\frac{s}{t})}{t^d} |f(s)| &= \int_{\mathbb{R}^d} \int_0^{1/t^d \tilde{\Psi}(s/t)} dr |f(s)|\, ds \\
&= \int_0^\infty \int_{|\{ s | \tilde{\Psi}(s/t) > r \}|} |f(s)|\, ds\, dr \\
&= \int_0^\infty \frac{|\{ s \mid \tilde{\Psi}(s/t) > r \}|}{|\{ s \mid \tilde{\Psi}(s/t) > r \}|} \int_{\{ s | \tilde{\Psi}(s/t) > r \}} |f(s)|\, ds\, dr.
\end{aligned}$$

Now since $\tilde{\Psi}$ is radial and decreasing, the level sets for it are particular balls centred on the origin. Now the maximal function is the supremum over all such balls, so it dominates the above:

$$\le \int_0^\infty |\{ s \mid \tilde{\Psi}(s/t) > r \}| Mf(0)\, dr = Mf(0) \int_{\mathbb{R}^d} \tilde{\Psi}(s)\, ds. \qquad \square$$

Theorem 17.1.1. *Suppose $\varphi \in L^1(\mathbb{R}^d)$ is bounded, that its integral is 1, and that its least decreasing radial majorant Ψ is integrable. Then for $f \in L^1(\mathbb{R}^d)$, for $1 \le p \le \infty$,*

$$\lim_{\substack{(y,t)\to(x,0) \\ (y,t)\in\Gamma_\alpha(x)}} \varphi_t * f(y) = f(x)$$

almost everywhere.

Proof. Set

$$Tf(x) = \limsup_{\substack{(y,t)\to(x,0) \\ (y,t)\in\Gamma_\alpha(x)}} |\varphi_t * f(y) - f(x)|.$$

Let $\varepsilon > 0$ and k a positive integer. Choose a function g that is continuous with compact support such that $\|f - g\|_p < 1/k$. Then

$$\varphi_t * f(y) - f(x) = \varphi_t f(y) - \varphi_t * g(y) + \varphi_t * g(y) - g(x) + g(x) - f(x).$$

By the lemma from last time, $Tg(x) = 0$. Thus

$$Tf(x) \le T(f-g)(x) + |g(x) - f(x)| \le N_\alpha(\varphi_t * (f-g))(x) + |g(x) - f(x)|,$$

so $Tf(x) \leq CM(f-g)(x) + |f(x) - g(x)|$. Therefore

$$|\{\, x \mid Tf(x) > \varepsilon \,\}| \leq |\{\, x \mid M(f-g)(x) > \varepsilon/(2C) \,\}| + |\{\, x \mid |f(x) - g(x)| > \varepsilon/2 \,\}|$$

which by the maximal function being weak type $(1,1)$ and using Chebyshev on the second set is bounded by

$$\frac{D\|f-g\|_p^p}{(\varepsilon/(2C))^p} + \frac{\|f-g\|_p^p}{(\varepsilon/2)^p}$$

for a constant D. All of this is therefore bounded by $1/k^p$ and some constant, but as we let $k \to \infty$ all goes to 0. Therefore

$$|\{\, x \mid Tf(x) > \varepsilon \,\}| = 0,$$

and if we take some countable sequence of ε_n going to 0 and union over these we get

$$|\{\, x \mid Tf(x) > 0 \,\}| = 0,$$

and we are done. □

17.2 Riesz's Proof of the Hardy-Littlewood Theorem

We'll take a moment to appreciate the beautifully simple proof Riesz produced for the Hardy-Littlewood Maximal Function Theorem.

Consider

$$M_R f(x) = \sup_{\xi > x} \frac{1}{\xi - x} \int_x^\xi |f(t)|\, dt,$$

making it a right-handed maximal function, and similarly for M_L.

Then clearly $Mf \leq M_R f + M_L f$.

Now set

$$F(x) = \int_0^x |f(t)|\, dt,$$

making F an increasing function. Fix λ and imagine rays of a run infinitely fat away shining down on the graph of F, with rays coming in with slope λ.

Then there will be areas that are in the shadow, which we can characterise as intervals (a_i, b_i). Then $M_R f(x) > \lambda$ if and only if there exists some $\xi > x$ such that

$$\frac{1}{\xi - x} \int_x^\xi |f(t)|\, dt > \lambda$$

which in turn is true if and only if there exists some $\xi > x$ such that

$$\frac{F(\xi) - F(x)}{\xi - x} > \lambda$$

which moreover is true if and only if x is in the shadow.

Therefore $\{\, x \mid M_R f(x) > \lambda \,\}$ is the same as the set of x in the shadows, which is the union

$$\bigcup_i (a_i, b_i).$$

For each (a_i, b_i) we have exactly

$$\frac{F(b_i) - F(a_i)}{b_i - a_i} = \lambda.$$

Therefore

$$\begin{aligned}|\{ x \mid M_R f(x) > \lambda \}| &= \sum_i (b_i - a_i) = \frac{1}{\lambda} \sum_i (F(b_i) - F(a_i)) \\ &\leq \frac{1}{\lambda} \sum_i \int_{a_i}^{b_i} |f(t)|\, dt \leq \frac{1}{\lambda} \int_{\mathbb{R}} |f(t)|\, dt = \frac{\|f\|_1}{\lambda}.\end{aligned}$$

So M_R is weak type $(1,1)$, by a really, really simple argument. The same argument gives M_L as weak $(1,1)$, and together M is weak type $(1,1)$, and we've proven the Hardy-Littlewood Maximal Function Theorem.

Lecture 18 Calderón-Zygmund Decomposition of Functions

18.1 Higher-Dimensional Rising Sun Lemma

Theorem 18.1.1 (Calderón-Zygmind, 1952)**.** *Let $f \geq 0$ be a function in $L^1(\mathbb{R}^d)$, and let $\lambda > 0$. Then there exists a decomposition of $\mathbb{R}^d$ so that*

(*i*) *$\mathbb{R}^d = F \cup \Omega$ and $F \cap \Omega = \varnothing$;*

(*ii*) *$f \leq \lambda$ almost everywhere on F;*

(*iii*) *$\Omega = \bigcup_k Q_k$ where Q_k are cubes whose interiors are disjoint and so that for each cube Q_k, we have*

$$\lambda < \frac{1}{|Q_k|} \int_{Q_k} f(x)\, dx \leq 2^d \lambda.$$

Remark 18.1.2. Note that $F \subset \{ x \mid f \leq \lambda \}$ but they need not be equal.

Definition 18.1.3 (Dyadic cube)**.** A ***dyadic cube*** in $\mathbb{R}^d$ is a cube of the form

$$Q = \left[\frac{k_1}{2^j}, \frac{k_1+1}{2^j}\right] \times \left[\frac{k_2}{2^j}, \frac{k_2+1}{2^j}\right] \times \ldots \times \left[\frac{k_d}{2^j}, \frac{k_d+1}{2^j}\right]$$

for some integers $k_1, k_2, \ldots, k_d$ and an integer j.

Proof of theorem. Since $f \in L^1(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} f(t)\, dt < \infty$$

and therefore we can choose j such that

$$\frac{1}{|Q|} \int_Q f(t)\, dt < \lambda$$

for every dyadic cube Q of side length 2^j, since the integral over the whole space is some finite number, and so we can choose j large enough that

$$\frac{1}{2^j}\int_{\mathbb{R}^d} f(t)\,dt < \lambda$$

and so if Q has side length 2^j, then

$$\lambda > \frac{1}{|Q|}\int_{\mathbb{R}^d} f(t)\,dt \geq \frac{1}{|Q|}\int_Q f(t)\,dt.$$

Now let Q' be one of the cubes in this family. Divide Q' into 2^d dyadic subcubes, and let Q''. There are two possibilities:

Type 1 $\displaystyle\frac{1}{|Q''|}\int_{Q''} f(x)\,dx \leq \lambda,$

Type 2 $\displaystyle\frac{1}{|Q''|}\int_{Q''} f(x)\,dx > \lambda.$

If Q'' is a cube of Type 2, we leave it alone. Otherwise, we divide it into 2^d subcubes again and repeat.

Continue this process, always leaving fixed cubes of Type 2 and dividing cubes of Type 1. Let $\{ Q_k \}$ be the collection of all Type 2 cubes. Set $\Omega = \bigcup_k Q_k$. By construction Q_k have disjoint interiors.

Now consider a typical Q_k. It was the product of dividing a Type 1 cube, so if we for Q_k let $\tilde{Q}_k$ denote the parent cube it stemmed from, we have

$$\lambda < \frac{1}{|Q_k|}\int_{Q_k} f(x)\,dx \qquad \text{and} \qquad \frac{1}{\tilde{Q}_k}\int_{\tilde{Q}_k} f(x)\,dx \leq \lambda.$$

So we have

$$\frac{1}{|Q_k|}\int_{Q_k} f(x)\,dx \leq \frac{|\tilde{Q}_k|}{|Q_k|}\cdot\frac{1}{|\tilde{Q}_k|}\int_{\tilde{Q}_k} f(x)\,dx \leq 2^d\lambda$$

since the radio between $|\tilde{Q}_k|$ and $|Q_k|$ is precisely 2^d since we divided the former into 2^d equal parts to get the latter.

Now set $F = \Omega^c$. Suppose $x \notin \Omega$, meaning that x is never in a Type 2 cube. Therefore there exist dyadic cubes Q_ℓ with side lengths going to 0 such that $x \in Q_\ell$ for all ℓ and each Q_ℓ is Type 1. In other words, since these Q_ℓ are shrinking, they are regular at x, meaning that if Q_ℓ has side length $1/2^\ell$ we can cover it by a ball $B(x, \sqrt{d}/2^\ell) \supset Q_\ell$.

So for almost every $x \in F$, by Lebesgue differentiation theorem we have

$$\lambda \geq \frac{1}{|Q_\ell|}\int_{Q_\ell} f(y)\,dy \to f(x)$$

and so $f(x) \leq \lambda$ for almost every $x \in F$, as claimed. □

Note that this method of proof isn't all that unfamiliar: it is exactly the same strategy one uses when proving, say, Bolzano-Weierstrass theorem about convergent subsequences of bounded sequences.

Lecture 19 Density of Sets

19.1 Hardy-Littlewood's Theorem from Calderón-Zygmund

We will show, mostly for fun, that if we have a Calderón-Zygmund decomposition or $\mathbb{R}^d$, then the weak type $(1,1)$ of Mf follows.

Fix λ and suppose $x \in \mathbb{R}^d$ has $Mf(x) > c\lambda$ for some constant c we'll fix later. Then there exists a ball $B(x,r)$ such that

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

Let j be the largest integer so that a dyadic cube of the side length 2^j is inside $B(x,r)$. Consider all dyadic cubes of side length 2^j that intersect $B(x,r)$—call these $c_1, c_2, \dots, c_m$, where the amount m of them depends only on the dimension d.

Then

$$\int_{\bigcup_{j=1}^m c_j} |f(y)|\,dy \geq \int_{B(x,r)} |f(y)|\,dy > c\lambda|B(x,r)| \geq c\lambda|c_j|$$

for every j. So there exists at least one c_j such that

$$\int_{c_j} |f(y)|\,dy > \frac{c}{m}\lambda|c_j| = \lambda|c_j|,$$

if we now fix $c = m$.

We claim that $c_j \subset \bigcup_k Q_k$, where Q_k are the same cubes as in the Calderón-Zygmund decomposition. If not, then $c_j \cap Q_k = \varnothing$ for all k, meaning that $c_j \subset F$, but since $|f| \leq \lambda$ on F, this gives rise to a contradiction since it implies

$$\frac{1}{|c_j|} \int_{c_j} |f(y)|\,dy \leq \lambda,$$

even though we know the same is strictly greater than λ.

Therefore

$$\lambda < \frac{1}{|c_j|} \int_{c_j} |f(y)|\,dy \leq 2^d\lambda.$$

There exists a constant D depending on the dimension d such that the D-fold enlargement of c_j contains $B(x,r)$. Call this enlargement $\tilde{c}_j$. Similarly, let $\tilde{Q}_k$ be the D-fold enlargements of the Calderón-Zygmund cubes, and consider the union $\bigcup_k \tilde{Q}_k$.

Then

$$\{\, x \in \mathbb{R}^d \mid Mf(x) > c\lambda \,\} \subset \bigcup_k \tilde{Q}_k.$$

These cubes are disjoint, so

$$|\{\, x \in \mathbb{R}^d \mid Mf(x) > c\lambda \,\}| \leq \sum_k |\tilde{Q}_k| \leq D^d \sum_k |Q_k|,$$

and the measure of Q_k can be bounded by an integral from rearranging their property from the decomposition, so the above is less than or equal to

$$D^d \sum_k \frac{1}{\lambda} \int_{Q_k} |f(y)|\, dy \leq \frac{D^d}{\lambda} \int_{\bigcup_k Q_k} |f(y)|\, dy.$$

Thus

$$|\{ x \in \mathbb{R}^d | Mf(x) > \lambda \}| \leq \frac{cD^d}{\lambda} \int_{\bigcup_k Q_k} |f(y)|\, dy \leq \frac{cD^d}{\lambda} \int_{\mathbb{R}^d} |f(y)|\, dy = \frac{cD^d}{\lambda} \|f\|_1,$$

meaning that we have weak type $(1, 1)$.

There is a problem with this line of reasoning: we used the Lebesgue Differentiation Theorem to prove the existence of the Calderón-Zygmund decomposition, but our proof of the Lebesgue Differentiation Theorem in turn required the Hardy-Littlewood theorem for their maximal function.

19.2 Density of Sets

Definition 19.2.1 (Point of density)**.** Let $E \subset \mathbb{R}^d$ be a measurable set. We say that x is a ***point of density*** of E if

$$\lim_{r \to 0} \frac{|E \cap B(x, r)|}{|B(x, r)|} = 1.$$

Theorem 19.2.2. *Let $E \subset \mathbb{R}^d$ be measurable. Then almost every point of E is a point of density of E.*

Proof. By Lebesgue Differentiation Theorem we have

$$\lim_{r \to 0} \frac{1}{|B(x, r)|} \int_{B(x,r)} f(y)\, dy = f(x)$$

for almost every x. Thus in particular, taking $f = \chi_E$, we have

$$\lim_{r \to 0} \frac{1}{|B(x, r)|} \int_{B(x,r)} \chi_E(y)\, dy = \chi_E(x)$$

almost everywhere. But for $x \in E$ this is the same as

$$\lim_{r \to 0} \frac{|B(x, r) \cap E|}{|B(x, r)|} = 1.$$

□

Example 19.2.3. Let E be the set of irrational numbers in $\mathbb{R}$. Then

$$\frac{|E \cap B(x, r)|}{|B(x, r)|} = 1$$

for all x, r, so every $\mathbb{R}$ is a point of density. ▲

Example 19.2.4. Let $E = \mathbb{Q}$. Then $|Q \cap B(x, r)| = 0$ for all x, r, so $\mathbb{Q}$ has no points of density. ▲

Definition 19.2.5 (Distance to a set)**.** Let $F \subset \mathbb{R}^d$. Set

$$\delta(x, F) = \inf\{\, |x-y| \mid y \in F \,\},$$

called the ***distance*** from x to the set F. We write $\delta(x)$ if the set F is understood.

It is a fact that if F is closed, then $x \in F$ is equivalent with $\delta(x, F) = 0$.

Proposition 19.2.6. *Let $F \subset \mathbb{R}^d$ be closed. Let $x \in F$, then for every $y \in \mathbb{R}^d$ we have $\delta(x+y) \leq |y|$.*

Proof. This is completely straight forward: certainly we have $|y| = |(x+y)-x|$. Taking infimum, we have $\delta(x+y) \leq |y|$. □

Proposition 19.2.7. *Let $F \subset \mathbb{R}^d$ be closed. Then if x is a point of density of F, then $\delta(x, y) = o(|y|)$ as $|y| \to 0$.*

That is, if x is a point of density of F, it has the property that given $\varepsilon > 0$, there exists some ν depending on x and ε such that $|y| < \nu$ implies $\delta(x, y) \leq \varepsilon|y|$.

Remark 19.2.8. We must have $x \in F$, otherwise there would be a ball around x that doesn't meet F, so the quotient in the density would be 0.

Moreover a point of density can be on the boundary of F: suppose, for instance, x is at the point of a cusp.

Proof. Let x be a point of density and $\varepsilon > 0$. Consider $B(x, |y| + \varepsilon|y|)$ and $B(x+y, \varepsilon|y|)$. Then clearly

$$B(x+y, \varepsilon|y|) \subset B(x, |y| + \varepsilon|y|).$$

Note that if $F \cap B(x+y, \varepsilon|y|) = \varnothing$, then

$$\begin{aligned}
\frac{|F \cap B(x, |y| + \varepsilon|y|)|}{|B(x, |y| + \varepsilon|y|)|} &\leq \frac{|B(x, |y| + \varepsilon|y|)| - |B(x+y, \varepsilon|y|)|}{|B(x, |y| + \varepsilon|y|)|} \\
&= \frac{C_d(|y| + \varepsilon|y|)^d - C_d(\varepsilon|y|)^d}{C_d(|y| + \varepsilon|y|)^d} \\
&= \frac{(1+\varepsilon)^d - \varepsilon^d}{(1+\varepsilon)^d} = 1 - \big(\frac{\varepsilon}{1+\varepsilon}\big)^d,
\end{aligned}$$

where C_d is the volume of the unit sphere in d dimensions.

Since x is a point of density of F, this is false, because the above quantity is *equal* to 1 for small radii. Thus there exists some ν such that $|y| < \nu$ implies $B(x+y, \varepsilon|y|)$ contains points of F, and therefore $\delta(x, y) < \varepsilon|y|$. □

Definition 19.2.9 (Marcinkiewicz integral)**.** Let F be a closed set in $\mathbb{R}^d$. Set $\delta(x) = \delta(x, F)$. For $x \in \mathbb{R}^d$, define

$$I(x) = \int_{|y|<1} \frac{\delta(x+y)}{|y|^{d+1}}\, dy$$

called the ***Marcinkiewicz integral***.

Theorem 19.2.10. *(i) If $x \notin F$, then $I(x) = \infty$.*

(ii) For almost every $x \in F$, $I(x) < \infty$.

Lecture 20 Marcinkiewicz Integral

20.1 Convergence of Marcinkiewicz Integral

Before we go on to prove the theorem about the convergence of Marcinkiewicz integrals stated last time, note that in $\mathbb{R}^d$,

$$\int_{|y|\leq 1} \frac{1}{|y|^{\alpha}}\, dy = \int_0^1 \int_{S^{d-1}} \frac{1}{r^{\alpha}} r^{d-1}\, d\sigma\, dr$$

here S^{d-1} is the $d-1$ dimensional sphere and σ is the surface area measure. This is then equal to

$$\sigma(S^{d-1}) \int_0^1 \frac{1}{r^{\alpha-d+1}}\, dr = \sigma(S^{d-1}) \begin{cases} \left.\frac{r^{d-\alpha}}{d-\alpha}\right|_0^1 & \text{if } d \neq \alpha \\ \left.\log(r)\right|_0^1 & \text{if } d = \alpha, \end{cases}$$

which is therefore infinite if $d \leq \alpha$ and otherwise $\sigma(S^{d-1})/(d-\alpha)$.

Note moreover that the previous bounds on $\delta(x+y)$ we've discussed are insufficient to prove the theorem, since $\delta(x+y) \leq |y|$ simply yields

$$I(x) \leq \int_{|y|<1} \frac{|y|}{|y|^{d+1}}\, dy = \int_{|y|<1} \frac{1}{|y|^{d}}\, dy = \infty,$$

so no information gained.

Similarly, using the second bound, saying that if $x \in F$ we have $\delta(d+y) = o(|y|)$, so in other words for every $\varepsilon > 0$ there exists a $\nu > 0$ such that $|y| < \nu$ implies $\delta(x+y) < \varepsilon|y|$. This means that if we pick a sequence ε_n going to 0, we can create a function $\gamma(y)$ such that $\delta(x+y) < \gamma(y)$.

Therefore

$$I(x) \leq \int_{|y|<1} \frac{o(|y|)}{|y|^{d+1}}\, dy = \int_{|y|<1} \frac{\gamma(y)|y|}{|y|^{d+1}}\, dy = \int_{|y|<1} \frac{\gamma(y)}{|y|^{d}}\, dy.$$

Therefore if, say, $\gamma(y) = |y|^{\beta}$ for $\beta > 0$, then the integral will converge, whereas if, perhaps, $\gamma(y) = 1/\log(1/|y|)$, the integral will be infinite.

In other words bounding the numerator by $o(|y|)$ is not enough either. All this to say, the theorem is quite subtle!

We'll prove the theorem using the following lemma:

Lemma 20.1.1. *Let F be a closed set whose complement has finite measure. Set*

$$I_*(x) = \int_{\mathbb{R}^d} \frac{\delta(x+y)}{|y|^{d+1}}\, dy.$$

Then $I_*(x) < \infty$ for almost every $x \in F$.

Moreover

$$\int_F I_*(x)\, dx \leq C|F^c| < \infty,$$

for some constant C.

Note that $I_*(x) \geq I(x)$ since the former integrates over a larger set.

Proof. We show the second part—doing so implies the former, since the integrand has to be finite for the entire integral to be finite. This goes more or less as usual: we use Fubini's theorem to switch the order of integration. First, however, we make the switch of variable $u = y + x$, so

$$\begin{aligned}\int_F I_*(x)\,dx &= \int_F \int_{\mathbb{R}^d} \frac{\delta(x+y)}{|y|^{d+1}}\,dy\,dx = \int_F \int_{\mathbb{R}^d} \frac{\delta(u)}{|u-x|^{d+1}}\,du\,dx \\ &= \int_F \int_{F^c} \frac{\delta(u)}{|u-x|^{d+1}}\,du\,dx = \int_{F^c} \int_F \frac{\delta(u)}{|u-x|^{d+1}}\,dx\,du,\end{aligned}$$

where the switch from $\mathbb{R}^d$ to F^c stems from $\delta(u) = 0$ on F.

Examining the inner integral, keeping $u \in F^c$ fixed. Thus $\delta(u) \leq |x-u|$, whereby

$$\begin{aligned}\int_F \frac{\delta(u)}{|u-x|^{d+1}}\,dx &\leq \int_{\{\,x||x-u|^{d+1}\,\}} \frac{\delta(u)}{|x-u|^{d+1}}\,dx = \delta(u) \int_{B(u,\delta(u))^c} \frac{1}{|x-u|^{d+1}}\,dx \\ &= \delta(u) \int_{B(o,\delta(u))^c} \frac{1}{|y|^{d+1}}\,dy = \delta(u) \int_{\delta(u)}^{\infty} \int_{S^{d-1}} \frac{1}{r^{d+1}} r^{d-1}\,d\sigma\,dr \\ &= \delta(u)\sigma(S^{d-1}) \int_{\delta(u)}^{\infty} \frac{1}{r^2}\,dr = \delta(u)\sigma(S^{d-1}) \frac{1}{\delta(u)} = C\end{aligned}$$

whereby

$$\int_F I_*(x)\,dx = \int_{F^c} \int_F \frac{\delta(u)}{|x-u|^{d+1}}\,dx\,du \leq \int_{F^c} C\,du = C|F^c|. \qquad \square$$

Proof of the Theorem. For (i), let $x \in F^c$. Then there exists some $r > 0$, and without loss of generality let us assume $r < 1$, such that $B(x,r) \subset F^c$ since F^c is open. Then if $|y| < r/2$, we have $\delta(x+y) \geq r/2$. Hence

$$I(x) = \int_{|y|<1} \frac{\delta(x+y)}{|y|^{d+1}}\,dy \geq \int_{|y|<r/2} \frac{\delta(x+y)}{|y|^{d+1}}\,dy \geq \frac{r}{2} \int_{|y|<r/2} \frac{1}{|y|^{d+1}}\,dy = \infty.$$

Next, for (ii), let M be a large positive integer. Set $F_M = F \cup B(0,M)^c$. Then by construction F_M^c has finite measure, so we can use the previous lemma. Let $\delta_M(x) = \delta(x, F_M)$ and so $\delta_M(x) \leq \delta(x)$ since F_M is bigger than F.

Also if $|y| \leq 1$, and $x \in B(0, M-2)$, then $\delta_M(x+y) = \delta(x+y)$. So for $x \in B(0, M-2)$,

$$I(x) = \int_{|y|<1} \frac{\delta(x+y)}{|y|^{d+1}}\,dy = \int_{|y|<1} \frac{\delta_M(x+y)}{|y|^{d+1}}\,dy \leq \int_{\mathbb{R}^d} \frac{\delta_M(x+y)}{|y|^{d+1}} = I_*(x)$$

for $x \in F_M$. So for almost every $x \in F_M \cap B(0, M-2) = F \cap B(0, M-2)$ we have $I(x) < \infty$, and so if we let M go to infinity, we have that for almost every $x \in F$, $I(x) < \infty$. $\square$

There are different varieties of this result. For instance,

$$I^\lambda(x) = \int_{|y|<1} \frac{\delta(x+y)^\lambda}{|y|^{d+\lambda}}\,dy$$

for $\lambda > 0$. In other words our result is the special case $\lambda = 1$.

Lecture 21 Integral Operators

21.1 Schur's Lemma

As a change of topic, we will now concern ourselves with certain integral operators for a bit. In particular, let (X, μ) and (Y, ν) be measure spaces, and let $k\colon X \times Y \to \mathbb{C}$ be a (jointly) measurable function. Define T as a map from the measurable functions on Y to the measurable functions on X by

$$Tf(x) = \int_Y k(x,y) f(y)\, d\nu(y).$$

A fairly fundamental result about these is this:

Theorem 21.1.1 (Schur's lemma)**.** *Let (X, μ) and (Y, ν) be measure spaces, and let $k\colon X \times Y \to \mathbb{C}$ be measurable. Set*

$$Tf(x) = \int_Y k(x,y) f(y)\, d\nu(y).$$

Suppose that there exists a constant C such that

(i) $\int_X |k(x,y)|\, d\mu(x) \le C$ *independent of y and*

(ii) $\int_Y |k(x,t)|\, d\nu(y) \le C$ *independent of x.*

Then T is bounded on L^p, i.e. for each p with $1 \le p \le \infty$, there exists a constant C_p^* such that $\|Tf\|_{L^p(X)} \le C_p^* \|f\|_{L^p(Y)}$.

Proof. First consider $p = \infty$. Then we have

$$|Tf(x)| \le \int_Y |k(x,y)||f(y)|\, d\nu(y) \le \|f\|_\infty \int_Y |k(x,y)|\, d\nu(y) \le C\|f\|_\infty,$$

so T maps a function in $L^\infty(Y)$ to a function in $L^\infty(X)$.

Next consider $p = 1$. Then

$$\begin{aligned}
\int_X |Tf(x)|\, d\mu(x) &= \int_X \left| \int_Y k(x,y) f(y)\, d\nu(y) \right| d\mu(x) \\
&\le \int_X \int_Y |k(x,y)||f(y)|\, d\nu(y)\, d\mu(x) \\
&= \int_Y |f(y)| \int_X |k(x,y)|\, d\mu(x)\, d\nu(y) \\
&\le C \int_Y |f(y)|\, d\nu(y) = C\|f\|_{L^1(Y)}.
\end{aligned}$$

Therefore T is (strong) type $(1,1)$ and (∞,∞), so by Marcinkiewicz interpolation T is strong type (p,p) for all $1 < p < \infty$. □

It is possible to fashion this into a new and improves version of the same result—we won't prove it, but it's similar:

Theorem 21.1.2. *Suppose k is as in the previous theorem. Suppose k satisfies*

(i) $\int_X |k(x,y)|^r \, d\mu(x) \le C$ *independent of y and*

(ii) $\int_Y |k(x,t)|^r \, d\nu(y) \le C$ *independent of x.*

Then $\|Tf\|_{L^q(X)} \le C^*_{p,q} \|f\|_{L^p(Y)}$ where $1/q = 1/p + 1/r - 1$.

We will now try to answer the question what happens when we don't have these strong convergence conditions on k. That is, define T in the same way, but suppose we have

$$\int_{\mathbb{R}^d} |k(x,y)| \, dx = \int_{\mathbb{R}^d} |k(x,y)| \, dy = \infty.$$

Let $f \in L^1(\mathbb{R})$ and consider

$$h(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(t)}{t - z} \, dt.$$

Then h is analytic in the upper half plane.

Note that we can bound it by

$$\int_{\mathbb{R}} \left| \frac{f(t)}{t-z} \right| dt \le \int_{\mathbb{R}} \frac{|f(t)|}{\operatorname{Im} z} \, dt = \frac{1}{\operatorname{Im} z} \int_{\mathbb{R}} |f(t)| \, dt.$$

Let $z = x + iy$, and we get

$$\begin{aligned} h(z) &= \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(t)}{t - (x+iy)} \, dt = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(t)}{(t-x) - iy} \cdot \frac{(t-x)+iy}{(t-x)+iy} \, dt \\ &= \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(t)(t-x)}{(t-x)^2 + y^2} + \frac{iyf(t)}{(t-x)^2+y^2} \, dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{f(t)y}{(t-x)^2+y^2} \, dt + \frac{-i}{2\pi} \int_{\mathbb{R}} \frac{f(t)(t-x)}{(t-x)^2+y^2} \, dt. \end{aligned}$$

Suppose f is real-valued. Then we have just decomposed h into real and imaginary parts.

It is a fact known as Fatou's theorem that

$$\lim_{\varepsilon \to 0^+} h(x + i\varepsilon)$$

exists for almost every x. Call this limit $h(x)$.

We claim that

$$\operatorname{Im} h(x) = \lim_{\varepsilon \to 0} -\frac{1}{2\pi} \int_{\{\, t \mid |t-x| > \varepsilon \,\}} \frac{f(t)}{t-x} \, dt.$$

I.e. the imaginary part of h is its own so-called ***Hilbert transform***.

Note for the record that

$$\int_{\mathbb{R}} \frac{1}{|t-x|} \, dt = \infty$$

for all x, so Schur's lemma doesn't work.

The proof of this claim largely boils down to symmetry. Fix x and $\varepsilon > 0$. Call $\{\, t \mid |t-x| > \varepsilon \,\} = T_\varepsilon$ and let us examine

$$\begin{aligned} &\left| \int_{T_\varepsilon} \frac{f(t)}{t-x}\,dt - \int_{\mathbb{R}} \frac{f(t)(t-x)}{(t-x)^2+\varepsilon^2}\,dt \right| \\ &\quad \le \left| \int_{T_\varepsilon} f(t)\Big(\frac{1}{t-x} - \frac{t-x}{(t-x)^2+\varepsilon^2}\Big)\,dt \right| + \left| \int_{T_\varepsilon} f(t)\frac{t-x}{(t-x)^2+\varepsilon^2} \right| \\ &\quad = \left| \int_{T_\varepsilon} f(t)\frac{(t-x)^2+\varepsilon^2-(t-x)^2}{(t-x)((t-x)^2+\varepsilon^2)}\,dt \right| + \left| \int_{T_\varepsilon} (f(t)-f(x))\frac{t-x}{(t-x)^2+\varepsilon^2}\,dt \right| \end{aligned}$$

since the area over which we are integrating is symmetric with respect to x, and the fraction in the second integral is odd. Bounding appropriately by ε we thus get

$$\le \left| \int_{T_\varepsilon} f(t)\frac{\varepsilon^2}{(t-x)((t-x)^2+\varepsilon^2)}\,dt \right| + \frac{1}{\varepsilon}\int_{T_\varepsilon} |f(t)-f(x)|\,dt = \mathrm{I} + \mathrm{II}.$$

By Lebesgue Differentiation theorem, II goes to 0 almost everywhere as $\varepsilon \to 0$. For the first integral, let $T_n = \{\, t \mid 2^{n+1}\varepsilon \ge |t-x| > 2^n\varepsilon \,\}$ and write

$$\begin{aligned} \mathrm{I} &\le \sum_{n=0}^{\infty} \int_{T_n} \frac{|f(t)-f(x)|\varepsilon^2}{|t-x||(t-x)^2+\varepsilon^2|}\,dt \le \sum_{n=0}^{\infty} \int_{T_n} \frac{|f(t)-f(x)|\varepsilon^2}{2^n\varepsilon((2^n\varepsilon)^2+\varepsilon^2)}\,dt \\ &= \sum_{n=0}^{\infty} \frac{1}{2^n\varepsilon} \int_{T_n} \frac{|f(t)-f(x)|}{2^{2n}+1}\,dt = \sum_{n=0}^{\infty} \frac{1}{(2^n\varepsilon)(2^{2n}+1)} \int_{T_n} |f(t)-f(x)|\,dt \\ &\le \sum_{n=0}^{N} \frac{\int_{T_n}|f(t)-f(x)|\,dt}{(2^n\varepsilon)(2^{2n}+1)} + \sum_{n=N+1}^{\infty} \frac{\int_{T_n}|f(t)-f(x)|\,dt}{(2^n\varepsilon)(2^{2n}+1)} \\ &\le \sum_{n=0}^{N} \frac{1}{(2^n\varepsilon)(2^{2n}+1)} \int_{T_n} |f(t)-f(x)|\,dt + \sum_{n=N+1}^{\infty} \frac{2Mf(x)}{2^{2n}+1}, \end{aligned}$$

which goes to 0 as $\varepsilon \to 0$, again by Lebesgue differentiation theorem.

To conclude, then, if we take $f \in L^1(\mathbb{R})$ to be real-valued and set

$$h(z) = \frac{1}{2\pi i}\int_{\mathbb{R}} \frac{f(t)}{t-z}\,dt,$$

we have that

$$\lim_{y\to 0} h(x+iy) := h(x)$$

exists almost everywhere and

$$\operatorname{Im} h(x) = \lim_{\varepsilon\to 0} \int_{\{\, t \mid |t-x|>\varepsilon \,\}} \frac{f(t)}{t-x}\,dt.$$

It is also possible to show that $\operatorname{Re} h(x) = f(x)$ almost everywhere.

Lecture 22 Integral Operators continued

22.1 Singular Integrals

Theorem 22.1.1. *Let $k(x) \in L^2(\mathbb{R}^d)$. Suppose*

(i) $|\hat{k}(\xi)| \leq B$ *for every* $\xi \in \mathbb{R}^d$*, and*

(ii) k *is* C^1 *on* $\mathbb{R}^d \setminus \{ 0 \}$ *and* $|\nabla k(x)| \leq B/|x|^{n+1}$.

For $f \in L^p \cap L^2$*, set*

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

Then for $1 < p < \infty$ *there exists a constant* $A_{p,d,B}$ *such that*

$$\|Tf\|_p \leq A_{p,d,B} \|f\|_p$$

for every $f \in L^p(\mathbb{R}^d)$.

This is due to Calderón and Zygmund in the early nineteen fifties, and uses their decomposition.

Ultimately we want to consider $k(x, y)$, of course—this then is a special case.

The integral operator T thus defined is a ***singular integral operator***, and

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

itself is a ***singular integral***. The name comes from the fact that k can have singularities:

Example 22.1.2. We saw last time that

$$Hf(x) = \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} \frac{1}{x-y} f(y)\, dy$$

works in this framework, and it certainly has a singularity, since $k(x) = 1/x$. ▲

The proof is in four parts:

1. Show that T is type $(2, 2)$.
2. Show that T is weak type $(1, 1)$.
3. Apply Macinkiewicz interpolation to get that T is type (p, p) for all $1 < p < 2$.
4. Use the fact that the dual space to L^p is L^q, with $1/p + 1/q = 1$, so it $1 < p < 2$, we have $2 < q < \infty$.

Proof of part 1. Let $f \in L^2$ and consider

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

Since we can think of the above as a convolution, we have $\widehat{Tf}(\xi) = \hat{k}(\xi)\hat{f}(\xi)$. Hence

$$\|Tf\|_2 = \|\widehat{Tf}\|_2 = \|\hat{k}\hat{f}\|_2 \leq \|\hat{k}\|_\infty \|\hat{f}\|_2 \leq B\|f\|_2.$$ □

Unfortunately, step 2 is not nearly as simple:

Proof of part 2. We want to show that there exists some constant C such that

$$|\{ x \in \mathbb{R}^d \mid |Tf(x)| > \lambda \}| \leq \frac{C}{\lambda} \int_{\mathbb{R}^d} |f(x)|\, dx$$

for every $\lambda > 0$.

Fix λ and apply the Calderón-Zygmund decomposition to $|f(y)|$, obtaining F and Ω, with Ω composed of cubes Q_j. Define

$$g(x) = \begin{cases} f(x) & \text{if } x \in F \\ \frac{1}{|Q_j|} \int_{Q_j} f(y)\, dy & \text{if } x \in Q_j \text{ for some } j \end{cases}$$

and

$$b(x) = \begin{cases} 0 & \text{if } x \in F \\ f(x) - \frac{1}{|Q_j|} \int_{Q_j} f(y)\, dy & \text{if } x \in Q_j \text{ for some } j. \end{cases}$$

Then $f(x) = g(x) + b(x)$, where we call g the good function and b the bad function. This is due to $|g(x)| \leq 2^d \lambda$ for all $x \in \mathbb{R}^d$.

On the other hand, for each Q_j we have

$$\frac{1}{|Q_j|} \int_{Q_j} b(x)\, dx = \frac{1}{|Q_j|} \int_{Q_j} f(x) - \frac{1}{|Q_j|} \int_{Q_j} f(y)\, dy\, dx = 0$$

so the average of b is 0 on each cube. By linearity of T we moreover have $Tf = Tg + Tb$, which means that $|Tf| \leq |Tg| + |Tb|$, whence

$$|\{ x \mid |Tf(x)| > \lambda \}| \leq |\{ x \mid |Tg(x)| > \lambda/2 \}| + |\{ x \mid |Tb(x)| > \lambda/2 \}|.$$

So we estimate those two measures. First, Tg, by

$$\begin{aligned} \|g\|_2^2 &= \int_{\mathbb{R}^d} |g(x)|^2\, dx = \int_F |g(x)|^2\, dx + \int_\Omega |g(x)|^2\, dx \\ &\leq \int_F \lambda |f(x)|\, dx + (2^d \lambda)^2 |\Omega| \leq \lambda \|f\|_1 + (2^d\lambda)^2 \frac{1}{\lambda} \int_\Omega |f(x)|\, dx \leq C\lambda \|f\|_1. \end{aligned}$$

There are two things of particular note here. In the first inequality, we have bounded one of the g by λ, and the other by f, on F. Secondly, the inequality from $|\Omega|$ to the integral is the Calderón-Zygmund decomposition rearranged to some for the measure of the cubes.

Now by part 1 above,

$$\|Tg\|_2^2 \leq C \|g\|_2^2 \leq C\lambda \|f\|_1,$$

implying that

$$|\{ x \mid |Tg(x)| > \lambda/2 \}| \leq \frac{1}{(\lambda/2)^2} \|Tg\|_2^2 \leq \frac{4}{\lambda^2} C\lambda \|f\|_1 = \frac{C}{\lambda} \|f\|_1.$$

It remains to ponder Tb. Set, for convenience,

$$b_j(x) = \begin{cases} b(x) & \text{if } x \in Q_j \\ 0 & \text{otherwise.} \end{cases}$$

Then $b(x) = \sum_j b_j(x)$, with b_j supported on Q_j, and on Q_j we have

$$b_j(x) = f(x) = \frac{1}{|Q_j|} \int_{Q_j} f(y)\, dy.$$

So $Tb(x) = \sum_j Tb_j(x)$, which works since, even though it's an infinite sum, the left-hand side is an integral which we decomposed into integrals over disjoint parts, so there is no problem with interchanging the sum and the integral.

Fixing j, we then have

$$Tb_j(x) = \int_{Q_j} k(x-y) b_j(y)\, dy.$$

Let c_j denote the centre of the cube Q_j. Then

$$Tb_j(x) = \int_{Q_j} (k(x-y) - k(x-c_j)) b_j(y)\, dy$$

since

$$\int_{Q_j} k(x-c_j) b_j(y)\, dy = 0$$

due to the first factor not depending on y, so it comes out in front, and b_j has zero average.

Let $\tilde{Q}_j$ be the cube with centre c_j but side lengths 2 times that of Q_j. Suppose $x \notin \tilde{Q}_j$. Then for $y \in Q_j$ we have

$$|k(x-y) - k(x-c_j)| \le |y - c_j||\nabla k(x - \tilde{y})|,$$

with $\tilde{y}$ on the line segment between y and c_j. Think of this as nothing more than the Mean Value Theorem, applied on that same line segment.

So

$$|k(x-y) - k(x-c_j)| \le (diam Q_j) \frac{CB}{|x-\tilde{y}|^{d+1}} \le (\operatorname{diam} Q_j) \frac{CB}{|x-c_j|^{d+1}}$$

since

$$|x - c_j| \le |x - \tilde{y}| + |\tilde{y} - c_j| \le |x - \tilde{y}| + \operatorname{diam} Q_j \le 2|x - \tilde{y}|.$$

Therefore

$$\begin{aligned} |Tb_j(x)| &\le C \frac{\operatorname{diam} Q_j}{|x-c_j|^{d+1}} \int_{Q_j} |b_i(y)|\, dy \\ &\le C \frac{\operatorname{diam} Q_j}{|x-c_j|^{d+1}} \int_{Q_j} \left| f(x) - \frac{1}{|Q_j|} \int_{Q_j} f(y)\, dy \right| dx \\ &\le 2C \frac{\operatorname{diam} Q_j}{|x-c_j|^{d+1}} \int_{Q_j} |f(y)|\, dy \le \frac{C \operatorname{diam} Q_j}{|x-c_j|^{d+1}} 2^d \lambda |Q_j| \\ &= D \frac{\operatorname{diam} Q_j}{|x-c_j|^{d+1}} \lambda |Q_j|. \end{aligned}$$

Therefore

$$\int_{\tilde{Q}_j^c} |Tb_j(x)|\, dx \le D \operatorname{diam} Q_j |Q_j| \lambda \int_{\tilde{Q}_j^c} \frac{1}{|x-c_j|^{d+1}}\, dx,$$

which, if we switch to polar coordinates, integrate over the complement of the ball inscribed in $\tilde{Q}_j$, call it $\tilde{B}_j^c$, and recentre at the origin, if bounded by

$$D \operatorname{diam} Q_j |Q_j| \lambda \int_{\tilde{B}_j^c} \frac{1}{r^{d+1}} r^{d-1}\, dr\, d\sigma = C \operatorname{diam} Q_j |Q_j| \lambda \frac{C}{\operatorname{diam} Q_j} = D|Q_j|\lambda.$$

□

Lecture 23 Integral Operators continued

23.1 Finishing the Proof

We continue from where we left off with the proof of Part 2:

Proof of Part 2, continued. Set $E = (\bigcup_j \tilde{Q}_j)^c$. Then $E \subset \tilde{Q}_j^c$ for every j, and

$$\begin{aligned} \int_E |Tb(x)|\, dx &\leq \sum_j \int_E |Tb_j(x)|\, dx \leq \sum_j \int_{\tilde{Q}_j^c} |Tb_j(x)|\, dx \\ &\leq \sum_j D|Q_j|\lambda \leq D\lambda \sum_j \frac{1}{\lambda} \int_{Q_j} |f(x)|\, dx \leq D\|f\|_1. \end{aligned}$$

So on E, Tb is in L^1. Hence

$$|\{\, x \in E \mid |Tb(x)| > \lambda/2 \,\}| \leq \frac{D\|f\|_1}{\lambda}$$

by Chebyshev. Moreover, $E^c = \bigcup_j \tilde{Q}_j$, and thus

$$|E^c| = \Big|\bigcup_j \tilde{Q}_j\Big| \leq \sum_j |\tilde{Q}_j| \leq C \sum_j |Q_j| \leq C \sum_j \frac{1}{\lambda} \int_{Q_j} |f(x)|\, dx \leq \frac{C}{\lambda}\|f\|_1.$$

Therefore

$$|\{\, x \in E^c \mid |Tb(x)| > \lambda/2 \,\}| \leq |E^c| \leq \frac{C}{\lambda}\|f\|_1,$$

so T is weak type $(1,1)$. □

Remark 23.1.1. The only place we used and will use the hypothesis $|\nabla k(x)| \leq B/|x|^{d+1}$ is the estimate

$$\int_{Q_j^c} |Tb_j(x)|\, dx \leq D\lambda|Q_j|.$$

We can replace it by

$$\int_{|x| \geq 2|y|} |k(x-y) - k(x)|\, dx \leq B.$$

Proof of Step 3. We have shown that T is type $(2,2)$ and weak type $(1,1)$, so by Marcinkiewicz interpolation T is type (p,p) for all $1 < p < 2$. □

Proof of Step 4. We want to show that the operator T is type (p,p) for $2 < p < \infty$. Let $f \in L^p \cap L^1$, $2 < p < \infty$. Let $\varphi \in L^q$, with $1/p + 1/q = 1$, such that $\|\varphi\|_q = 1$ and φ is continuous with compact support. Then

$$\begin{aligned}\int_{\mathbb{R}^d} Tf(x)\varphi(x)\,dx &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x-y)f(y)\,dy\varphi(x)\,dx \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x-y)f(y)\varphi(x)\,dx\,dy,\end{aligned}$$

where we used Fubini's theorem in the last equality. To see why this is valid, consider

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |k(x-y)\varphi(x)f(y)|\,dx\,dy &= \int_{\mathbb{R}^d}|f(y)|\int_{\mathbb{R}^d}|k(x-y)\varphi(x)|\,dx\,dy \\ &= \int_{\mathbb{R}^d}|f(y)|\Big(\int_{\mathbb{R}^d}|k(x-y)|^2\,dx\Big)^{1/2}\Big(\int_{\mathbb{R}^d}|\varphi(x)|^2\,dx\Big)^{1/2}dy \\ &= \int_{\mathbb{R}^d}|f(y)|\,dy\|\varphi\|_2\varphi k_2 = \|f\|_1\|\varphi\|_2\varphi k_2 < \infty.\end{aligned}$$

Set $\tilde{k}(x) = k(-x)$. Then

$$\int_{\mathbb{R}^d} Tf(x)\varphi(x)\,dx = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\tilde{k}(y-x)\varphi(x)\,dxf(y)\,dy = \int_{\mathbb{R}^d}\tilde{T}\varphi(y)f(y)\,dy$$

where $\tilde{T}$ is the operator with kernel $\tilde{k}$. Note that $\tilde{k}$ satisfies all the same estimates as k, so $\tilde{T}$ is type (p,p) for $1 < p < 2$. Therefore

$$\left|\int_{\mathbb{R}^d} Tf(x)\varphi(x)\,dx\right| = \left|\int_{\mathbb{R}^d}\tilde{T}\varphi(x)f(x)\,dx\right| \leq \|\tilde{T}f\|_1\|f\|_p\,eqC\|\varphi\|_1\|f\|_p \leq C\|f\|_p.$$

Since

$$\|Tf\|_p = \sup\left\{\left|\int_{\mathbb{R}^d} Tf(x)\varphi(x)\,dx\right| \;\middle|\; \varphi \in L^q, \|\varphi\|_q = 1, \varphi \text{ compact support}\right\}$$

we get $\|Tf\|_p \leq C\|f\|_p$. □

Theorem 23.1.2. *Suppose $k(x)$, $x \in \mathbb{R}^d$, satisfies*

(*i*) $|k(x)| \leq B/|x|^d$, $|x| > 0$;

(*ii*) $\displaystyle\int_{|x|\geq 2|y|}|k(x-y)-k(x)|\,dx \leq B$ *for all* $y \in \mathbb{R}^d$;

(*iii*) $\displaystyle\int_{R_1<|x|<R_2} k(x)\,dx = 0$ *whenever* $0 < R_1 < |x| < R_2 < \infty$.

For $f \in L^p$, $1 < p < \infty$, set

$$T_\varepsilon f(x) = \int_{|y|>\varepsilon} f(x-y)k(y)\,dy.$$

Then there exists a constant $A = A_{p,d,B}$ such that $\|T_\varepsilon f\|_p \leq A\|f\|_p$ independent of ε and

$$\lim_{\varepsilon\to 0} T_\varepsilon f(x) = Tf(x)$$

exists in L^p norm and $\|Tf\|_p \leq A\|f\|_p$.

Turns out that in fact the above limit also exists pointwise almost everywhere, but that is much more work.

Example 23.1.3. Consider the ***Hilbert transform***

$$Hf(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|y| \geq \varepsilon} f(x-y) \frac{1}{y} \, dy,$$

i.e. the above with the kernel $k(x) = 1/(\pi x)$. Then k satisfies the hypotheses of the theorem; clearly $|k(x)| \leq B/|x|$, where $B = 1/\pi$, the last property is satisfies by symmetry since k is odd, and for (ii) we have

$$\begin{aligned} \int_{|x| \geq 2|y|} |k(x-y) - k(x)| \, dx &= \frac{1}{\pi} \int_{|x| \geq 2|y|} \left| \frac{1}{x-y} - \frac{1}{x} \right| dx \\ &= \frac{1}{\pi} \int_{|x| \geq 2|y|} \frac{|y|}{|x-y||x|} \, dx \\ &\leq \frac{1}{\pi} \int_{|x| \geq 2|y|} \frac{|y|}{(|x| - |y|)|x|} \, dx \\ &\leq \frac{|y|}{\pi} \int_{|x| \geq 2|y|} \frac{1}{(|x| - |y|)|x|} \, dx \\ &\leq 2 \frac{|y|}{\pi} \int_{2|y|}^{\infty} \frac{1}{|x|^2} \, dx = C. \end{aligned}$$

▲

Lemma 23.1.4. *Suppose k satisfies (i)–(iii) of the last theorem. Set*

$$k_\varepsilon(x) = \begin{cases} k(x) & \text{if } |x| \geq \varepsilon \\ 0 & \text{if } |x| < \varepsilon. \end{cases}$$

Then $k_\varepsilon \in L^2(\mathbb{R}^d)$ and $\|\hat{k}_\varepsilon\| \leq CB$ where B is as before, and C is a constant independent of ε and B.

Proof. It suffices to consider $\varepsilon = 1$. Suppose we know $|\hat{k}(\xi)| \leq CB$, and consider $\tilde{k}(x) = \varepsilon^d k(\varepsilon x)$. Then $\tilde{k}$ satisfies (i)–(iii) of the theorem with the same constant B. To see (i), consider

$$|\tilde{k}(x)| = |\varepsilon^d k(\varepsilon x)| \leq \varepsilon^d \frac{B}{|\varepsilon x|^d} = \frac{B}{|x|^d},$$

and similarly (ii) and (iii). Then

$$\frac{1}{\varepsilon^d} (\tilde{k})_1(x/\varepsilon) = \begin{cases} \frac{1}{\varepsilon^d} \hat{k}(x/\varepsilon) & \text{if } |x/\varepsilon| > 1 \\ 0 & \text{if } |x/\varepsilon| \leq 1. \end{cases}$$

and

$$\hat{k}_\varepsilon(\xi) = \left| \widehat{\frac{1}{\varepsilon^d} \tilde{k}_1(\cdot/\varepsilon)}(\xi) \right| = |\widehat{(\tilde{k})_1}(\varepsilon y)| \leq CB.$$

□

Lecture 24 Integral Operators continued

24.1 Proof of the Lemma

We start by finishing up the proof of the lemma from last time.

Proof continued. As discussed it suffices to consider k_1. Then we have

$$\begin{aligned}\hat{k}_1(y) &= \lim_{\mathbb{R}\to\infty}\int_{|x|\le R} e^{ix\cdot y}k_1(x)\,dx\\ &= \int_{|x|\le 2\pi/|y|} e^{-ix\cdot y}k_1(x)\,dx + \lim_{R\to\infty}\int_{2\pi/|y|\le|x|\le R} e^{-ix\cdot y}k_1(x)\,dx\\ &= \mathrm{I}+\mathrm{II}.\end{aligned}$$

We can bound I using property (*iii*) from the assumptions of our theorem, i.e.

$$\mathrm{I} = \int_{|x|\le 2\pi/|y|} e^{-ix\cdot y}k_1(x)\,dx = \int_{|x|\le 2\pi/|y|}(e^{-ix\cdot y}-1)k_1(x)\,dx$$

meaning that

$$\begin{aligned}|\mathrm{I}| &\le \int_{|x|\le 2\pi/|y|}|e^{-ix\cdot y}-1||k_1(x)|\,dx\\ &\le \int_{|x|\le 2\pi/|y|}|x||y||k_1(x)|\,dx \le |y|\int_{|x|\le 2\pi/|y|}|x|\frac{B}{|x|^d}\,dx\\ &= B|y|\int_{|x|\le 2\pi/|y|}\frac{1}{|x|^{d-1}}\,dx = B|y|\int_{S^{d-1}}\int_0^{2\pi/|y|}\frac{1}{r^{d-1}}r^{d-1}\,dr\,d\sigma\\ &= C_dB,\end{aligned}$$

so I is bounded by a constant independent of ε. Note that we used the property that $|e^{i\theta}-1| = |e^{i\theta}-e^{i0}| \le \theta$.

For II, set $z = y\pi/|y|^2$, and note that $e^{-iz\cdot y} = -1$. Then

$$\begin{aligned}\int_{2\pi/|y|\le|x|\le R} e^{-ix\cdot y}k_1(x)\,dx &= \int_{2\pi/|y|\le|x-z|\le R} k_1(x-z)e^{-i(x-z)\cdot y}\,dx\\ &= -\int_{\le|x-z|\le R} k_1(x-z)e^{-ix\cdot y}\,dx.\end{aligned}$$

Therefore our integral is bounded by

$$\frac{1}{2}\Big(\int_{2\pi/|y|\le|x|\le R} e^{-ix\cdot y}k_1(x)\,dx - \int_{2\pi/|y|\le|x-z|\le R} k_1(x-z)e^{-ix\cdot y}\,dx\Big)$$

which in turn we can write as

$$\begin{aligned}&\frac{1}{2}\Big(\int_{2\pi/|y|\le|x|\le R} e^{-ix\cdot y}k_1(x)\,dx - \int_{2\pi/|y|\le|x|\le R} e^{-ix\cdot y}k_1(x-z)\,dx+\\ &+\int_{2\pi/|y|\le|x|\le R} e^{-ix\cdot y}k_1(x-z)\,dx - \int_{2\pi/|y|\le|x-z|\le R} k_1(x-z)e^{-ix\cdot y}\,dx\Big).\end{aligned}$$

Call the first line II_a, and the second line II_b.

Now

$$|\mathrm{II}_a| \le \frac{1}{2}\int_{2\pi/|y|\le|x|\le R} |k_1(x) - k_1(x-z)|\,dx$$
$$= \frac{1}{2}\int_{2|z|\le|x|\le R} |k_1(x) - k_1(x-z)|\,dx \le B$$

by assumption. For II_b, we have the same integrand but different regions, so let

$$\mathcal{R} = \{\, x \mid 2\pi/|y| \le |x| \le R\,\} \triangle \{\, x \mid 2\pi/|y| \le |x-z| \le R\,\}$$

and therefore

$$|\mathrm{II}_b| \le \left| \int_{\mathcal{R}} e^{-ix\cdot y} k_1(x-z)\,dx \right|$$

and we can integrate over two large annuli which captures this symmetric difference (plus a bit more, but we have control over it):

$$|\mathrm{II}_b| \le C\int_{\pi/|y|\le|u|\le 3\pi/|y|} \frac{B}{|u|^d}\,du + C\int_{R-\pi/|y|\le|u|^d\le R+\pi/|y|} \frac{B}{|u|^d}\,du$$
$$= C\int_{\pi/|y|}^{3\pi/|y|}\int_{S^{d-1}} \frac{B}{r^d} r^{d-1}\,d\sigma\,dr + C\int_{R-\pi/|y|}^{R+\pi/|y|}\int_{S^{d-1}} \frac{B}{r^d} r^{d-1}\,d\sigma\,dr$$
$$= CB\int_{\pi/|y|}^{3\pi/|y|} \frac{1}{r}\,dr + CB\int_{R-\pi/|y|}^{R+|y|} \frac{1}{r}\,dr$$
$$= CB\log(3) + CB\log\Big(\frac{R+\pi/|y|}{R-\pi/|y|}\Big) \le CB,$$

and we are done. □

Lecture 25 Integral Operators continued

25.1 Finalising the Proof

We finish off the course by proving Theorem 23.1.2:

Proof. We need to show that each k_ε satisfies the second property of the previous theorem, namely that

$$\int_{|x|\ge 2|y|} |k_\varepsilon(x-y) - k_\varepsilon(x)|\,dx \le B.$$

To accomplish this, note that we can write this integral as

$$\int\limits_{\substack{|x|\ge 2|y| \\ |x-y|>\varepsilon \\ |y|>\varepsilon}} |k(x-y) - k(x)|\,dx + \int\limits_{\substack{|x|\ge 2|y| \\ |x-y|<\varepsilon \\ |y|>\varepsilon}} |k(x)|\,dx + \int\limits_{\substack{|x|\ge 2|y| \\ |x-y|>\varepsilon \\ |y|<\varepsilon}} |k(x-y)|\,dx,$$

which we denote I, II, and III respectively.

The first one is bounded by B by hypothesis. For the second one we have $1/|x| > 1/\varepsilon$, so

$$\mathrm{II} \le \int_{|x-y|<\varepsilon} \frac{B}{\varepsilon^d}\,dx = C_d\varepsilon^d\frac{B}{\varepsilon^d} = C_dB.$$

Moreover

$$\mathrm{III} \le \int_{|x|<\varepsilon} \frac{B}{\varepsilon^d}\,dx = C_d\varepsilon^d\frac{B}{\varepsilon^d} = C_dB.$$

Hence by the previous theorem $\|T_\varepsilon f\|_p \le A_{B,p,d}\|f\|_p$ independent of ε. Now we need to prove the existence of the limit.

Consider $f \in C^\infty$ with compact support. Then

$$\begin{aligned} T_\varepsilon f(x) &= \int_{|y|\ge\varepsilon} k(y)f(x-y)\,dy \\ &= \int_{|y|\ge 1} k(y)f(x-y)\,dy + \int_{1\ge|y|\ge\varepsilon} k(y)(f(x-y)-f(x))\,dy = \mathrm{I}+\mathrm{II}, \end{aligned}$$

where we can add $f(x)$ as we do since it's constant with respect to y, meaning that in the second integral we'd just be integrating k over a ring, where it is zero by hypothesis.

Taking $1/p + 1/q = 1$ and using Hölder's inequality we have

$$\begin{aligned} |\mathrm{I}| &= \left|\int_{\mathbb{R}^d} k_1(y)f(x-y)\,dy\right| \le \Big(\int_{\mathbb{R}^d} |k_1(y)|^q\,dy\Big)^{1/q}\Big(\int_{\mathbb{R}^d} |f(x-y)|^p\,dy\Big)^{1/p} \\ &= C\|f\|_p. \end{aligned}$$

For the second one, we have by the Mean value theorem that

$$\begin{aligned} |\mathrm{II}| &\le \int_{1\ge|y|\ge\varepsilon} |k(y)|\|\nabla f\|_\infty |y|\,dy \le \|\nabla f\|_\infty \int_{1\ge|y|\ge\varepsilon} \frac{B}{|y|^d}|y|\,dy \\ &\le B\omega_d\|\nabla f\|_\infty, \end{aligned}$$

where by ω_d we mean the volume of the unit ball in $\mathbb{R}^d$. So if $\varepsilon_1 > \varepsilon_2$, then

$$\begin{aligned} |T_{\varepsilon_1}f(x) - T_{\varepsilon_2}f(x)| &\le \int_{\varepsilon_1\ge|y|\ge\varepsilon_2} |k(y)|\|\nabla f\|_\infty|y|\,dy \le CB\int_{\varepsilon_1\ge|y|\ge\varepsilon_2} \frac{1}{|y|^{d-1}}\,dy \\ &= CB\int_{S^{d-1}}\int_{\varepsilon_1}^{\varepsilon_2} \frac{1}{r^{d-1}}r^{d-1}\,dr\,d\sigma \le CB|\varepsilon_2-\varepsilon_1|. \end{aligned}$$

Hence $T_\varepsilon f(x)$ converges uniformly in ε. Now consider $f \in L^p$ and let $\eta > 0$. Write $f = f_1 + f_2$, with $f_1 \in C^\infty$ with compact support and $\|f_2\|_p < \eta$, which is possible since smooth functions with compact support are dense in L^p.

Then

$$T_\varepsilon f(x) = T_\varepsilon f_1(x) + T_\varepsilon f_2(x)$$

by linearity of T_ε, so

$$\begin{aligned} \|T_{\varepsilon_1}f(x) - T_{\varepsilon_2}f(x)\| &\le \|T_{\varepsilon_1}f_1(x) - T_{\varepsilon_2}f_1(x)\|_p + \|T_{\varepsilon_1}f_2(x) - T_{\varepsilon_2}f_2(x)\|_p \\ &\le \|T_{\varepsilon_1}f_1(x) - T_{\varepsilon_2}f_1(x)\|_p + A\|f_2\|_p + A\|f_2\|_p \\ &\le \|T_{\varepsilon_1}f_1(x) - T_{\varepsilon_2}f_1(x)\|_p + 2A\eta. \end{aligned}$$

By the above we can choose ε_1 and ε_2 small enough so that the first part is bounded by η, and so $T_\varepsilon f$ is a Cauchy sequence in L^p, in the sense that we can choose any sequence of ε_n that will be Cauchy. Hence

$$\lim_{\varepsilon \to 0} T_\varepsilon f = Tf$$

exists in L^p. □

Note that we did not show that this limit exists pointwise almost everywhere. This is true, but to do it we need to create

$$T^* f(x) = \sup_{\varepsilon > 0} \|T_\varepsilon f(x)\|$$

and do similar maximal function computations we did before.

Index