HARMONIC ANALYSIS

Contents

1.	Introduction	2
2.	Hardy-Littlewood maximal function	3
3.	Approximation by convolution	16
4.	Muckenhoupt weights	29
4.1.	Calderón-Zygmund decomposition	36
4.2.	Connection of A_p to weak and strong type estimates	41
5.	Fourier transform	45
5.1.	On rapidly decreasing functions	45
5.2.	On L^1	53
5.3.	On L^2	54
5.4.	On L^p, $1 < p < 2$	55
6.	Singular integrals	57
6.1.	Hilbert transform	58
6.2.	Riesz transform	77
6.3.	Singular integrals of convolution type	82
6.4.	Singular integrals of nonconvolution type: Calderón-Zygmund operators	99

Date: 2.12.2010.

1. Introduction

This lecture note contains a sketch of the lectures. More illustrations and examples are presented during the lectures.

The tools of the harmonic analysis have a wide spectrum of applications in mathematical theory. The theory has strong real world applications at the background as well:

- Signal processing: Fourier transform, Fourier multipliers, Singular integrals.
- Solving PDEs: Poisson integral, Hilbert transform, Singular integrals.
- Regularity of PDEs: Hardy-Littlewood maximal function, approximation by convolution, Calderón-Zygmund decomposition, BMO.

Example 1.1. *We consider a problem*

$$\Delta u = f \quad \text{in} \quad \mathbf{R}^n$$

where $f \in L^p(\mathbf{R}^n)$. The solution u is of the form

$$u(x) = C \int_{\mathbf{R}^n} \frac{f(y)}{|x-y|^{n-2}} \, \mathrm{d}y.$$

One of the questions in the regularity theory of PDEs is, does u have the second derivatives in L^p i.e.

$$\frac{\partial^2 u}{\partial x_i \partial x_j} \in L^p(\mathbf{R}^n)?$$

If we formally differentiate u, we get

$$\frac{\partial^2 u}{\partial x_i \partial x_j} = C \int_{\mathbf{R}^n} f(y) \underbrace{\frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{|x-y|^{n-2}}}_{|\cdot| \le C/|x-y|^n} \, \mathrm{d}y.$$

It follows that $\int_{\mathbf{R}^n} f(y) \frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{|x-y|^{n-2}} \, \mathrm{d}y$ defines a singular integral $Tf(x)$. A typical theorem in the theory of singular integrals says

$$||Tf||_p \le C \, ||f||_p$$

and thus we can deduce that $\frac{\partial^2 u}{\partial x_i \partial x_j} \in L^p(\mathbf{R}^n)$.

Example 1.2. *Suppose that we have three different signals f_1, f_2, f_3 with different frequencies but only one channel, and that we receive*

$$f = f_1 + f_2 + f_3$$

from the channel. The Fourier transform $\mathcal{F}(f)$ gives us a spectrum of the signal f with three spikes in $|\mathcal{F}(f)|$. We would like to recover the *signal f_1. Thus we take a multiplier (filter)*

$$a_1(y) := \chi_{(a,b)}(y) = \begin{cases} 1, & y \in (a,b), \\ 0, & \textit{otherwise}, \end{cases}$$

where the interval (a,b) contains the frequency of f_1. Thus formally by taking the inverse Fourier transform, we get

$$f_1 = \mathcal{F}^{-1}(a_1 \mathcal{F}(f)) =: Tf(x).$$

This, again formally, defines an operator T which turns out to be of the form

$$c \int_{\mathbf{R}} \frac{\sin(Cy)}{y} f(x-y)\,\mathrm{d}y$$

with some constants c, C. This operator is of a convolution type. However, $\sin(Cy)/y$ is not integrable over the whole $\mathbf{R}$, so this requires some care!

2. Hardy-Littlewood maximal function

Definition 2.1. Let $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ and m a Lebesgue measure. A Hardy-Littlewood maximal function $Mf : \mathbf{R}^n \mapsto [0,\infty]$ is

$$Mf(x) = \sup_{Q \ni x} \frac{1}{m(Q)} \int_Q |f(y)|\,\mathrm{d}y =: \sup_{Q \ni x} \fint_Q |f(y)|\,\mathrm{d}y,$$

where the supremum is taken over all the cubes Q with sides parallel to the coordinate axis and that contain the point x. Above we used the shorthand notation

$$\fint_Q f(x)\,\mathrm{d}x = \frac{1}{m(Q)} \int_Q f(x)\,\mathrm{d}x$$

for the integral average.

Notation 2.2. We denote an open cube by

$$Q = Q(x,l) = \{y \in \mathbf{R}^n \; : \; \max_{1 \le i \le n} |y_i - x_i| < l/2\},$$

$$l(Q) \text{ is a side length of the cube } Q,$$

$$m(Q) = l(Q)^n,$$

$$\mathrm{diam}(Q) = l(Q)\sqrt{n}.$$

Example 2.3. $f : \mathbf{R} \to \mathbf{R}$, $f(x) = \chi_{(0,1)}(x)$

$$Mf(x) = \begin{cases} \frac{1}{x}, & x > 1, \\ 1, & 0 \le x \le 1, \\ \frac{1}{1-x}, & x < 0. \end{cases}$$

Observe that $f \in L^1(\mathbf{R})$ but $Mf \notin L^1(\mathbf{R})$.

Remark 2.4. (i) Mf is defined at every point $x \in \mathbf{R}^n$ and if $f = g$ almost everywhere (a.e.), then $Mf(x) = Mg(x)$ at every $x \in \mathbf{R}^n$.

(ii) It may well be that $Mf = \infty$ for every $x \in \mathbf{R}^n$. Let for example $n = 1$ and $f(x) = x^2$.

(iii) There are several definitions in the literature which are often equivalent. Let

$$\tilde{M}f(x) = \sup_{l>0} ⨍_{Q(x,l)} |f(y)| \, \mathrm{d}y,$$

where the supremum is taken over all cubes $Q(x,l)$ centered at x. Then clearly

$$\tilde{M}f(x) \leq Mf(x)$$

for all $x \in \mathbf{R}^n$. On the other hand, if Q is a cube such that $x \in Q$, then $Q = Q(x_0, l_0) \subset Q(x, 2l_0)$ and

$$\begin{aligned} ⨍_Q |f(x)| \, \mathrm{d}y &\leq \frac{m(Q(x,2l_0))}{m(Q(x,l_0))} \frac{1}{m(Q(x,2l_0))} \int_{Q(x,2l_0)} |f(y)| \, \mathrm{d}y \\ &\leq 2^n \tilde{M}f(x) \end{aligned}$$

because

$$\frac{m(Q(x,2l_0))}{m(Q(x,l_0))} = \frac{(2l_0)^n}{l_0^n} = 2^n.$$

It follows that $Mf(x) \leq 2^n \tilde{M}f(x)$ and

$$\tilde{M}f(x) \leq Mf(x) \leq 2^n \tilde{M}f(x)$$

for every $x \in \mathbf{R}^n$. We obtain a similar result, if cubes are replaced for example with balls.

Next we state some immediate properties of the maximal function. The proofs are left for the reader.

Lemma 2.5. *Let $f, g \in L^1_{loc}(\mathbf{R}^n)$. Then*

(i)

$$Mf(x) \geq 0 \text{ for all } x \in \mathbf{R}^n \text{ (positivity).}$$

(ii)

$$M(f+g)(x) \leq Mf(x) + Mg(x) \text{ (sublinearity)}$$

(iii)

$$M(\alpha f)(x) = |\alpha| \, Mf(x), \; \alpha \in \mathbf{R} \text{ (homogeneity).}$$

(iv)

$$M(\tau_y f) = (\tau_y Mf)(x) = Mf(x+y) \text{ (translation invariance).}$$

Lemma 2.6. *If $f \in C(\mathbf{R}^n)$, then*

$$|f(x)| \leq Mf(x)$$

for all $x \in \mathbf{R}^n$.

Proof. Let $f \in C(\mathbf{R}^n)$, $x \in \mathbf{R}^n$. Then

$$\forall \varepsilon > 0 \ \exists \delta > 0 \text{ s.t. } |f(x) - f(y)| < \varepsilon \text{ whenever } |x - y| < \delta.$$

From this and the triangle inequality, it follows that

$$\left| \fint_Q |f(x)| \, \mathrm{d}y - |f(x)| \right| \overset{\fint_Q 1 \, \mathrm{d}y = 1}{=} \left| \fint_Q \Big(|f(y)| - |f(x)| \Big) \, \mathrm{d}y \right|$$
$$\leq \fint_Q ||f(y)| - |f(x)|| \, \mathrm{d}y \leq \fint_Q |f(y) - f(x)| \, \mathrm{d}y < \varepsilon$$

whenever $\operatorname{diam}(Q) = \sqrt{n}\, l(Q) < \delta$. Thus

$$|f(x)| = \lim_{Q \ni x, l(Q) \to 0} \fint_Q |f(x)| \, \mathrm{d}y \leq \sup_{Q \ni x} \fint_Q |f(x)| \, \mathrm{d}y = Mf(x). \qquad \square$$

Remember that $f : \mathbf{R}^n \to [-\infty, \infty]$ is *lower semicontinuous* if

$$\{x \in \mathbf{R}^n \, : \, f(x) > \lambda\} = f^{-1}((\lambda, \infty])$$

is open for all $\lambda \in \mathbf{R}$. Thus for example, χ_U is lower semicontinuous whenever $U \subset \mathbf{R}^n$ is open. It also follows that if f is lower semicontinuous then it is measurable.

Lemma 2.7. *Mf is lower semicontinuous and thus measurable.*

Proof. We denote

$$E_\lambda = \{x \in \mathbf{R}^n \, : \, Mf(x) > \lambda\}, \ \lambda > 0.$$

Whenever $x \in E_\lambda$ it follows that there exists $Q \ni x$ such that

$$\fint_Q |f(y)| \, \mathrm{d}y > \lambda.$$

Further

$$Mf(z) \geq \fint_Q |f(y)| \, \mathrm{d}y > \lambda$$

for every $z \in Q$, and thus

$$Q \subset E_\lambda. \qquad \square$$

Lemma 2.8. *If $f \in L^\infty(\mathbf{R}^n)$, then $Mf \in L^\infty(\mathbf{R}^n)$ and*

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

Proof.

$$\fint_{Q(x)} |f(y)| \, \mathrm{d}y \leq ||f||_\infty \fint_Q 1 \, \mathrm{d}x = ||f||_\infty ,$$

for every $x \in \mathbf{R}^n$. From this it follows that

$$||Mf||_\infty \leq ||f||_\infty . \qquad \square$$

Lemma 2.9. *Let E be a measurable set. Then for each $0 < p < \infty$, we have*

$$\int_E |f(x)|^p \, \mathrm{d}x = p \int_0^\infty \lambda^{p-1} m(\{x \in E \,:\, |f(x)| > \lambda\}) \, \mathrm{d}\lambda$$

Proof. Sketch:

$$\begin{aligned}\int_E |f(x)|^p \, \mathrm{d}x &= \int_{\mathbf{R}^n} \chi_E(x) p \int_0^{|f(x)|} \lambda^{p-1} \, \mathrm{d}\lambda \, \mathrm{d}x \\ &\overset{\text{Fubini}}{=} p \int_0^\infty \lambda^{p-1} \int_{\mathbf{R}^n} \chi_{\{x \in E \,:\, |f(x)| > \lambda\}}(x) \, \mathrm{d}x \, \mathrm{d}\lambda \\ &= p \int_0^\infty \lambda^{p-1} m(\{x \in E \,:\, |f(x)| > \lambda\}) \, \mathrm{d}\lambda. \qquad \square\end{aligned}$$

Definition 2.10. Let $f : \mathbf{R}^n \to [-\infty, \infty]$ be measurable. The function f belongs to weak $L^1(\mathbf{R}^n)$ if there exists a constant C such that $0 \leq C < \infty$ such that

$$m(\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}) \leq \frac{C}{\lambda}$$

for all $\lambda > 0$.

7.9.2010

Remark 2.11. (i) $L^1(\mathbf{R}^n) \subset$ weak $L^1(\mathbf{R}^n)$ because

$$\begin{aligned}m(\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}) &= \int_{\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}} 1 \, \mathrm{d}x \\ &\leq \int_{\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}} \underbrace{\frac{|f(x)|}{\lambda}}_{\geq 1} \, \mathrm{d}x \leq \frac{||f||_1}{\lambda},\end{aligned}$$

for every $\lambda > 0$.

(ii) weak $L^1(\mathbf{R}^n)$ is not included into $L^1(\mathbf{R}^n)$. This can be seen by considering

$$f : \mathbf{R}^n \to [0, \infty], \; f(x) = |x|^{-n}.$$

Indeed,

$$\begin{aligned}\int_{B(0,1)} |f(x)| \, \mathrm{d}x = \int_{B(0,1)} |x|^{-n} \, \mathrm{d}x &= \int_0^1 \int_{\partial B(0,r)} r^{-n} \, \mathrm{d}S(x) \, \mathrm{d}r \\ &= \int_0^1 r^{-n} \underbrace{\int_{\partial B(0,r)} 1 \, \mathrm{d}S(x)}_{\omega_{n-1} r^{n-1}} \, \mathrm{d}r \\ &= \omega_{n-1} \int_0^1 \frac{1}{r} \, \mathrm{d}r = \infty,\end{aligned}$$

that is $||f||_1 = \infty$ and thus $f \notin L^1(\mathbf{R}^n)$. On the other hand for every $\lambda > 0$

$$m(\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}) = m(B(0, \lambda^{-1/n})) = \frac{\Omega_n}{\lambda}$$

where Ω_n is a measure of a unit ball. Hence $f \in$ weak $L^1(\mathbf{R}^n)$.

Theorem 2.12 (Hardy-Littlewood I)**.** *If $f \in L^1(\mathbf{R}^n)$, then Mf is in weak $L^1(\mathbf{R}^n)$ and*

$$m(\{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}) \leq \frac{5^n}{\lambda} ||f||_1$$

for every $0 < \lambda < \infty$.

In other words, the maximal functions maps L^1 to weak L^1.

The proof of this theorem uses the Vitali covering theorem.

Theorem 2.13 (Vitali covering)**.** *Let $\mathcal{F}$ be a family of cubes Q s.t.*

$$\operatorname{diam}(\bigcup_{Q \in \mathcal{F}} Q) < \infty.$$

Then there exist a countable number of disjoint cubes $Q_i \in \mathcal{F}$, $i = 1, 2, \ldots$ s.t.

$$\bigcup_{Q \in \mathcal{F}} Q \subset \bigcup_{i=1}^{\infty} 5Q_i$$

Here $5Q_i$ is a cube with the same center as Q_i whose side length is multiplied by 5.

Proof. The idea is to choose cubes inductively at round i by first throwing away the ones intersecting the cubes $Q_1, \ldots, Q_{i-1}$ chosen at the earlier rounds and then choosing the largest of the remaining cubes not yet chosen. Because the largest cube was chosen at every round, it follows that $\cup_{j=1}^{i-1} 5Q_j$ will cover the cubes thrown away. However, implementing this intuitive idea requires some care because there can be infinitely many cubes in the family $\mathcal{F}$. In particular, it may not be possible to choose largest one, but we choose almost the largest one.

To work out the details, suppose that $Q_1, \ldots, Q_{i-1} \in \mathcal{F}$ are chosen. Define

$$l_i = \sup\{l(Q) \,:\, Q \in \mathcal{F} \text{ and } Q \cap \bigcup_{j=1}^{i-1} Q_j = \emptyset\}. \tag{2.14}$$

Observe first that $l_i < \infty$, due to $\operatorname{diam}(\bigcup_{Q \in \mathcal{F}} Q) < \infty$. If there is no a cube $Q \in \mathcal{F}$ such that

$$Q \cap \bigcup_{j=1}^{i-1} Q_j = \emptyset,$$

then the process will end and we have found the cubes $Q_1, \dots, Q_{i-1}$. Otherwise we choose $Q_i \in \mathcal{F}$ such that

$$l(Q_i) > \frac{1}{2} l_i \quad \text{and} \quad Q_i \cap \bigcup_{j=1}^{i-1} Q_j = \emptyset.$$

This is also how we choose the first cube. Observe further that this is possible since $0 < l_i < \infty$. We have chosen the cubes so that they are disjoint and it suffices to show the covering property.

Choose an arbitrary $Q \in \mathcal{F}$. Then it follows that this Q intersects at least one of the chosen cubes $Q_1, Q_2, \dots$, because otherwise

$$Q \cap Q_i = \emptyset \quad \text{for every} \quad i = 1, 2, \dots$$

and thus the sup in (2.14) must be at least $l(Q)$ so that

$$l_i \geq l(Q) \quad \text{for every} \quad i = 1, 2, \dots.$$

It follows that

$$l(Q_i) > \frac{1}{2} l_i \geq \frac{1}{2} l(Q) > 0$$

for every $i = 1, 2, \dots$, so that

$$m(\bigcup_i^{\infty} Q_i) = \sum_{i=1}^{\infty} m(Q_i) = \infty,$$

where we also used the fact that the cubes are disjoint. This contradicts the fact that $m(\bigcup_i^{\infty} Q_i) < \infty$ since $\bigcup_i^{\infty} Q_i$ is a bounded set according to assumption $\operatorname{diam}(\bigcup_{Q \in \mathcal{F}} Q) < \infty$. Thus we have shown that Q intersects a cube in Q_i, $i = 1, 2, \dots$. Then there exists a smallest index i so that

$$Q \cap Q_i \neq \emptyset.$$

implying

$$Q \cap \bigcup_{j=1}^{i-1} Q_j = \emptyset.$$

Furthermore, according to the procedure

$$l(Q) \leq l_i < 2l(Q_i)$$

and thus $Q \subset 5Q_i$ and moreover

$$\bigcup_{Q \in \mathcal{F}} Q \subset \bigcup_{i=1}^{\infty} 5Q_i. \qquad \square$$

Proof of Theorem 2.12. Remember the notation

$$E_\lambda = \{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}, \quad \lambda > 0$$

so that $x \in E_\lambda$ implies that there exits a cube $Q_x \ni x$ such that

$$\fint_{Q_x} |f(y)| \,\mathrm{d}y > \lambda \tag{2.15}$$

If Q_x would cover E_λ, then the result would follow by the estimate

$$m(E_\lambda) \le m(Q) \le \int_{\mathbf{R}^n} \frac{|f(y)|}{\lambda} \,\mathrm{d}y.$$

However, this is not usually the case so we have to cover E_λ with cubes. But then the overlap of cubes needs to be controlled, and here we utilize the Vitali covering theorem.

In application of the Vitali covering theorem, there is also a technical difficulty that E_λ may not be bounded. This problem is treated by looking at the

$$E_\lambda \cap B(0,k).$$

Let $\mathcal{F}$ be a collection of cubes with the property (2.15), and $x \in E_\lambda \cap B(0,k)$. Now for every $Q \in \mathcal{F}$ it holds that

$$l(Q)^n = m(Q) < \frac{1}{\lambda} \int_Q |f(y)| \,\mathrm{d}y \le \frac{||f||_1}{\lambda},$$

so that

$$l(Q) \le \Big(\frac{||f||_1}{\lambda}\Big)^{1/n} < \infty.$$

Thus $\operatorname{diam}(\bigcup_{Q\in\mathcal{F}} Q) < \infty$ and the Vitali covering theorem implies

$$\bigcup_{Q\in\mathcal{F}} Q \subset \bigcup_{i=1}^{\infty} 5Q_i.$$

Combining the facts, we have

$$\begin{aligned}
m(E_\lambda \cap B(0,k)) &\le m(\bigcup_{Q\in\mathcal{F}}^{\infty} Q) \le \sum_{i=1}^{\infty} m(5Q_i) = 5^n \sum_{i=1}^{\infty} m(Q_i) \\
&\overset{(2.15)}{\le} \frac{5^n}{\lambda} \sum_{i=1}^{\infty} \int_{Q_i} |f(y)| \,\mathrm{d}y \\
&\overset{\text{cubes are disjoint}}{=} \frac{5^n}{\lambda} \int_{\cup_{i=1}^{\infty} Q_i} |f(y)| \,\mathrm{d}y \le \frac{5^n}{\lambda} ||f||_1 .
\end{aligned}$$

Then we pass to the original E_λ

$$m(E_\lambda) = \lim_{k\to\infty} m(E_\lambda \cap B(0,k)) \le \frac{5^n}{\lambda} ||f||_1 . \qquad \square$$

Remark 2.16. Observe that $f \in L^1(\mathbf{R}^n)$ implies that $Mf(x) < \infty$ a.e. $x \in \mathbf{R}^n$ because

$$\begin{aligned} m(\{x \in \mathbf{R}^n : Mf(x) = \infty\} &\le m(\{x \in \mathbf{R}^n : Mf(x) > \lambda\}) \\ &\le \frac{5^n}{\lambda} \|f\|_1 \to 0 \end{aligned}$$

as $\lambda \to \infty$.

Definition 2.17. (i)

$$f \in L^1(\mathbf{R}^n) + L^p(\mathbf{R}^n), \qquad 1 \le p \le \infty$$

if

$$f = g + h, \quad g \in L^1(\mathbf{R}^n), \quad h \in L^p(\mathbf{R}^n)$$

(ii)

$$T : L^1(\mathbf{R}^n) + L^p(\mathbf{R}^n) \to \text{ measurable functions}$$

is subadditive, if

$$|T(f+g)(x)| \le |Tf(x)| + |Tg(x)| \text{ a.e. } x \in \mathbf{R}^n$$

(iii) T is of strong type (p,p), $1 \le p \le \infty$, if there exists a constant C independent of functions $f \in L^p(\mathbf{R}^n)$ s.t.

$$\|Tf\|_p \le C \|f\|_p .$$

for every $f \in L^p(\mathbf{R}^n)$

(iv) T is of weak type (p,p), $1 \le p < \infty$, if there exists a constant C independent of functions $f \in L^p(\mathbf{R}^n)$ s.t.

$$m(\{x \in \mathbf{R}^n : Tf(x) > \lambda\}) \le \frac{C}{\lambda^p} \|f\|_p^p$$

for every $f \in L^p(\mathbf{R}^n)$.

Remark 2.18. (i) Observe that the maximal operator is subadditive, of weak type (1,1) that is

$$m(\{x \in \mathbf{R}^n : Mf(x) > \lambda\}) \le \frac{5^n}{\lambda} \|f\|_1 ,$$

of strong type (∞, ∞)

$$\|Mf\|_\infty \le C \|f\|_\infty ,$$

and *nonlinear*.

(ii) Strong (p,p) implies weak (p,p):

$$\begin{aligned} m(\{x \in \mathbf{R}^n : Tf(x) > \lambda\}) &\overset{\text{Chebysev}}{\le} \frac{1}{\lambda^p} \int_{\mathbf{R}^n} |Tf|^p \, \mathrm{d}x \\ &\overset{\text{strong } (p,p)}{\le} \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f|^p \, \mathrm{d}x. \end{aligned}$$

Theorem 2.19 (Hardy-Littlewood II)**.** *If $f \in L^p(\mathbf{R}^n)$, $1 < p \leq \infty$, then $Mf \in L^p(\mathbf{R}^n)$ and there exists $C = C(n,p)$ (meaning C depends on n, p) such that*

$$||Mf||_p \leq C\, ||f||_p\, .$$

This is not true, when $p = 1$, cf. Example 2.3. The proof is based on the interpolation (Marcinkiewicz interpolation theorem, proven below) between weak $(1,1)$ and strong (∞,∞). In the proof of the Marcinkiewicz interpolation theorem, we use the following auxiliary lemma. 9.9.2010

Lemma 2.20. *Let $1 \leq p \leq q \leq \infty$. Then*

$$L^p(\mathbf{R}^n) \subset L^1(\mathbf{R}^n) + L^q(\mathbf{R}^n).$$

Proof. Let $f \in L^p(\mathbf{R}^n)$, $\lambda > 0$. We split f into two part as $f = f_1 + f_2$ by setting

$$f_1(x) = f\chi_{\{x \in \mathbf{R}^n \,:\, |f(x)| \leq \lambda\}}(x) = \begin{cases} f(x), & |f(x)| \leq \lambda \\ 0, & |f(x)| > \lambda, \end{cases}$$

$$f_2(x) = f\chi_{\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}}(x) = \begin{cases} f(x), & |f(x)| > \lambda \\ 0, & |f(x)| \leq \lambda. \end{cases}$$

We will show that $f_1 \in L^q$ and $f_2 \in L^1$

$$\begin{aligned} \int_{\mathbf{R}^n} |f_1(x)|^q \,\mathrm{d}x &= \int_{\mathbf{R}^n} |f_1(x)|^{q-p}\, |f_1(x)|^p \,\mathrm{d}x \\ &\overset{|f_1| \leq \lambda}{\leq} \lambda^{q-p} \int_{\mathbf{R}^n} |f_1(x)|^p \,\mathrm{d}x \\ &\overset{|f_1| \leq |f|}{\leq} \lambda^{q-p}\, ||f||_p^p < \infty, \end{aligned}$$

$$\begin{aligned} \int_{\mathbf{R}^n} |f_2(x)| \,\mathrm{d}x &= \int_{\mathbf{R}^n} |f_2|^{1-p}\, |f_2|^p \,\mathrm{d}x \\ &\overset{|f_2| > \lambda \text{ or } f_2 = 0}{\leq} \lambda^{1-p} \int_{\mathbf{R}^n} |f_2|^p \,\mathrm{d}x \\ &\overset{|f_2| \leq |f|}{\leq} \lambda^{1-p}\, ||f||_p^p < \infty. \end{aligned}$$

□

Theorem 2.21 (Marcinkiewicz interpolation theorem)**.** *Let $1 < q \leq \infty$,*

$$T : L^1(\mathbf{R}^n) + L^q(\mathbf{R}^n) \to \textit{measurable functions}$$

is subadditive, and

(i) T is of weak type $(1,1)$
(ii) T is of weak type (q,q), if $q < \infty$, and
T is of strong type (q,q), if $q = \infty$.

Then T is of strong type (p,p) for every $1 < p < q$ that is

$$||Tf||_p \le C\,||f||_p$$

for every $f \in L^p(\mathbf{R}^n)$.

Proof. **Case $q < \infty$.** Let $f = f_1 + f_2$ where as before

$$f_1 = f\chi_{\{|f|\le\lambda\}} \quad \text{and} \quad f_2 = f\chi_{\{|f|>\lambda\}}$$

and recall that $f_1 \in L^q$ and $f_2 \in L^1$. Subadditivity implies

$$|Tf| \le |Tf_1| + |Tf_2|$$

for a.e. $x \in \mathbf{R}^n$. Thus

$$\begin{aligned}
m(\{x \in \mathbf{R}^n \,:\, |Tf(x)| > \lambda\}) &\le m(\{x \in \mathbf{R}^n \,:\, |Tf_1(x)| > \lambda/2\}) \\
&\quad + m(\{x \in \mathbf{R}^n \,:\, |Tf_2(x)| > \lambda/2\}) \\
&\le \Big(\frac{C_1}{\lambda/2}\,||f_1||_q\Big)^q + \frac{C_2}{\lambda/2}\,||f_2||_1 \\
&\le \frac{(2C_1)^q}{\lambda^q}\int_{\{x\in\mathbf{R}^n\,:\,|f(x)|\le\lambda\}} |f(x)|^q \,\mathrm{d}x \\
&\quad + \frac{2C_2}{\lambda}\int_{\{x\in\mathbf{R}^n\,:\,|f(x)|>\lambda\}} |f(x)| \,\mathrm{d}x.
\end{aligned}$$

Then by Lemma 2.9, it follows that

$$\begin{aligned}
\int_{\mathbf{R}^n} |Tf|^p \,\mathrm{d}x &= p\int_0^\infty \lambda^{p-1} m(\{x \in \mathbf{R}^n \,:\, |Tf(x)| > \lambda)\,\mathrm{d}\lambda \\
&\le (2C_1)^q p \int_0^\infty \lambda^{p-q-1}\int_{\{x\in\mathbf{R}^n\,:\,|f(x)|\le\lambda\}} |f(x)|^q \,\mathrm{d}x\,\mathrm{d}\lambda \\
&\quad + 2pC_2\int_0^\infty \lambda^{p-2}\int_{\{x\in\mathbf{R}^n\,:\,|f(x)|>\lambda\}} |f(x)| \,\mathrm{d}x\,\mathrm{d}\lambda.
\end{aligned}$$

Further by Fubini's theorem

$$\begin{aligned}
\int_0^\infty \lambda^{p-q-1}\int_{\{x\in\mathbf{R}^n\,:\,|f(x)|\le\lambda\}} |f(x)|^q \,\mathrm{d}x\,\mathrm{d}\lambda &= \int_{\mathbf{R}^n} |f(x)|^q \int_{|f(x)|}^\infty \lambda^{p-q-1}\,\mathrm{d}\lambda\,\mathrm{d}x \\
&= \frac{1}{q-p}\int_{\mathbf{R}^n} |f(x)|^q\,|f(x)|^{p-q} \,\mathrm{d}x \\
&= \frac{1}{q-p}\int_{\mathbf{R}^n} |f(x)|^p \,\mathrm{d}x
\end{aligned}$$

and

$$\begin{aligned}\int_0^\infty \lambda^{p-2} \int_{\{x\in\mathbf{R}^n \,:\, |f(x)|>\lambda\}} |f(x)| \,\mathrm{d}x\,\mathrm{d}\lambda &= \int_{\mathbf{R}^n} |f(x)| \int_0^{|f(x)|} \lambda^{p-2}\,\mathrm{d}\lambda\,\mathrm{d}x \\ &= \frac{1}{p-1}\int_{\mathbf{R}^n} |f(x)|^{p-1}\,|f(x)|\,\mathrm{d}x \\ &= \frac{1}{p-1}\int_{\mathbf{R}^n} |f(x)|^p\,\mathrm{d}x.\end{aligned}$$

Thus we arrive at

$$||Tf||_p^p \le p\Big(\frac{2C_2}{p-1} + \frac{(2C_1)^q}{q-p}\Big)\,||f||_p^p\,.$$

Case $q=\infty$. Suppose that

$$||Tg||_\infty \le C_2\,||g||_\infty$$

for every $g\in L^\infty(\mathbf{R}^n)$. We again split $f\in L^p(\mathbf{R}^n)$ as $f=f_1+f_2$ where

$$f_1 = f\chi_{\{|f|\le\lambda/(2C_2)\}} \quad\text{and}\quad f_2 = f\chi_{\{|f|>\lambda/(2C_2)\}}$$

and by Lemma 2.20, $f_1\in L^\infty$ and $f_2\in L^1$. We have a.e.

$$|Tf_1(x)| \le ||Tf_1||_\infty \le C_2\,||f_1||_\infty \le C_2\frac{\lambda}{2C_2} = \frac{\lambda}{2}.$$

Thus

$$\begin{aligned}m(\{x\in\mathbf{R}^n \,:\, |Tf(x)|>\lambda\}) \le{}& \underbrace{m(\{x\in\mathbf{R}^n \,:\, |Tf_1(x)|>\lambda/2\})}_{=0} \\ &+ m(\{x\in\mathbf{R}^n \,:\, |Tf_2(x)|>\lambda/2\}).\end{aligned}$$

It follows that

$$\begin{aligned}m(\{x\in\mathbf{R}^n \,:\, |Tf(x)|>\lambda\}) &\le m(\{x\in\mathbf{R}^n \,:\, |Tf_2(x)|>\lambda/2\}) \\ &\overset{\text{weak }(1,1)}{\le} \frac{C_1}{\lambda/2}\int_{\mathbf{R}^n} |f_2(x)|\,\mathrm{d}x \\ &= \frac{2C_1}{\lambda}\int_{\{x\in\mathbf{R}^n \,:\, |f(x)|>\lambda/(2C_2)\}} |f(x)|\,\mathrm{d}x.\end{aligned}$$

Then by using Lemma 2.9 again, we see that

$$\begin{aligned}\int_{\mathbf{R}^n} |Tf(x)|^p\,\mathrm{d}x &= p\int_0^\infty \lambda^{p-1} m(\{x\in\mathbf{R}^n \,:\, |Tf(x)|>\lambda\})\,\mathrm{d}\lambda \\ &\le 2C_1 p\int_0^\infty \lambda^{p-2}\int_{\{x\in\mathbf{R}^n \,:\, |f(x)|>\lambda/(2C_2)\}} |f(x)|\,\mathrm{d}x\,\mathrm{d}\lambda \\ &\overset{\text{Fubini}}{=} 2^p C_2^{p-1} C_1 \frac{p}{p-1}\int_{\mathbf{R}^n} |f(x)|^p\,\mathrm{d}x.\end{aligned}$$

□

Example 2.22 (Proof of the Sobolev's inequality via the maximal function)**.** *Suppose that* $u \in C_0^\infty(\mathbf{R}^n)$*. We immediately have*

$$u(x) = -\int_0^\infty \frac{\partial}{\partial r} u(x + r\omega)\,\mathrm{d}r,$$

where $\omega \in \partial B(0,1)$*. Integrating this over the whole unit sphere*

$$\begin{aligned}
\omega_{n-1} u(x) &= \int_{\partial B(0,1)} u(x)\,\mathrm{d}S(\omega) \\
&= -\int_{\partial B(0,1)} \int_0^\infty \frac{\partial}{\partial r} u(x + r\omega)\,\mathrm{d}r\,\mathrm{d}S(\omega) \\
&= -\int_{\partial B(0,1)} \int_0^\infty \nabla u(x + r\omega) \cdot \omega\,\mathrm{d}r\,\mathrm{d}S(\omega) \\
&= -\int_0^\infty \int_{\partial B(0,1)} \nabla u(x + r\omega) \cdot \omega\,\mathrm{d}S(\omega)\,\mathrm{d}r
\end{aligned}$$

and changing variables so that $y = x + r\omega,\ \ \mathrm{d}S(y) = r^{n-1}\,\mathrm{d}S(\omega)$*,* $\omega = (y - x)/\left|y - x\right|, r = \left|y - x\right|$ *we get*

$$\omega_{n-1} u(x) = -\int_0^\infty \int_{\partial B(0,r)} \nabla u(y) \cdot \frac{y - x}{\left|y - x\right|^n}\,\mathrm{d}S(y)\,\mathrm{d}r$$

so that

$$u(x) = -\frac{1}{\omega_{n-1}} \int_{\mathbf{R}^n} \frac{\nabla u(y) \cdot (x - y)}{\left|x - y\right|^n}\,\mathrm{d}y.$$

Further

$$\left|u(x)\right| \leq \frac{1}{\omega_{n-1}} \int_{\mathbf{R}^n} \frac{\left|\nabla u(y)\right|}{\left|x - y\right|^{n-1}}\,\mathrm{d}y$$

which is so called Riesz potential. We split this into a bad part and a good part as $\int_{\mathbf{R}^n} = \int_{B(x,r)} + \int_{\mathbf{R}^n \setminus B(x,r)}$*. By estimating the bad part over the sets* $B(x, 2^{-i}r) \setminus B(x, 2^{-i-1}r)$ *as*

$$\begin{aligned}
\int_{B(x,r)} \frac{\left|\nabla u(y)\right|}{\left|x - y\right|^{n-1}}\,\mathrm{d}y &= \sum_{i=0}^\infty \int_{B(x,2^{-i}r)\setminus B(x,2^{-i-1}r)} \frac{\left|\nabla u(y)\right|}{\left|x - y\right|^{n-1}}\,\mathrm{d}y \\
&\leq \sum_{i=0}^\infty \int_{B(x,2^{-i}r)\setminus B(x,2^{-i-1}r)} \frac{\left|\nabla u(y)\right|}{(2^{-i-1}r)^{n-1}}\,\mathrm{d}y \\
&\leq \sum_{i=0}^\infty \frac{2^{-i}r}{2^{-i}r} \int_{B(x,2^{-i}r)} 2^{n-1} \frac{\left|\nabla u(y)\right|}{(2^{-i}r)^{n-1}}\,\mathrm{d}y \\
&\leq C \sum_{i=0}^\infty 2^{n-1} 2^{-i} r \fint_{B(x,2^{-i}r)} \left|\nabla u(y)\right|\,\mathrm{d}y \\
&\leq C 2^{n-1} r M \left|\nabla u\right|(x) \sum_{i=0}^\infty 2^{-i}
\end{aligned}$$

we get

$$\int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\,\mathrm{d}y \le CrM\,|\nabla u|\,(x). \tag{2.23}$$

On the other hand, for the good part we use Hölder's inequality with the powers p *and* $p/(p-1)$ *, where* $p<n$*, as*

$$\int_{\mathbf{R}^n\setminus B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\,\mathrm{d}y$$
$$\le \Big(\int_{\mathbf{R}^n\setminus B(x,r)} |\nabla u(y)|^p\,\mathrm{d}y\Big)^{1/p}\Big(\int_{\mathbf{R}^n\setminus B(x,r)} |x-y|^{(1-n)p/(p-1)}\,\mathrm{d}y\Big)^{(p-1)/p}.$$

Then we calculate

$$\Big(\int_{\mathbf{R}^n\setminus B(x,r)} |x-y|^{(1-n)p/(p-1)}\,\mathrm{d}y\Big)^{(p-1)/p}$$
$$=\Big(\int_r^\infty \omega_{n-1}\rho^{n-1}\rho^{(1-n)p/(p-1)}\,\mathrm{d}\rho\Big)^{(p-1)/p}$$
$$=\Big(\omega_{n-1}\int_r^\infty \rho^{(1-n)/(p-1)}\,\mathrm{d}\rho\Big)^{(p-1)/p}=\Big(\omega_{n-1}\int_r^\infty \rho^{-1+(p-n)/(p-1)}\,\mathrm{d}\rho\Big)^{(p-1)/p}.$$

Combining the previous calculations, we get 14.9.2010

$$\int_{\mathbf{R}^n\setminus B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\,\mathrm{d}y \le C\,||\nabla u||_p\, r^{1-\frac{n}{p}}, \tag{2.24}$$

with $p<n$*. Choosing* $r=\Big(||\nabla u||_p/(M\,|\nabla u|\,(x))\Big)^{p/n}$ *as well as combining the estimates* (2.23) *and* (2.24)*, we get*

$$\begin{aligned}|u(x)| &\le C\int_{\mathbf{R}^n} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\,\mathrm{d}y\\ &\le C\,||\nabla u||_p^{p/n}\, M\,|\nabla u|\,(x)^{(n-p)/n}.\end{aligned}$$

Then we take the power[1] $np/(n-p)$ *on both sides and end up with*

$$|u(x)|^{np/(n-p)} \le C\,||\nabla u||_p^{p^2/(n-p)}\, M\,|\nabla u|\,(x)^p.$$

By recalling Hardy-Littlewood II, we obtain

$$\begin{aligned}\int_{\mathbf{R}^n} |u(x)|^{np/(n-p)}\,\mathrm{d}x &\le C\,||\nabla u||_p^{p^2/(n-p)}\int_{\mathbf{R}^n} M\,|\nabla u|\,(x)^p\,\mathrm{d}x\\ &\le C\,||\nabla u||_p^{p^2/(n-p)}\,||\nabla u||_p^p \le C\,||\nabla u||_p^{np/(n-p)}.\end{aligned}$$

This is so called Sobolev's inequality

$$\Big(\int_{\mathbf{R}^n} |u(x)|^{p^*}\,\mathrm{d}x\Big)^{1/p^*} \le C\Big(\int_{\mathbf{R}^n} |\nabla u(x)|^p\,\mathrm{d}x\Big)^{1/p},$$

which holds for every $u\in C_0^\infty(\mathbf{R}^n)$ *and* $p<n$*.*

[1] This is sometimes denoted by $p^*=np/(n-p)$ and called a Sobolev conjugate. It satisfies $1/p+1/p^*=1/n$.

3. Approximation by convolution

Definition 3.1 (Convolution)**.** Suppose that $f, g : \mathbf{R}^n \to [-\infty, \infty]$ are Lebesgue-measurable functions. The convolution

$$(f * g)(x) = \int_{\mathbf{R}^n} f(y)g(x-y)\,\mathrm{d}y$$

is defined if $y \mapsto f(y)g(x-y)$ is integrable for almost every $x \in \mathbf{R}^n$.

Observe that: $f, g \in L^1(\mathbf{R}^n)$ does not imply $fg \in L^1(\mathbf{R}^n)$ which can be seen by considering for example $f = g = \frac{\chi_{(0,1)}(x)}{\sqrt{x}}$.

Theorem 3.2 (Minkowski's/Young's inequality)**.** *If $f \in L^p(\mathbf{R}^n)$, $1 \le p \le \infty$ and $g \in L^1(\mathbf{R}^n)$, then $(f * g)(x)$ exists for almost all $x \in \mathbf{R}^n$ and*

$$||f * g||_p \le ||f||_p \, ||g||_1 \, .$$

Proof. **Case $p = 1$:** Because

$$|(f * g)(x)| \le \int_{\mathbf{R}^n} |f(y)|\,|g(x-y)|\,\mathrm{d}y$$

we have

$$\begin{aligned}
\int_{\mathbf{R}^n} |(f * g)(x)|\,\mathrm{d}x &\le \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} |f(y)|\,|g(x-y)|\,\mathrm{d}y\,\mathrm{d}x \\
&\overset{\text{Fubini}}{=} \int_{\mathbf{R}^n} |f(y)| \Big(\int_{\mathbf{R}^n} |g(x-y)|\,\mathrm{d}x \Big)\,\mathrm{d}y \\
&= \int_{\mathbf{R}^n} |f(y)|\,\mathrm{d}y \int_{\mathbf{R}^n} |g(x)|\,\mathrm{d}x \\
&= ||f||_1 \, ||g||_1 \, .
\end{aligned}$$

Case $p = \infty$:

$$\begin{aligned}
|(f * g)(x)| &\le \int_{\mathbf{R}^n} |f(y)|\,|g(x-y)|\,\mathrm{d}y \\
&\le \operatorname*{ess\,sup}_{y \in \mathbf{R}^n} |f(x)| \int_{\mathbf{R}^n} |g(x-y)|\,\mathrm{d}y \\
&= ||f||_\infty \, ||g||_1 \, .
\end{aligned}$$

Case $1 < p < \infty$: Set

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Then

$$\begin{aligned}
|(f*g)(x)| &\le \int_{\mathbf{R}^n} |f(y)|\,|g(x-y)| \,\mathrm{d}y \\
&= \int_{\mathbf{R}^n} |f(y)|\,|g(x-y)|^{1/p}\,|g(x-y)|^{1/p'} \,\mathrm{d}y \\
&\overset{\text{Hölder}}{\le} \Big(\int_{\mathbf{R}^n} |f(y)|^p\,|g(x-y)| \,\mathrm{d}y\Big)^{1/p}\Big(\int_{\mathbf{R}^n} |g(x-y)| \,\mathrm{d}y\Big)^{1/p'} \\
&= \Big(\int_{\mathbf{R}^n} |f(y)|^p\,|g(x-y)| \,\mathrm{d}y\Big)^{1/p} ||g||_1^{1/p'}.
\end{aligned}$$

Thus

$$\begin{aligned}
\int_{\mathbf{R}^n} |(f*g)(x)|^p \,\mathrm{d}x &\le ||g||_1^{p/p'} \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} |f(y)|^p\,|g(x-y)| \,\mathrm{d}y\,\mathrm{d}x \\
&\overset{\text{Fubini}}{=} ||g||_1^{p/p'} \int_{\mathbf{R}^n} |f(y)|^p \int_{\mathbf{R}^n} |g(x-y)| \,\mathrm{d}x\,\mathrm{d}y \\
&= ||g||_1^{p/p'}\,||g||_1\,||f||_p^p = ||g||_1^p\,||f||_p^p,
\end{aligned}$$

because

$$\frac{p}{p'} + 1 = p(\frac{1}{p'} + \frac{1}{p}) = p. \qquad \square$$

We state the following simple properties of convolution without a proof.

Lemma 3.3 (Basic properties of convolution)**.** *Let $f, g, h \in L^1(\mathbf{R}^n)$. Then*

(i) $f * g = g * f$.
(ii) $f * (g * h) = (f * g) * h$.
(iii) $(\alpha f + \beta g) * h = \alpha(f * h) + \beta(g * h),\ \alpha, \beta \in \mathbf{R}^n$.

For $\phi \in L^1(\mathbf{R}^n)$, $\varepsilon > 0$, we denote

$$\phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\phi(\frac{x}{\varepsilon}),\ x \in \mathbf{R}^n. \tag{3.4}$$

Example 3.5. *(i) Let $\phi(x) = \frac{\chi_{B(0,1)}(x)}{m(B(0,1))}$. Then*

$$\phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\frac{\chi_{B(0,1)}(\frac{x}{\varepsilon})}{m(B(0,1))} = \frac{\chi_{B(0,\varepsilon)(x)}}{m(B(0,\varepsilon))}.$$

Then for $f \in L^1(\mathbf{R}^n)$, a mollification

$$\begin{aligned}
(f * \phi_\varepsilon)(x) &= \int_{\mathbf{R}^n} f(y)\phi_\varepsilon(x-y)\,\mathrm{d}y \\
&= \fint_{B(x,\varepsilon)} f(y)\,\mathrm{d}y.
\end{aligned}$$

turns out to be useful. Observe also that $||\phi_\varepsilon||_1 = 1$ *for any* $\varepsilon > 0$ *so that*

$$||f * \phi_\varepsilon||_1 \leq ||f||_1 \, ||\phi_\varepsilon||_1 = ||f||_1 \, .$$

(ii)

$$\varphi = \begin{cases} \exp\left(\frac{1}{|x|^2-1}\right), & x \in B(0,1) \\ 0, & \textit{else.} \end{cases}$$

It holds that $\varphi \in C_0^\infty(\mathbf{R}^n)$ *and thus also* $\varphi \in L^1(\mathbf{R}^n)$*. Let*

$$\phi = \frac{\varphi}{||\varphi||_1}.$$

Then $\phi_\varepsilon \in C_0^\infty(\mathbf{R}^n)$, $\operatorname{spt}(\phi_\varepsilon) \subset \overline{B}(0,\varepsilon)$*, and*

$$\begin{aligned} \int_{\mathbf{R}^n} \phi_\varepsilon(x) \, \mathrm{d}x &= \frac{1}{\varepsilon^n} \int_{\mathbf{R}^n} \phi(x/\varepsilon) \, \mathrm{d}x \\ &\overset{y=\frac{x}{\varepsilon},\ \mathrm{d}x=\varepsilon^n \, \mathrm{d}y}{=} \frac{1}{\varepsilon^n} \int_{\mathbf{R}^n} \phi(y) \varepsilon^n \, \mathrm{d}y \\ &= \int_{\mathbf{R}^n} \phi(y) \, \mathrm{d}y \\ &= \int_{\mathbf{R}^n} \frac{\varphi(y)}{||\varphi||_1} \, \mathrm{d}y = \frac{||\varphi||_1}{||\varphi||_1} = 1, \end{aligned}$$

for all $\varepsilon > 0$*. The function* ϕ_ε *is called a standard mollifier in this case. As before, if* $f \in L^1(\mathbf{R}^n)$*, then*

$$||f * \phi_\varepsilon||_1 \leq ||f||_1 \, .$$

Lemma 3.6. *Let* $\phi \in L^1(\mathbf{R}^n)$ *and recall that* $\phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\phi(\frac{x}{\varepsilon})$*. Then*

(i)

$$\int_{\mathbf{R}^n} \phi_\varepsilon(x) \, \mathrm{d}x = \int_{\mathbf{R}^n} \phi(x) \, \mathrm{d}x$$

for every $\varepsilon > 0$*.*

(ii)

$$\lim_{\varepsilon \to 0} \int_{\mathbf{R}^n \setminus B(0,r)} |\phi_\varepsilon(x)| \, \mathrm{d}x = 0$$

for every $r > 0$*.*

Proof. (i) Change of variables, see above.

(ii) We calculate

$$\begin{aligned}\int_{\mathbf{R}^n\setminus B(0,r)} |\phi_\varepsilon(x)| \,\mathrm{d}x &= \frac{1}{\varepsilon^n}\int_{\mathbf{R}^n\setminus B(0,r)} |\phi(x/\varepsilon)| \,\mathrm{d}x \\ &\overset{y=\frac{x}{\varepsilon},\ \mathrm{d}x=\varepsilon^n\,\mathrm{d}y}{=} \int_{\mathbf{R}^n\setminus B(0,r/\varepsilon)} \phi(y)\,\mathrm{d}y \\ &= \int_{\mathbf{R}^n} \phi(y)\chi_{\mathbf{R}^n\setminus B(0,r/\varepsilon)}\,\mathrm{d}y \to 0\end{aligned}$$

as $\varepsilon \to 0$ by Lebesgue's dominated convergence theorem. □

Theorem 3.7. *Let* $\phi \in L^1(\mathbf{R}^n)$,

$$a = \int_{\mathbf{R}^n} \phi(x)\,\mathrm{d}x$$

and $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$. *Then*

$$||\phi_\varepsilon * f - af||_p \to 0$$

as $\varepsilon \to 0$.

Notice that the statement is invalid if $p = \infty$.

Proof. We will work out the details below, but the idea in the proof is that by using the definition of the convolution together with Hölder's inequality and Fubini's theorem, we obtain

$$\begin{aligned}&\int_{\mathbf{R}^n} |(f*\phi_\varepsilon)(x) - af(x)|^p \,\mathrm{d}x \\ &\qquad\le ||\phi||_1^{p/p'} \int_{\mathbf{R}^n} |\phi_\varepsilon(y)| \Big(\int_{\mathbf{R}^n} |f(x-y)-f(x)|^p \,\mathrm{d}x\Big)\,\mathrm{d}y \\ &\qquad= ||\phi||_1^{p/p'} \int_{B(0,r)} |\phi_\varepsilon(y)| \Big(\int_{\mathbf{R}^n} |f(x-y)-f(x)|^p \,\mathrm{d}x\Big)\,\mathrm{d}y \\ &\qquad\quad + ||\phi||_1^{p/p'} \int_{\mathbf{R}^n\setminus B(0,r)} |\phi_\varepsilon(y)| \Big(\int_{\mathbf{R}^n} |f(x-y)-f(x)|^p \,\mathrm{d}x\Big)\,\mathrm{d}y \\ &\qquad= I_1 + I_2,\end{aligned} \tag{3.8}$$

where $1/p+1/p' = 1$. The first term on the right hand side, I_1, is small when r is small because intuitively then $f(x-y)$ only differs little from $f(x)$. On the other hand, the second integral, I_2, is small for small enough $\varepsilon > 0$ for any r because ϕ_ε gets more and more concentrated. 16.9.2010

Next we work out the details. By the previous lemma

$$af(x) = f(x)\int_{\mathbf{R}^n} \phi(y)\,\mathrm{d}y = \int_{\mathbf{R}^n} f(x)\phi_\varepsilon(y)\,\mathrm{d}y.$$

Thus

$$\begin{aligned}
&\int_{\mathbf{R}^n} |(f*\phi_\varepsilon)(x) - af(x)|^p \,\mathrm{d}x \\
&\quad = \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} (f(x-y)-f(x))\phi_\varepsilon(y)\,\mathrm{d}y \right|^p \mathrm{d}x \\
&\quad \leq \int_{\mathbf{R}^n} \Big(\int_{\mathbf{R}^n} |f(x-y)-f(x)|\, |\phi_\varepsilon(y)|^{1/p}\, |\phi_\varepsilon(y)|^{1/p'} \,\mathrm{d}y \Big)^p \mathrm{d}x \\
&\quad \overset{\text{Hölder}}{\leq} \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} |(f(x-y)-f(x))|^p \,|\phi_\varepsilon(y)| \,\mathrm{d}y \Big(\int_{\mathbf{R}^n} |\phi_\varepsilon(y)| \,\mathrm{d}y \Big)^{p/p'} \mathrm{d}x \\
&\quad \overset{\text{Fubini}}{=} ||\phi||_1^{p/p'} \int_{\mathbf{R}^n} |\phi_\varepsilon(y)| \Big(\int_{\mathbf{R}^n} |f(x-y)-f(x)|^p \,\mathrm{d}x \Big) \mathrm{d}y.
\end{aligned}$$

This confirms (3.8), and we start estimating I_2 and I_1.

Fix $\eta > 0$. First we estimate I_1. By a well-known result in L^p-theory, $C_0(\mathbf{R}^n)$ (compactly supported continuous functions) are dense in $L^p(\mathbf{R}^n)$ meaning that we can choose $g \in C_0(\mathbf{R}^n)$ such that

$$\int_{\mathbf{R}^n} |f(x) - g(x)|^p \,\mathrm{d}x < \eta.$$

Moreover, as g is uniformly continuous because it is compactly supported, so that we can choose small enough $r > 0$ to have

$$\int_{\mathbf{R}^n} |g(x-y) - g(x)|^p \,\mathrm{d}x < \eta,$$

for any $y \in B(0,r)$. Also recall that by convexity of $x^p, p > 1$ for some $a, b \in \mathbf{R}$ we have $|a+b|^p \leq (|a| + |b|)^p = (\frac{1}{2}2\,|a| + \frac{1}{2}2\,|b|)^p \leq \frac{1}{2}(2\,|a|)^p + \frac{1}{2}(2\,|b|)^p = 2^{p-1}\,|a|^p + 2^{p-1}\,|b|^p$. By using these tools, and by adding and subtracting g, we can estimate

$$\begin{aligned}
&\int_{\mathbf{R}^n} |f(x-y) - f(x)|^p \,\mathrm{d}x \\
&\quad \leq \int_{\mathbf{R}^n} |f(x-y) - g(x-y) + g(x-y) - g(x) + g(x) - f(x)|^p \,\mathrm{d}x \\
&\quad \overset{\text{convexity}}{\leq} C \int_{\mathbf{R}^n} |f(x-y) - g(x-y)|^p \,\mathrm{d}x \\
&\qquad + C \int_{\mathbf{R}^n} |g(x-y) - g(x)|^p \,\mathrm{d}x + C \int_{\mathbf{R}^n} |g(x) - f(x)|^p \,\mathrm{d}x \leq 3\eta
\end{aligned}$$

for any $y \in B(0,r)$. Thus

$$\begin{aligned}
I_1 &= ||\phi||_1^{p/p'} \int_{B(0,r)} |\phi_\varepsilon(y)| \Big(\int_{\mathbf{R}^n} |f(x-y) - f(x)|^p \,\mathrm{d}x \Big) \mathrm{d}y \\
&\leq ||\phi||_1^{p/p'} \int_{B(0,r)} |\phi_\varepsilon(y)|\, 3\eta \,\mathrm{d}y \leq C\eta.
\end{aligned}$$

Next we estimate I_2. By the previous lemma (Lemma 3.6 (ii)), for any $r > 0$, there exists $\varepsilon' > 0$ such that

$$\int_{\mathbf{R}^n \setminus B(0,r)} |\phi_\varepsilon(y)| \, \mathrm{d}y < \eta,$$

for every $0 < \varepsilon < \varepsilon'$. Thus since

$$\begin{aligned}\int_{\mathbf{R}^n} |f(x-y) - f(x)|^p \, \mathrm{d}x \leq & 2^{p-1} \int_{\mathbf{R}^n} |f(x-y)|^p \, \mathrm{d}x \\ & + 2^{p-1} \int_{\mathbf{R}^n} |f(x)|^p \, \mathrm{d}x < \infty\end{aligned}$$

for $f \in L^p$, we see that

$$\begin{aligned} I_2 &= ||\phi||_1^{p/p'} \int_{\mathbf{R}^n \setminus B(0,r)} |\phi_\varepsilon(y)| \Big(\int_{\mathbf{R}^n} |f(x-y) - f(x)|^p \, \mathrm{d}x \Big) \, \mathrm{d}y \\ &\leq C \int_{\mathbf{R}^n \setminus B(0,r)} |\phi_\varepsilon(y)| \, \mathrm{d}y < C\eta, \end{aligned}$$

where $C = ||\phi||_1^{p/p'} \, 2^p \, ||f||_p^p$. Thus for any $\eta > 0$ we get an estimate

$$\int_{\mathbf{R}^n} |(f * \phi_\varepsilon)(x) - af(x)|^p \, \mathrm{d}x \leq I_1 + I_2 \leq C\eta$$

with C independent of η, by first choosing small enough r so that I_1 is small, and then for this fixed $r > 0$ by choosing ε small enough so that I_2 is small. □

Remark 3.9. Similarly, we can prove that for $\phi \in L^1(\mathbf{R}^n)$ and $a = \int_{\mathbf{R}^n} \phi \, \mathrm{d}x$, we have

(i) If $f \in C(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$, then

$$f * \phi_\varepsilon \to af$$

as $\varepsilon \to 0$ uniformly on compact subsets of $\mathbf{R}^n$.

(ii) If $f \in L^\infty(\mathbf{R}^n)$ is in addition uniformly continuous, then $f * \phi_\varepsilon$ converges uniformly to af in the whole of $\mathbf{R}^n$, that is,

$$||f * \phi_\varepsilon - af||_\infty \to 0$$

as $\varepsilon \to 0$.

Theorem 3.10. *Let $\phi \in L^1(\mathbf{R}^n)$ be such that*

(i) $\phi(x) \geq 0$ a.e. $x \in \mathbf{R}^n$.

(ii) ϕ is radial, i.e. $\phi(x) = \tilde{\phi}(|x|)$

(iii) ϕ is radially decreasing, i.e.,

$$|x| > |y| \quad \Rightarrow \quad \phi(x) \leq \phi(y).$$

Then there exists $C = C(n, \phi)$ such that

$$\sup_{\varepsilon} |((f * \phi_\varepsilon)(x)| \leq CMf(x)$$

for all $x \in \mathbf{R}^n$ and $f \in L^p$, $1 \leq p \leq \infty$.

Proof. First we will show by a direct computation utilizing the definition of convolution, that this holds for radial functions with relatively simple structure. Then we obtain the general case by approximation argument. To this end, let us first assume that ϕ is a radial function of the form

$$\phi(x) = \sum_{i=1}^{k} a_i \chi_{B(0,r_i)}, \quad a_i > 0.$$

Then

$$\int_{\mathbf{R}^n} \phi(x) \, \mathrm{d}x = \sum_{i=1}^{k} a_i m(B(0, r_i))$$

Thus we can calculate

$$\begin{aligned}
|(f * \phi_\varepsilon)(x)| &= \left| \int_{\mathbf{R}^n} f(x-y)\phi_\varepsilon(y) \, \mathrm{d}y \right| \\
&= \left| \frac{1}{\varepsilon^n} \int_{\mathbf{R}^n} f(x-y)\phi(\frac{y}{\varepsilon}) \, \mathrm{d}y \right| \\
&\overset{z=y/\varepsilon,\, \mathrm{d}y=\varepsilon^n \, \mathrm{d}z}{=} \left| \int_{\mathbf{R}^n} f(x-\varepsilon z)\phi(z) \, \mathrm{d}z \right| \\
&= \left| \sum_{i=1}^{k} \int_{B(0,r_i)} f(x-\varepsilon z) a_i \, \mathrm{d}z \right| \\
&\leq \sum_{i=1}^{k} a_i \int_{B(0,r_i)} |f(x-\varepsilon z)| \, \mathrm{d}z \\
&= \sum_{i=1}^{k} a_i m(B(0,r_i)) \fint_{B(0,r_i)} |f(x-\varepsilon z)| \, \mathrm{d}z.
\end{aligned}$$

By a change of variables $y = x - \varepsilon z$, $z = (x-y)/\varepsilon$, $\mathrm{d}z = \mathrm{d}y/\varepsilon^n$ we see that

$$\begin{aligned}
\fint_{B(0,r_i)} |f(x-\varepsilon z)| \, \mathrm{d}z &= \frac{1}{\varepsilon^n m(B(0,r_i))} \int_{B(x,\varepsilon r_i)} |f(y)| \, \mathrm{d}y \\
&= \frac{1}{m(B(0,\varepsilon r_i))} \int_{B(x,\varepsilon r_i)} |f(y)| \, \mathrm{d}y \\
&\leq \frac{m(Q(x,2\varepsilon r_i))}{m(B(0,\varepsilon r_i))} \frac{1}{m(Q(x,2\varepsilon r_i))} \int_{Q(x,2\varepsilon r_i)} |f(y)| \, \mathrm{d}y \\
&\leq C(n)Mf(x).
\end{aligned}$$

Combining the facts, we get

$$\begin{aligned}|(f * \phi_\varepsilon)(x)| &\le \sum_{i=1}^{k} a_i m(B(0,r_i))C(n)Mf(x)\\ &= C(n)\,||\phi||_1\, Mf(x).\end{aligned}$$

Next we go to the general case. As ϕ is nonnegative, radial, and radially decreasing, there exists a sequence $\phi_j, j = 1, 2, \dots$ of function as above such that $\phi_1 \le \phi_2 \le \dots$ and

$$\phi_j(x) \to \phi(x) \quad \text{a.e.} \quad x \in \mathbf{R}^n,$$

as $j \to \infty$. Now

$$\begin{aligned}|(f * \phi_\varepsilon)(x)| &\le \int_{\mathbf{R}^n} |f(x-y)|\, \phi_\varepsilon(x)\,\mathrm{d}x\\ &= \int_{\mathbf{R}^n} |f(x-y)| \lim_{j\to\infty} (\phi_j)_\varepsilon(y)\,\mathrm{d}y\\ &\overset{\text{MON}}{=} \lim_{j\to\infty} \int_{\mathbf{R}^n} |f(x-y)|\,(\phi_j)_\varepsilon(y)\,\mathrm{d}y\\ &\le C(n) \lim_{j\to\infty} ||\phi_j||_1\, Mf(x)\\ &\overset{\text{MON}}{=} C(n)\,||\phi||_1\, Mf(x)\end{aligned}$$

for every $x \in \mathbf{R}^n$. In the calculation above, MON stands for the Lebesgue monotone convergence theorem. □

Remark 3.11. If ϕ is not radial or nonnegative, then we can use radial majorant

$$\tilde{\phi}(x) = \sup_{|y|\ge|x|} |\phi(y)|$$

which is nonnegative, radial and radially decreasing. Thus if $\tilde{\phi} \in L^1(\mathbf{R}^n)$, then the previous theorem, as well as the next theorem holds.

Theorem 3.12. *Let $\phi \in L^1(\mathbf{R}^n)$ be as in Theorem 3.10 that is*
(i) $\phi(x) \ge 0$ a.e. $x \in \mathbf{R}^n$.
(ii) ϕ is radial, i.e. $\phi(x) = \tilde{\phi}(|x|)$
(iii) ϕ is radially decreasing, i.e.,

$$|x| > |y| \quad \Rightarrow \quad \phi(x) \le \phi(y).$$

and $a = ||\phi||_1$. If $f \in L^p(\mathbf{R}^n)$, $1 \le p \le \infty$, then

$$\lim_{\varepsilon\to 0}(f * \phi_\varepsilon)(x) = af(x)$$

for almost all $x \in \mathbf{R}^n$.

21.9.2010

Proof. The sketch of the proof: By a density of continuous functions in L^p, we can choose $g \in C_0(\mathbf{R}^n)$ so that $||f-g||_p$ is small. By adding and subtracting g, we can estimate

$$\begin{aligned} |(f*\phi_\varepsilon)(x) - af(x)| \le &\ |\phi_\varepsilon * (f-g)(x) - a(f-g)(x)| \\ &+ |(g*\phi_\varepsilon)(x) - ag(x)| \,. \end{aligned} \tag{3.13}$$

Since $g \in C_0(\mathbf{R}^n)$, the second term tends to zero as $\varepsilon \to 0$. Thus we can focus attention on the first term on the right hand side. By Theorem 3.10, we can estimate

$$\begin{aligned} |(f*\phi_\varepsilon)(x) - af(x)| &\le |\phi_\varepsilon * (f-g)(x) - a(f-g)(x)| \\ &\le M(f-g)(x) + a\,|(f-g)(x)| \,. \end{aligned}$$

Finally, we can show by using the weak type estimates that the quantities on the right hand side get small almost everywhere.

Details: **Case** $1 \le p < \infty$**:**
As sketched above the weak type estimates play a key role. Theorem Hardy-Littlewood I (Theorem 2.12) implies

$$m(\{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}) \le \frac{C}{\lambda}\,||f||_1 \tag{3.14}$$

for $\lambda > 0$, and Hardy-Littlewood II (Theorem 2.19) imply

$$m(\{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}) \overset{\text{Chebyshev}}{\le} \frac{C}{\lambda^p}\,||Mf||_p^p \overset{\text{H-L II}}{\le} C\,||f||_p^p \,. \tag{3.15}$$

As g is continuous at $x \in \mathbf{R}^n$ it follows that for every $\eta > 0$ there exists $\delta > 0$ such that

$$|g(x-y) - g(x)| < \eta \quad \text{whenever} \quad |y| < \delta.$$

Thus

$$\begin{aligned} |(g*\phi_\varepsilon)(x) - ag(x)| &\le \int_{\mathbf{R}^n} |g(x-y) - g(x)|\,\phi_\varepsilon(y)\,\mathrm{d}y \\ &\le \eta \underbrace{\int_{B(0,\delta)} \phi_\varepsilon(y)\,\mathrm{d}y}_{\le ||\phi||_1} + 2\,||g||_\infty \underbrace{\int_{\mathbf{R}^n \setminus B(0,\delta)} \phi_\varepsilon(x)\,\mathrm{d}y}_{\to 0 \text{ as } \varepsilon \to 0 \text{ by Lemma 3.6}} \,. \end{aligned}$$

Since η was arbitrary, it follows that

$$\lim_{\varepsilon \to 0} |(g*\phi_\varepsilon)(x) - ag(x)| = 0$$

for all $x \in \mathbf{R}^n$.

This in mind we can estimate

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0} |(f*\phi_\varepsilon)(x) - af(x)| \\
&\qquad \le \limsup_{\varepsilon\to 0} |((f-g)*\phi_\varepsilon)(x) - a(f-g)(x)| \\
&\qquad\quad + \underbrace{\limsup_{\varepsilon\to 0} |(g*\phi_\varepsilon)(x) - ag(x)|}_{=0} \\
&\qquad \le \sup_{\varepsilon>0} |((f-g)*\phi_\varepsilon)(x)| + a\,|(f-g)(x)| \\
&\qquad \overset{\text{Theorem 3.10}}{\le} CM(f-g)(x) + a\,|(f-g)(x)|\,.
\end{aligned}
\tag{3.16}
$$

Next we define

$$A_i = \{x \in \mathbf{R}^n \,:\, \limsup_{\varepsilon\to 0} |(f*\phi_\varepsilon)(x) - af(x)| > \frac{1}{i}\}.$$

By the previous estimate,

$$A_i \subset \{x \in \mathbf{R}^n \,:\, CM(f-g)(x) > \frac{1}{2i}\} \cup \{x \in \mathbf{R}^n \,:\, a\,|f(x)-g(x)| > \frac{1}{2i}\},$$

for $i = 1, 2, \dots$. Let $\eta > 0$, and let $g \in C_0(\mathbf{R}^n)$ be such that (density)

$$||f-g||_p \le \eta.$$

This and the previous inclusion imply

$$
\begin{aligned}
m(A_i) &\le m(\{x \in \mathbf{R}^n \,:\, CM(f-g)(x) > \frac{1}{2i}\}) + m(\{x \in \mathbf{R}^n \,:\, a\,|f(x)-g(x)| > \frac{1}{2i}\}) \\
&\overset{(3.14),(3.15)}{\le} Ci^p\,||f-g||_p^p + Ci^p\,||f-g||_p^p \\
&\le Ci^p\,||f-g||_p^p \le Ci^p\eta^p
\end{aligned}
$$

for every η, $i = 1, 2, \dots$. Thus

$$m(A_i) = 0$$

and

$$m(\cup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} m(A_i) = 0.$$

This gives us

$$m(\{x \in \mathbf{R}^n \,:\, \limsup_{\varepsilon\to 0} |(f*\phi_\varepsilon)(x) - af(x)| > 0\}) = 0$$

which proofs the claim

$$\lim_{\varepsilon\to 0} |(f*\phi_\varepsilon)(x) - af(x)| = 0 \quad \text{a.e. } x \in \mathbf{R}^n.$$

Case $p = \infty$: Now $f \in L^\infty(\mathbf{R}^n)$. We show that

$$\lim_{\varepsilon\to 0} (f*\phi_\varepsilon)(x) = af(x)$$

for almost every $x \in B(0,r), r > 0$. Let

$$f_1(x) = f\chi_{B(0,r+1)}(x) = \begin{cases} f(x), & x \in B(0,r+1) \\ 0, & \text{otherwise,} \end{cases}$$

and $f_2 = f - f_1$. Now $f_1 \in L^1(\mathbf{R}^n)$ and by the previous case

$$\lim_{\varepsilon \to 0} (f_1 * \phi_\varepsilon)(x) = af_1(x)$$

for almost every $x \in \mathbf{R}^n$. By utilizing this, we obtain for almost every $x \in B(0,r)$ that

$$\begin{aligned} \lim_{\varepsilon \to 0} (f * \phi_\varepsilon)(x) &= \lim_{\varepsilon \to 0} (f_1 * \phi_\varepsilon)(x) + \lim_{\varepsilon \to 0} (f_2 * \phi_\varepsilon)(x) \\ &= af(x) + \lim_{\varepsilon \to 0} (f_2 * \phi_\varepsilon)(x), \end{aligned}$$

and it remains to show that $\lim_{\varepsilon \to 0} (f_2 * \phi_\varepsilon)(x) = 0$ for almost all $x \in B(0,r)$. To this end, let $x \in B(0,r)$ so that $f_2(x-y) = 0$ for $y \in B(0,1)$ and calculate

$$\begin{aligned} |(f_2 * \phi_\varepsilon)(x)| &= \left| \int_{\mathbf{R}^n} f_2(x-y)\phi_\varepsilon(y)\,\mathrm{d}y \right| \\ &= \left| \int_{\mathbf{R}^n \setminus B(0,1)} f_2(x-y)\phi_\varepsilon(y)\,\mathrm{d}y \right| \\ &= ||f_2||_\infty \int_{\mathbf{R}^n \setminus B(0,1)} \phi_\varepsilon(y)\,\mathrm{d}y \to 0 \end{aligned}$$

as $\varepsilon \to 0$. □

By choosing

$$\phi(x) = \chi_{B(0,1)}(x)/m(B(0,1)),$$

so that

$$\phi_\varepsilon(x) = \chi_{B(0,\varepsilon)}/(\varepsilon^n m(B(0,1))) = \chi_{B(0,\varepsilon)}/m(B(0,\varepsilon)),$$

we immediately obtain

Theorem 3.17 (Lebesgue density theorem)**.** *If $f \in L^1_{loc}(\mathbf{R}^n)$, then*

$$\lim_{r \to 0} \fint_{B(x,R)} f(y)\,\mathrm{d}y = f(x)$$

for almost every $x \in \mathbf{R}^n$.

Example 3.18. *Let*

$$\phi(x) = P(x) = \frac{C(n)}{(1+|x|^2)^{(n+1)/2}}$$

where the constant is chosen so that

$$\int_{\mathbf{R}^n} P(x)\,\mathrm{d}x = 1.$$

Next we define

$$P_t(x) = \frac{1}{t^n} P(\frac{x}{t}) = C(n) \frac{t}{(|x|^2 + t^2)^{(n+1)/2}},\ t > 0$$

and

$$u(x,t) = (f * P_t)(x) = \int_{\mathbf{R}^n} P_t(x-y) f(y) \, \mathrm{d}y.$$

This is called the Poisson integral for f. It has the following properties

(i) $\Delta u = \frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x_1^2} + \ldots + \frac{\partial^2 u}{\partial x_2^2} = 0$ *and*

(ii) $\lim_{t\to 0} u(x,t) = f(x)$ *for almost every $x \in \mathbf{R}^n$ by Theorem 3.12.*

Let

$$\mathbf{R}^{n+1}_+ = \{(x_1, x_2, \ldots, t) \in \mathbf{R}^{n+1} \,:\, t > 0\}$$

denote the upper half space. As stated above u is harmonic in $\mathbf{R}^{n+1}_+$ so that $u(x,t) = \int_{\mathbf{R}^n} P_t(x-y) f(y) \, \mathrm{d}y$ solves

$$\begin{cases} \Delta u(x,t) = 0, & (x,t) \in \mathbf{R}^{n+1}_+ \\ u(x,0) = f(x), & x \in \partial \mathbf{R}^{n+1}_+ = \mathbf{R}^n, \end{cases}$$

where the boundary condition is obtained in the sense

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

almost everywhere on $\mathbf{R}^n$. As $(x,t) \to (x,0)$ along a perpendicular axis, we call this radial convergence.

Question Does the Poisson integral converge better than radially?

Definition 3.19. Let $x \in \mathbf{R}^n$ and $\alpha > 0$. Then

(i) We define a cone

$$\Gamma_\alpha(x) = \{(y,t) \in \mathbf{R}^{n+1}_+ \,:\, |x-y| < \alpha t\}.$$

(ii) Function $u(x,t)$ converges nontangentially, if $u(y,t) \to f(x)$ and $(y,t) \to (x,0)$ so that (y,t) remains inside the cone $\Gamma_\alpha(x)$.

Theorem 3.20. *Let $f \in L^p(\mathbf{R}^n)$, $1 \le p \le \infty$, and $u(x,t) = (f * P_t)(x)$. Then for every $\alpha > 0$, there exists $C = C(n, \alpha)$ such that*

$$u^*_\alpha(x) := \sup_{(y,t)\in\Gamma_\alpha(x)} |u(y,t)| \le CMf(x)$$

for every $x \in \mathbf{R}^n$.

u^* is called a nontangential maximal function. 23.9.2010

Proof. First we show that

$$P_t(y-z) \le C(\alpha, n) P_t(x-z) \quad \text{for every } (y,t) \in \Gamma_\alpha(x),\ z \in \mathbf{R}^n.$$

To establish this, we calculate

$$
\begin{aligned}
|x-z|^2 &\le (|x-y|+|y-z|)^2 \\
&\overset{\text{convexity}}{\le} 2(|x-y|^2+|y-z|^2) \\
&\le 2((\alpha t)^2+|y-z|^2).
\end{aligned}
$$

Thus

$$
\begin{aligned}
|x-z|^2+t^2 &\le (2\alpha^2+1)t^2+2\,|y-z|^2 \\
&\le \max(2,2\alpha^2+1)(|y-z|^2+t^2)
\end{aligned}
$$

so that

$$
\frac{|x-z|^2+t^2}{\max(2,2\alpha^2+1)} \le (|y-z|^2+t^2).
$$

We apply this and deduce

$$
\begin{aligned}
P_t(y-z) &= C(n)\frac{t}{(|y-z|^2+t^2)^{(n+1)/2}} \\
&\le C(n)\max(2,2\alpha^2+1)^{(n+1)/2}\frac{t}{(|x-z|^2+t^2)^{(n+1)/2}} \\
&= C(n,\alpha)P_t(x-z).
\end{aligned}
$$

Utilizing this result we attack the original question and estimate

$$
\begin{aligned}
|u(y,t)| &\le \int_{\mathbf{R}^n} |f(z)|\,P_t(y-z)\,\mathrm{d}z \\
&\le C(\alpha,n)\int_{\mathbf{R}^n} |f(z)|\,P_t(x-z)\,\mathrm{d}z \\
&= C(\alpha,n)(|f|*P_t)(x) \\
&\le C(\alpha,n)\sup_{t>0}(|f|*P_t)(x) \\
&\overset{\text{Theorem 3.10}}{\le} C(\alpha,n)Mf(x).
\end{aligned}
$$

This concludes the proof giving

$$
\sup_{(x,t)\in\Gamma_\alpha(x)} |u(y,t)| \le cMf(x).
$$

□

Corollary 3.21. *If $f\in L^p(\mathbf{R}^n)$, $1\le p\le\infty$, then*

$$
(f*P_t)(y)\to f(x)
$$

nontangentially for almost every $x\in\mathbf{R}^n$.

Proof. Replace in (3.16) the use of Theorem 3.10 by the above estimate.

□

Remark 3.22. By considering a discontinuous $f \in L^p$, we see that $(f * P_{t_n})(y_n)$ does not converge to $f(x)$ for every sequence $(y_n, t_n) \to (x, 0)$. The cone is not the whole of the half space i.e. α must be finite!

Nevertheless, if $f \in C(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$, it follows that

$$u(y,t) = (f * P_t)(y) \to f(x)$$

when $(y,t) \to (x,0)$ in $\mathbf{R}^{n+1}_+$ without further restrictions. This is a consequence of Remark 3.9.

4. Muckenhoupt weights

A weight is a function $w \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$, such that $w \geq 0$ a.e. We have already seen that strong (p,p) property for a Hardy-Littlewood maximal function is an important tool in many applications. Next we study the question in the weighted case:

Let $1 < p < \infty$. Which weights $w \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ satisfy

$$\int_{\mathbf{R}^n} (Mf(x))^p w(x)\,\mathrm{d}x \leq C \int_{\mathbf{R}^n} |f(x)|^p\, w(x)\,\mathrm{d}x? \tag{4.1}$$

for every $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. As before

$$Mf(x) = \sup_{Q \ni x} \frac{1}{m(Q)} \int_Q |f(y)|\,\mathrm{d}y$$

is a Hardy-Littlewood maximal function.

This estimate implies the weak (p,p) estimate. Indeed,

$$\begin{aligned}
\int_{\{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}} w(x)\,\mathrm{d}x &\leq \int_{\{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}} \Big(\frac{Mf(x)}{\lambda}\Big)^p w(x)\,\mathrm{d}x \\
&\leq \frac{1}{\lambda^p} \int_{\mathbf{R}^n} (Mf(x))^p w(x)\,\mathrm{d}x \qquad (4.2) \\
&\overset{(4.1)}{\leq} \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f(x)|^p\, w(x)\,\mathrm{d}x.
\end{aligned}$$

If we define a measure

$$\mu(E) := \int_E w(x)\,\mathrm{d}x$$

then the weighted strong (p,p) estimate (4.1) can be written as

$$\int_{\mathbf{R}^n} (Mf(x))^p\,\mathrm{d}\mu \leq C \int_{\mathbf{R}^n} |f(x)|^p\,\mathrm{d}\mu \tag{4.3}$$

First, we derive some consequences for the weighted **weak** (p,p) estimate. Thus we also obtain some necessary conditions for the question: Which weights $w \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ satisfy weak (p,p) type estimate?

Lemma 4.4. *Suppose that the weighted weak* (p,p) *estimate* (4.2) *holds for some* p, $1 \le p < \infty$. *Then*

$$\Big(\frac{1}{m(Q)}\int_Q |f(x)|\,\mathrm{d}x\Big)^p \le \frac{C}{\mu(Q)}\int_Q |f(x)|^p\,\mathrm{d}\mu$$

for all cubes $Q \subset \mathbf{R}^n$ *and* $f \in L^1_{loc}(\mathbf{R}^n)$.

Proof. Fix a cube. If $\int_Q |f(x)|\,\mathrm{d}x = 0$ or $\int_Q |f(x)|\,\mathrm{d}\mu(x) = \infty$ then the result immediately follows. Thus we may assume

$$\frac{1}{m(Q)}\int_Q |f(x)|\,\mathrm{d}x > \lambda > 0$$

which implies according to the definition of the maximal function that

$$Mf(x) > \lambda > 0$$

for every $x \in Q$. In other words,

$$Q \subset \{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\}$$

so that

$$\begin{aligned}\mu(Q) &\le \mu(\{x \in \mathbf{R}^n \,:\, Mf(x) > \lambda\})\\ &\overset{(4.2)}{\le} \frac{C}{\lambda^p}\int_{\mathbf{R}^n} |f(x)|^p\,\mathrm{d}\mu.\end{aligned}$$

If we replace f by $f\chi_Q$ then this gives

$$\mu(Q) \le \frac{C}{\lambda^p}\int_Q |f(x)|^p\,\mathrm{d}\mu,$$

and by recalling the definition of λ we get the claim. □

Remark 4.5. By analyzing the previous result, we see some of the properties of weights we are studying. Let us choose $f = \chi_E$, $E \subset Q$ a measurable set, in the previous lemma. Then the lemma gives

$$\mu(Q)\Big(\frac{m(E)}{m(Q)}\Big)^p \le C\mu(E). \tag{4.6}$$

This implies

(i) Either $w = 0$ a.e. or $w > 0$ a.e. in Q

Indeed, otherwise it would hold for

$$E = \{x \in Q \,:\, w(x) = 0\}$$

that

$$m(E), m(Q \setminus E) > 0$$

(if "$w = 0$ a.e. in Q" is false, then $m(Q \setminus E) > 0$ and similarly for the other case) and further by $m(Q \setminus E) > 0$ it follows that

$$\mu(Q) > 0.$$

Then the right hand side would be zero (clearly $\mu(E) = \int_E w(x)\,\mathrm{d}x = \int_{\{w=0\}} w\,\mathrm{d}x = 0$) whereas the left hand side would be positive. A contradiction.

(ii) By choosing $Q = Q(x, 2l)$ and $E = Q(x, l)$, we see that

$$\mu(Q(x, 2l)) \leq C\mu(Q(x, l)),$$

because $m(Q(x, l))/m(Q(x, 2l)) = 2^n$. Measures with this property are called *doubling measures.*

(iii) Either $w = \infty$ a.e. or $w \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$.

If there would be a set

$$E \subset Q \text{ such that } w(x) < \infty \text{ and } m(E) > 0,$$

by (4.6) it follows that $\mu(Q) = \int_Q w(x)\,\mathrm{d}x$ is finite, and thus

$$w \in L^1(Q)$$

and by choosing larger cubes, we get $w \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. Thus the result follows.

Observe that $w \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ was one of our assumptions when defining weights, but it would be possible to take the weak type estimate as a starting point and then derive this as a result as shown above.

Next we derive **a necessary condition for weak** $(1, 1)$ estimate to hold.

Case $p = 1$: We shall use notation

$$\operatorname*{ess\,inf}_{x \in Q} w(x) := \sup\{m \in \mathbf{R} \,:\, w(x) \geq m \text{ a.e. } x \in Q\}$$

and define a set

$$E_\varepsilon = \{x \in Q \,:\, w(x) < \operatorname*{ess\,inf}_{y \in Q} w(y) + \varepsilon\}$$

for some $\varepsilon > 0$. By definition of ess inf, we have $m(E_\varepsilon) > 0$.

Now by (4.6),

$$\begin{aligned} \frac{\mu(Q)}{m(Q)} &\leq C\frac{\mu(E_\varepsilon)}{m(E_\varepsilon)} \\ &\overset{\text{def of } \mu}{=} \frac{C}{m(E_\varepsilon)} \int_{E_\varepsilon} w(x)\,\mathrm{d}x \leq C(\operatorname*{ess\,inf}_{y \in Q} w(y) + \varepsilon). \end{aligned}$$

By passing to a zero with ε, and recalling that $\mu(Q) = \int_Q w(x)\,\mathrm{d}x$, we get *Muckenhoupt A_1-condition*

$$\frac{1}{m(Q)} \int_Q w(x)\,\mathrm{d}x \leq C \operatorname*{ess\,inf}_{y \in Q} w(y). \tag{4.7}$$

If this condition holds we denote $w \in A_1$.

Lemma 4.8. *A weight w satisfies Muckenhoupt A_1-condition if and only if*

$$Mw(x) \leq Cw(x)$$

for almost every $x \in \mathbf{R}^n$.

On the other hand from the Lebesgue density theorem, we get $w(x) \leq Mw(x)$ for almost every $x \in \mathbf{R}^n$ so that

$$w(x) \leq Mw(x) \leq Cw(x).$$

Proof. "$\Leftarrow$" Suppose that $Mw(x) \leq Cw(x)$ for almost every $x \in \mathbf{R}^n$. Then

$$\frac{1}{m(Q)} \int_Q w(y)\,\mathrm{d}y \leq Cw(x) \text{ a.e. } x \in Q,$$

and thus

$$\frac{1}{m(Q)} \int_Q w(y)\,\mathrm{d}y \leq C \operatorname*{ess\,inf}_{x \in Q} w(x).$$

28.9.2010 "$\Rightarrow$" Suppose that $w \in A_1$ so that $\frac{1}{m(Q)} \int_Q w(y)\,\mathrm{d}y \leq C \operatorname{ess\,inf}_{x \in Q} w(x)$. We shall show that

$$m(\{x \in \mathbf{R}^n \,:\, Mw(x) > Cw(x)\}) = 0.$$

Choose a point $x \in \{x \in \mathbf{R}^n \,:\, Mw(x) > Cw(x)\}$ so that $Mw(x) > Cw(x)$. Then there exists a cube $Q \ni x$ such that

$$\frac{1}{m(Q)} \int_Q w(y)\,\mathrm{d}y > Cw(x).$$

Without loss of generality we may choose this cube so that the corners lie in the rational points. Thus

$$Cw(x) < \frac{1}{m(Q)} \int_Q w(y)\,\mathrm{d}y \overset{A_1}{\leq} C \operatorname*{ess\,inf}_{y \in Q} w(y)$$

so that

$$w(x) < \operatorname*{ess\,inf}_{y \in Q} w(y).$$

For this cube, we denote by

$$E_Q = \{x \in Q \,:\, w(x) < \operatorname*{ess\,inf}_{y \in Q} w(y)\})$$

which is of measure zero. Now we repeat the process for each $x \in \{x \in \mathbf{R}^n \,:\, Mw(x) > Cw(x)\}$ and as we restricted ourselves to a countable family of cubes with corners at rational points, we have

$$m(\bigcup E_Q) = 0$$

because countable union of zero measurable sets has a measure zero. $\square$

Observe/recall that uncountable union of zero measurable sets is not necessarily zero measurable, cf. $m(\cup_{x\in(0,1)}\{x\}) = 1$. Therefore the restriction on the countable set of cubes was necessary above.

Example 4.9. $w(x) = |x|^{-\alpha}$, $0 \le \alpha < n$, $x \in \mathbf{R}^n$, *belongs to* A_1. *Indeed, let* $x \in \mathbf{R}^n \setminus \{0\}$, $x \in Q$. *Then by choosing a radius* $r = l(Q)\sqrt{n}$, *we see that*

$$Q \subset B(x, r).$$

We calculate

$$\begin{aligned}
\frac{1}{m(Q)} \int_{Q \ni x} w(y)\,\mathrm{d}y &\le \frac{C}{B(x,r)} \int_{B(x,r)} w(y)\,\mathrm{d}y \\
&\overset{z = \frac{y}{|x|}, y = z|x|,\ \mathrm{d}y = |x|^n\ \mathrm{d}z}{=} \frac{C}{r^n} \int_{B(\frac{x}{|x|}, \frac{r}{|x|})} ||x|\, z|^{-\alpha} |x|^n\ \mathrm{d}z \\
&= \frac{C\,|x|^{-\alpha}}{\left(\frac{r}{|x|}\right)^n} \int_{B(\frac{x}{|x|}, \frac{r}{|x|})} |z|^{-\alpha}\ \mathrm{d}z \\
&\le C w(x) \underbrace{Mw\Big(\frac{x}{|x|}\Big)}_{<\infty}.
\end{aligned}$$

Thus by taking a supremum over Q *such that* $x \in Q$, *we see that*

$$Mw(x) \le Cw(x),$$

so that by Lemma 4.8, $w \in A_1$. *Also calculate* $⨍_{B(0,r)} w\,\mathrm{d}x$.

Next we derive **a necessary condition for weak** (p,p) estimate to hold.
Lemma 4.4 gives us the estimate

$$\mu(Q)\Big(\frac{1}{m(Q)} \int_Q |f(x)|\ \mathrm{d}x\Big)^p \le C \int_Q |f(x)|^p\ \mathrm{d}\mu.$$

We choose $f(x) = w^{1-p'}(x)$, where $1/p' + 1/p = 1$ i.e. $p' = p/(p-1)$. Recalling that $\mu(Q) = \int_Q w(x)\,\mathrm{d}x$, we get

$$\begin{aligned}
\int_Q w(x)\,\mathrm{d}x \Big(\frac{1}{m(Q)} \int_Q w^{1-p'}(x)\,\mathrm{d}x\Big)^p &\le C \int_Q w^{(1-p')p}(x) w(x)\,\mathrm{d}x \\
&= C \int_Q w(x)^{(1-p')p+1}\,\mathrm{d}x.
\end{aligned}$$

A short calculation $((1-p')p + 1 = (1 - p/(p-1))p + 1 = ((p-1-p)/(p-1))p + 1 = -p/(p-1) + 1 = 1 - p')$ shows that

$$(1-p')p + 1 = 1 - p'$$

so that if we divide by the integral on the right hand side the above inequality, we get

$$\frac{1}{m(Q)}\int_Q w(x)\,\mathrm{d}x\Big(\frac{1}{m(Q)}\int_Q w^{1-p'}(x)\,\mathrm{d}x\Big)^{p-1}\le C, \tag{4.10}$$

or

$$\frac{1}{m(Q)}\int_Q w(x)\,\mathrm{d}x\Big(\frac{1}{m(Q)}\int_Q w^{1/(1-p)}(x)\,\mathrm{d}x\Big)^{p-1}\le C.$$

This is called the *Muckenhoupt A_p-condition.*

Observe that above, we implicitly use $w^{1-p'}\in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. If this is not the case, we can consider

$$f=(w+\varepsilon)^{1-p'},$$

derive the above estimate, and let finally $\varepsilon\to 0$. After this argument, as $w>0$ a.e., (4.10) implies that $w^{1-p'}\in L^1_{\mathrm{loc}}(\mathbf{R}^n)$.

Example 4.11. *$w(x)=|x|^{-\alpha}$, $0\le\alpha<n$, $x\in\mathbf{R}^n$, belongs to A_p. It might also be instructive to calculate*

$$\frac{1}{m(B(0,r)}\int_{B(0,r)} w\,\mathrm{d}x\Big(\frac{1}{m(B(0,r))}\int_{B(0,r)} w^{1/(1-p)}\,\mathrm{d}x\Big)^{p-1}.$$

Let us collect the above definitions.

Definition 4.12 (Muckenhoupt 1972)**.** Let $w\in L^1_{\mathrm{loc}}(\mathbf{R}^n)$, $w>0$ a.e. Then w satisfies A_1-condition if there exists $C>0$ s.t.

$$\int_Q w(x)\,\mathrm{d}x\le C\operatorname*{ess\,inf}_{y\in Q} w(y).$$

for all cubes $Q\subset\mathbf{R}^n$. For $1<p<\infty$, w satisfies A_p-condition if there exists $C>0$ s.t.

$$\frac{1}{m(Q)}\int_Q w(x)\,\mathrm{d}x\Big(\frac{1}{m(Q)}\int_Q w^{1-p'}(x)\,\mathrm{d}x\Big)^{p-1}\le C$$

for all cubes $Q\subset\mathbf{R}^n$.

Remark 4.13. (i) $1-p'=1/(1-p)<0$, $w^{1-p'}\in L^1_{\mathrm{loc}}(\mathbf{R}^n)$

(ii) Let $p=2$. Then

$$\frac{1}{m(Q)}\int_Q w(x)\,\mathrm{d}x\frac{1}{m(Q)}\int_Q\frac{1}{w(x)}\,\mathrm{d}x\le C$$

(iii)

$$\begin{aligned} m(Q)&=\int_Q w^{1/p}w^{-1/p}\,\mathrm{d}x\\ &\overset{\text{Hölder}}{\le}\Big(\int_Q w^{p(1/p)}\,\mathrm{d}x\Big)^{1/p}\Big(\int_Q w^{p'(-1/p)}\,\mathrm{d}x\Big)^{1/p'}\\ &=\Big(\int_Q w\,\mathrm{d}x\Big)^{1/p}\Big(\int_Q w^{1-p'}\,\mathrm{d}x\Big)^{1/p'}. \end{aligned}$$

Dividing by $m(Q) = m(Q)^{1/p} m(Q)^{1/p'}$ and then taking power p on both sides we get

$$\frac{1}{m(Q)} \int_Q w \, \mathrm{d}x \Big(\frac{1}{m(Q)} \int_Q w^{1-p'} \, \mathrm{d}x \Big)^{p-1} \geq 1 \tag{4.14}$$

so that

$$\Big(\frac{1}{m(Q)} \int_Q w^{1-p'} \, \mathrm{d}x \Big)^{1-p} \leq \frac{1}{m(Q)} \int_Q w(x) \, \mathrm{d}x.$$

This was (a consequence of) Hölder's inequality. On the other hand, by looking at the A_p condition, we see that the inequality is reversed. Thus A_p condition is a reverse Hölder's inequality.

Theorem 4.15. *$A_p \subset A_q$, $1 \leq p < q$.*

Proof. **Case $1 < p < \infty$.** We recall that $q' - 1 = 1/(q-1)$.

$$\begin{aligned}
&\Big(\frac{1}{m(Q)} \int_Q \Big(\frac{1}{w} \Big)^{\frac{1}{q-1}} \mathrm{d}x \Big)^{q-1} \\
&\quad \overset{\text{Hölder}}{\leq} \Big(\frac{1}{m(Q)} \Big)^{q-1} \Big(\int_Q \Big(\frac{1}{w} \Big)^{\frac{1}{q-1} \frac{q-1}{p-1}} \mathrm{d}x \Big)^{(q-1)\frac{p-1}{q-1}} m(Q)^{(q-1)(1-\frac{p-1}{q-1})} \\
&\quad = C \Big(\int_Q \Big(\frac{1}{w} \Big)^{1/(p-1)} \mathrm{d}x \Big)^{p-1} m(Q)^{1-p} \\
&\quad \overset{w \in A_p}{\leq} \Big(\frac{1}{m(Q)} \int_Q w \, \mathrm{d}x \Big)^{-1}
\end{aligned}$$

which proves the claim in this case.
Case $p = 1$.

$$\begin{aligned}
\Big(\frac{1}{m(Q)} \int_Q \Big(\frac{1}{w} \Big)^{1/(q-1)} \mathrm{d}x \Big)^{q-1} &\leq \operatorname*{ess\,sup}_Q \frac{1}{w} \\
&= \frac{1}{\operatorname{ess\,inf}_Q w} \overset{w \in A_1}{\leq} \frac{C}{⨍_Q w \, \mathrm{d}x}.
\end{aligned}$$

□

Theorem 4.16. *Let $1 \leq p < \infty$, and $w \in L^1_{loc}(\mathbf{R}^n)$, $w < 0$ a.e. Then $w \in A_p$ if and only if*

$$\Big(\frac{1}{m(Q)} \int_Q |f(x)| \, \mathrm{d}x \Big)^p \leq \frac{C}{\mu(Q)} \int_Q |f(x)|^p \, \mathrm{d}\mu.$$

for every $f \in L^1_{loc}(\mathbf{R}^n)$ and $Q \subset \mathbf{R}^n$.

Proof. **Case $1 < p < \infty$.**

"$\Leftarrow$" was already proven before (4.10).

"⇒" First we use Hölder's inequality

$$\begin{aligned}\frac{1}{m(Q)}\int_Q |f(x)|\,\mathrm{d}x &= \frac{1}{m(Q)}\int_Q |f(x)|\,w(x)^{1/p}\Big(\frac{1}{w(x)}\Big)^{1/p}\mathrm{d}x \\ &\le \frac{1}{m(Q)}\Big(\int_Q |f(x)|^p\,w(x)\,\mathrm{d}x\Big)^{1/p}\Big(\int_Q \Big(\frac{1}{w(x)}\Big)^{p'/p}\mathrm{d}x\Big)^{1/p'},\end{aligned}$$

for $1/p'+1/p=1$. By taking the power p on both sides, using the definition of μ, arranging terms, using $p/p'=p-1$, $-p'/p=1/(1-p)$, and A_p condition, we get

$$\begin{aligned}\mu(Q)\Big(\frac{1}{m(Q)}\int_Q |f(x)|\,\mathrm{d}x\Big)^p &\le \frac{1}{m(Q)^p}\Big(\int_Q |f(x)|^p\,w(x)\,\mathrm{d}x\Big) \\ &\qquad \underbrace{\cdot\int_Q w(x)\,\mathrm{d}x\Big(\int_Q w(x)^{1/(1-p)}\,\mathrm{d}x\Big)^{p-1}}_{\overset{w\,\in\,A_p}{\le}\,Cm(Q)^p} \\ &\le C\int_Q |f(x)|^p\,\mathrm{d}\mu.\end{aligned}$$

Case $p=1$**.**

"⇐" was already proven before (4.7).

"⇒" Let $w\in A_1$ i.e.

$$\frac{1}{m(Q)}\int_Q w(x)\,\mathrm{d}x \le C\operatorname*{ess\,inf}_{x\in Q} w(x).$$

Then

$$\begin{aligned}\mu(Q)\frac{1}{m(Q)}\int_Q |f(x)|\,\mathrm{d}x &\le \frac{1}{m(Q)}\int_Q |f(x)|\,\mu(Q)\,\mathrm{d}x \\ &\overset{w\,\in\,A_1}{\le} \int_Q |f(x)|\operatorname*{ess\,inf}_{x\in Q} w(x)\,\mathrm{d}x \\ &\le C\int_Q |f(x)|\,w(x)\,\mathrm{d}x \\ &\le C\int_Q |f(x)|\,\mathrm{d}\mu. \qquad \square\end{aligned}$$

We aim at proving that the weighted weak/strong type estimate and A_p condition are equivalent. To establish this, we next study Calderón-Zygmund decomposition. It is an important tool both in harmonic analysis and in the theory of PDEs.

30.9.2010

4.1. **Calderón-Zygmund decomposition.** In this section we integrate with respect to the measure m only, and thus we recall the notation $⨍_Q = \frac{1}{m(Q)}\int_Q$.

Next we introduce dyadic cubes, which are generated using powers of 2.

Definition 4.17 (Dyadic cubes)**.** A dyadic interval on $\mathbf{R}$ is

$$[m2^{-k}, (m+1)2^{-k})$$

where $m, k \in \mathbb{Z}$. A dyadic cube in $\mathbf{R}^n$ is

$$\prod [m_j 2^{-k}, (m_j+1)2^{-k})$$

where $m_1, m_2, \ldots, m_n, k \in \mathbb{Z}$.

Observe that corners lie at $2^{-k}\mathbb{Z}^n$ and side length is 2^{-k}. Dyadic cubes have an important property that they are either disjoint or one is contained into another.

Notations

$$D_k = \text{"a collection of dyadic cubes with side length } 2^{-k}. \text{"}$$

A collection of all the dyadic cubes is denoted by

$$D = \bigcup_{k \in \mathbb{Z}} D_k.$$

Theorem 4.18 (Local Calderón-Zygmund decomposition)**.** *Let $Q_0 \subset \mathbf{R}^n$ be a dyadic cube, and $f \in L^1(Q_0)$. Then if*

$$\lambda \geq \fint_{Q_0} |f(x)| \, \mathrm{d}x$$

there exists a collection of dyadic cubes

$$F_\lambda = \{Q_j \, : j = 1, 2, \ldots\}$$

such that

(i)

$$Q_j \cap Q_k = \emptyset \textit{ when } j \neq k,$$

(ii)

$$\lambda < \fint_{Q_j} |f(x)| \, \mathrm{d}x \leq 2^n \lambda, \quad j = 1, 2, \ldots,$$

and

(iii)

$$|f(x)| \leq \lambda \textit{ for a.e. } x \in Q_0 \setminus \cup_{j=1}^{\infty} Q_j.$$

Remark 4.19. Naturally, if $|f(x)| \leq \lambda$, then $F_\lambda = \emptyset$. Notice also the assumption that Q_0 is dyadic could be dropped, and that if the condition $\lambda \geq \fint_{Q_0} |f(x)| \, \mathrm{d}x$ does not hold, then we can choose a larger cube to begin with so that this condition is satisfied.

Proof of Theorem 4.18. Clearly, $Q_0 \notin F_\lambda$ because of our assumption. We split Q_0 into 2^n dyadic cubes with side length $l(Q_0)/2$. Then we choose to F_λ, the cubes for which

$$\lambda < \fint_Q |f(x)| \, \mathrm{d}x.$$

Observe that (i) holds because we use dyadic cubes, and because of the estimate

$$\begin{aligned}\fint_Q |f(x)| \, \mathrm{d}x &\leq \frac{m(Q_0)}{m(Q)} \fint_{Q_0} |f(x)| \, \mathrm{d}x \\ &\leq 2^n \fint_{Q_0} |f(x)| \, \mathrm{d}x \leq 2^n \lambda,\end{aligned} \tag{4.20}$$

also the upper bound in (ii) holds. For the cubes that were not chosen i.e. for which

$$\fint |f(x)| \, \mathrm{d}x \leq \lambda,$$

we continue the process. Then the estimate (ii) holds for all the cubes that were chosen at some round. On the other hand, according to Lebesgue's density theorem

$$|f(x)| = \lim_{k\to\infty} \fint_{Q^{(k)}} |f(y)| \, \mathrm{d}y \overset{Q^{(k)} \text{ was not chosen}}{\leq} \lambda$$

for a.e. $x \in \mathbf{R}^n \setminus \cup_{Q\in F_\lambda} Q$. $\square$

Next we prove a global version of the Calderón-Zygmund decomposition. The idea in the proof is similar to the local version, but as we work in the whole of $\mathbf{R}^n$, there is no initial cube Q_0.

Theorem 4.21 (Global Calderón-Zygmund decomposition)**.** *Let $f \in L^1(\mathbf{R}^n)$ and $\lambda > 0$. Then there exists a collection of dyadic cubes*

$$F_\lambda = \{Q_j : j = 1, 2, \ldots\}$$

such that

(i)

$$Q_j \cap Q_k = \emptyset \text{ when } j \neq k,$$

(ii)

$$\lambda < \fint_{Q_j} |f(x)| \, \mathrm{d}x \leq 2^n \lambda, \quad j = 1, 2, \ldots,$$

and

(iii)

$$|f(x)| \leq \lambda \text{ for a.e. } x \in \mathbf{R}^n \setminus \cup_{j=1}^\infty Q_j.$$

Proof. We study a subcollection

$$F_\lambda \subset D$$

of dyadic cubes, which are the largest possible cubes such that

$$\fint_Q |f(x)| \,\mathrm{d}x > \lambda \tag{4.22}$$

holds. In other words, $Q \in F_\lambda$ if $Q \in D_k$ for some k, if (4.22) holds and for all the larger dyadic cubes $\tilde{Q}$, $Q \subset \tilde{Q}$, it holds that

$$\fint_{\tilde{Q}} |f(y)| \,\mathrm{d}y \leq \lambda.$$

The largest cube exists, if (4.22) holds for Q, because

$$\fint_{\tilde{Q}} |f(x)| \,\mathrm{d}x \leq \frac{||f||_1}{m(\tilde{Q})} \to 0$$

as $m(\tilde{Q}) \to \infty$ because $f \in L^1(\mathbf{R}^n)$. As the cubes in F_λ are maximal, they are disjoint, because if this were not the case the smaller cube would be contained to larger one as they are dyadic and thus we could replace it by the larger one. A similar calculation as in (4.20) shows that also the upper bound in (*ii*) holds. The proof is completed similarly as in the local version: (*iii*) is a consequence of Lebesgue's density theorem Theorem 3.17. □

Example 4.23. *Calderón-Zygmund decomposition for*

$$f : \mathbf{R} \to [0, \infty], \; f(x) = |x|^{-1/2}$$

with $\lambda = 1$.

Example 4.24. *By using the Calderón-Zygmund decomposition, we can split any* $f \in L^1(\mathbf{R}^n)$ *into a good and a bad part as (further details during the lecture)*

$$f = g + b$$

as

$$g = \begin{cases} f(x), & x \in \mathbf{R}^n \setminus \cup_{j=1}^\infty Q_j, \\ \fint_{Q_j} f(y) \,\mathrm{d}y, & x \in Q_j \in F_\lambda \end{cases}$$

and

$$b(x) = \sum_{j=1}^\infty b_j(x),$$

$$b_j(x) = (f(x) - \fint_{Q_j} f(y) \,\mathrm{d}y)\chi_{Q_j}(x).$$

Observe that $g \leq 2^n\lambda$ *and* $\int_{Q_j} b(y) \,\mathrm{d}y = 0$. *Split* $f : \mathbf{R} \to [0, \infty], \; f(x) = |x|^{-1/2}$ *in this way with* $\lambda = 1$.

Lemma 4.25. *Let* $f \in L^1(\mathbf{R}^n)$ *and*

$$F_\lambda = \{Q_j : j = 1, 2, \dots\}$$

Calderón-Zygmund decomposition with $\lambda > 0$ *from Theorem 4.21. Then*

$$\{x \in \mathbf{R}^n : Mf(x) > 4^n\lambda\} \subset \cup_{j=1}^\infty 3Q_j.$$

Proof. The Calderón-Zygmund decomposition gives bounds for the averages, so our task is passing from the averages to the maximal function. To this end, let

$$x \in \mathbf{R}^n \setminus \cup_{j=1}^\infty 3Q_j$$

and $Q \subset \mathbf{R}^n$ is a cube (not necessarily dyadic) s.t. $x \in Q$. If we choose, k so that

$$2^{-k-1} \leq l(Q) < 2^{-k},$$

then there exists at the most 2^n dyadic cubes $R_1, \dots R_l \in D_k$ such that

$$R_m \cap Q \neq \emptyset,\ m = 1, \dots, l.$$

Because R_m and Q intersect, $Q \subset 3R_m$. On the other hand R_m is not contained to any $Q_j \in F_\lambda$, because otherwise we would have $x \in Q \subset 3Q_j$ which contradicts our assumption $x \in \mathbf{R}^n \setminus \cup_{j=1}^\infty 3Q_j$. As R_m is not in F_λ, it follows by definition that

$$⨍_{R_m} |f(y)|\,\mathrm{d}y \leq \lambda$$

for $m = 1, \dots, l$. Thus

$$\begin{aligned} ⨍_Q |f(y)|\,\mathrm{d}y &= \frac{1}{m(Q)} \sum_{m=1}^{l} \int_{R_m \cap Q} |f(y)|\,\mathrm{d}y \\ &\leq \sum_{m=1}^{l} \frac{m(R_m)}{m(Q)} \frac{1}{m(R_m)} \int_{R_m} |f(y)|\,\mathrm{d}y \\ &\leq l2^n\lambda \leq 4^n\lambda. \end{aligned}$$

Moreover,

$$Mf(x) = \sup_{Q \ni x} ⨍_Q |f(y)|\,\mathrm{d}y \leq 4^n\lambda$$

for every $x \in \mathbf{R}^n \setminus \cup_{j=1}^\infty 3Q_j$. Thus

$$\mathbf{R}^n \setminus \cup_{j=1}^\infty 3Q_j \subset \{x \in \mathbf{R}^n : Mf(x) \leq 4^n\lambda\}. \qquad \square$$

5.10.2010

Corollary 4.26. *Let* $f \in L^1(\mathbf{R}^n)$ *and*

$$F_\lambda = \{Q_j : j = 1, 2, \dots\}$$

Calderón-Zygmund decomposition with $\lambda > 0$ *from Theorem 4.21. Then*

(i)

$$\{x \in \mathbf{R}^n : Mf(x) > 4^n\lambda\} \subset \cup_{j=1}^{\infty} 3Q_j.$$

and

(ii)

$$\cup_{j=1}^{\infty} Q_j \subset \{x \in \mathbf{R}^n : Mf(x) > \lambda\}.$$

Proof. (i) The previous lemma.

(ii) $Q_j \in F_\lambda$ implies

$$\fint_{Q_j} |f(y)| \,\mathrm{d}y > \lambda$$

and thus

$$Mf(x) > \lambda$$

for every $x \in Q_j$. Thus

$$\cup_{j=1}^{\infty} Q_j \subset \{x \in \mathbf{R}^n : Mf(x) > \lambda\}. \qquad \square$$

4.2. **Connection of A_p to weak and strong type estimates.** Now, we return to A_p-weights.

Theorem 4.27. *Let $w \in L^1_{loc}(\mathbf{R}^n)$, and $1 \le p < \infty$. Then the following are equivalent*

(i) $w \in A_p$.

(ii)

$$\mu(\{x \in \mathbf{R}^n : Mf(x) > \lambda\}) \le \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f(x)|^p \,\mathrm{d}\mu$$

for every $f \in L^1_{loc}(\mathbf{R}^n),\ \lambda > 0$.

Proof. It was shown above (4.10) in case $1 < p < \infty$ and in the case $p = 1$ above (4.7), that $(ii) \Rightarrow (i)$.

Then we aim at showing that $(i) \Rightarrow (ii)$. The idea is to use Lemma 4.25 and to estimate

$$\mu(\{x \in \mathbf{R}^n : Mf(x) > 4^n\lambda\}) \le \sum_{j=1}^{\infty} \mu(3Q_j), \tag{4.28}$$

for Calderón-Zygmund cubes at the level λ and for $f \in L^1(\mathbf{R}^n)$. Further, we have shown that $w \in A_p$ implies that μ is a doubling measure. Thus

$$\begin{aligned}
\mu(3Q_j) &\le \mu(Q_j) \\
&\overset{\text{Theorem 4.16}}{\le} C\Big(\fint_{Q_j} |f(x)| \,\mathrm{d}x\Big)^{-p} \int_{Q_j} |f(x)|^p \,\mathrm{d}\mu(x) \\
&\overset{Q_j \text{ is a Calderón-Zygmund cube}}{\le} \frac{C}{\lambda^p} \int_{Q_j} |f(x)|^p \,\mathrm{d}\mu(x).
\end{aligned}$$

Using this in (4.28), we get

$$\begin{aligned}\mu(\{x \in \mathbf{R}^n : Mf(x) > 4^n\lambda\}) &\le \sum_{j=1}^{\infty} \mu(3Q_j) \\ &\le \frac{C}{\lambda^p} \sum_{j=1}^{\infty} \int_{Q_j} |f(x)|^p \, \mathrm{d}\mu(x) \\ &\overset{Q_j \text{ are disjoint}}{\le} \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f(x)|^p \, \mathrm{d}\mu(x),\end{aligned}$$

and then replacing $4^n\lambda$ by λ gives the result.

However, in the statement, we only assumed that $f \in L^1_{\text{loc}}(\mathbf{R}^n)$ and in the above argument that $f \in L^1(\mathbf{R}^n)$. We treat this difficulty by considering

$$f_i = f\chi_{B(0,i)}, i = 1, 2, \ldots,$$

and then passing to a limit $i \to \infty$ with the help of Lebesgue's monotone convergence theorem. To be more precise, repeating the above argument, we get

$$\mu(\{x \in \mathbf{R}^n : Mf_i(x) > 4^n\lambda\}) \le \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f_i(x)|^p \, \mathrm{d}\mu(x).$$

Since

$$\{x \in \mathbf{R}^n : Mf(x) > 4^n\lambda\} = \cup_{i=1}^{\infty} \{x \in \mathbf{R}^n : Mf_i(x) > 4^n\lambda\}$$

the basic properties of measure and the above estimate imply

$$\begin{aligned}\mu(\{x \in \mathbf{R}^n : Mf(x) > 4^n\lambda\}) &= \lim_{i\to\infty} \mu(\{x \in \mathbf{R}^n : Mf_i(x) > 4^n\lambda\}) \\ &\le \lim_{i\to\infty} \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f_i(x)|^p \, \mathrm{d}\mu \\ &\overset{\text{MON}}{=} \frac{C}{\lambda^p} \int_{\mathbf{R}^n} |f(x)|^p \, \mathrm{d}\mu.\end{aligned}$$

□

Next we show that $w \in A_p$ satisfies a reverse Hölder's inequality. First, by the usual Hölder's inequality, we get

$$\begin{aligned}\frac{1}{m(Q)} \int_Q |f(x)| \, \mathrm{d}x &\le \frac{1}{m(Q)} \Big(\int_Q |f(x)|^p \, \mathrm{d}x \Big)^{1/p} \Big(\int_Q 1^{p'} \, \mathrm{d}x \Big)^{1/p'} \\ &\le m(Q)^{\frac{1}{p'}-1} \Big(\int_Q |f(x)|^p \, \mathrm{d}x \Big)^{1/p} \\ &\le \Big(\fint_Q |f(x)|^p \, \mathrm{d}x \Big)^{1/p}.\end{aligned}$$

Similarly

$$\Big(\fint_Q |f(x)|^p \, \mathrm{d}x \Big)^{1/p} \le C \Big(\fint_Q |f(x)|^q \, \mathrm{d}x \Big)^{1/q}, \; q > p.$$

Thus it is natural, to call inequality in which the power on the left hand side is larger the *reverse Hölder inequality.* Reverse Hölder inequalities tell, in general, that a function is more integrable than it first appears. We will need the following deep result of Gehring (1973). We skip the lengthy proof.

Lemma 4.29 (Gehring's lemma)**.** *Suppose that for p, $1 < p < \infty$, there exists $C \geq 1$ such that*

$$\Big(\fint_Q |f(x)|^p \,\mathrm{d}x\Big)^{1/p} \leq C \fint_Q |f(x)| \,\mathrm{d}x$$

for all cubes $Q \subset \mathbf{R}^n$. Then there exists $q > p$ such that

$$\Big(\fint_Q |f(x)|^q \,\mathrm{d}x\Big)^{1/q} \leq C \fint_Q |f(x)| \,\mathrm{d}x$$

for all cubes $Q \subset \mathbf{R}^n$.

Theorem 4.30 (reverse Hölder's inequality)**.** *Suppose that $w \in A_p$, $1 \leq p < \infty$. Then there exists $\delta > 0$ and $C > 0$ s.t.*

$$\Big(\frac{1}{m(Q)} \int_Q w^{1+\delta} \,\mathrm{d}x\Big)^{1/(1+\delta)} \leq \frac{C}{m(Q)} \int_Q w \,\mathrm{d}x$$

for all cubes $Q \subset \mathbf{R}^n$.

Proof. Since $w \in A_p$, we have

$$\frac{1}{m(Q)} \int_Q w \,\mathrm{d}x \Big(\frac{1}{m(Q)} \int_Q w^{1/(1-p)} \,\mathrm{d}x\Big)^{p-1} \leq C.$$

On the other hand Hölder's inequality implies for any measurable $f > 0$ (choose $p = p' = 2$ in (4.14)) that

$$\frac{1}{m(Q)} \int_Q f \,\mathrm{d}x \Big(\frac{1}{m(Q)} \int_Q \frac{1}{f} \,\mathrm{d}x\Big) \geq 1.$$

Then we set $f = w^{1/(p-1)}$ and get

$$1 \leq \frac{1}{m(Q)} \int_Q w^{1/(p-1)} \,\mathrm{d}x \Big(\frac{1}{m(Q)} \int_Q \Big(\frac{1}{w}\Big)^{1/(p-1)} \,\mathrm{d}x\Big).$$

Combining the inequalities for w , we get

$$\begin{aligned} &\frac{1}{m(Q)} \int_Q w \,\mathrm{d}x \Big(\frac{1}{m(Q)} \int_Q w^{1/(1-p)} \,\mathrm{d}x\Big)^{p-1} \\ &\qquad \leq \Big(\frac{C}{m(Q)} \int_Q w^{1/(p-1)} \,\mathrm{d}x\Big)^{p-1} \Big(\frac{1}{m(Q)} \int_Q w^{1/(1-p)} \,\mathrm{d}x\Big)^{p-1}. \end{aligned}$$

so that

$$\frac{1}{m(Q)} \int_Q w \,\mathrm{d}x \leq \Big(\frac{C}{m(Q)} \int_Q w^{1/(p-1)} \,\mathrm{d}x\Big)^{p-1}$$

or recalling f

$$\Big(\frac{1}{m(Q)}\int_Q f^{p-1}\,\mathrm{d}x\Big)^{1/(p-1)} \le \frac{C}{m(Q)}\int_Q f\,\mathrm{d}x.$$

Now, we may suppose that $p > 2$ because due to Theorem 4.15, we have $A_p \subset A_q$, $1 \le p < q$, and by this assumption $p-1 > 1$. By Gehring's lemma Lemma 4.29, there exists $q > p-1$ such that

$$\Big(\frac{1}{m(Q)}\int_Q f^{q}\,\mathrm{d}x\Big)^{1/q} \le \frac{C}{m(Q)}\int_Q f\,\mathrm{d}x$$

or again recalling f and taking power $p-1$ on both sides

$$\Big(\frac{1}{m(Q)}\int_Q w^{q/(p-1)}\,\mathrm{d}x\Big)^{(p-1)/q} \le \Big(\frac{C}{m(Q)}\int_Q w^{1/(p-1)}\,\mathrm{d}x\Big)^{p-1}.$$

The right hand side is estimated by using Hölder's inequality as

$$\Big(\frac{1}{m(Q)}\int_Q w^{1/(p-1)}\,\mathrm{d}x\Big)^{p-1} \le \frac{1}{m(Q)}\int_Q w\,\mathrm{d}x$$

and the proof is completed by choosing δ such that $1+\delta = q/(p-1)$. □

Theorem 4.31. *If $w \in A_p$, then $w \in A_{p-\varepsilon}$ for some $\varepsilon > 0$.*

Proof. First we observe that if $w \in A_p$, then (4. Exercise, problem 4)

$$w^{1-p'} \in A_{p'}.$$

Utilizing the previous theorem (Theorem 4.30) for $\left(\frac{1}{w}\right)^{p'-1} = \left(\frac{1}{w}\right)^{1/(p-1)}$, we see that

$$\Big(\frac{1}{m(Q)}\int_Q \Big(\frac{1}{w}\Big)^{(1+\delta)/(p-1)}\,\mathrm{d}x\Big)^{(p-1)/(1+\delta)} \le \Big(\frac{C}{m(Q)}\int_Q \Big(\frac{1}{w}\Big)^{1/(p-1)}\,\mathrm{d}x\Big)^{p-1}.$$

Now we can choose $\varepsilon > 0$ such that

$$\frac{p-1}{1+\delta} = (p-\varepsilon)-1$$

We utilize this and multiply the previous inequality by $\frac{1}{m(Q)}\int_Q w\,\mathrm{d}x$ to have

$$\begin{aligned}
&\frac{1}{m(Q)}\int_Q w\,\mathrm{d}x\Big(\frac{1}{m(Q)}\int_Q \Big(\frac{1}{w}\Big)^{1/((p-\varepsilon)-1)}\,\mathrm{d}x\Big)^{(p-\varepsilon)-1}\\
&\quad\le \frac{1}{m(Q)}\int_Q w\,\mathrm{d}x\Big(\frac{C}{m(Q)}\int_Q \Big(\frac{1}{w}\Big)^{1/(p-1)}\,\mathrm{d}x\Big)^{p-1}\\
&\quad\overset{w\in A_p}{\le} C.
\end{aligned}$$

Thus $w \in A_{p-\varepsilon}$. □

Next we answer the original question.

Theorem 4.32 (Muckenhoupt)**.** *Let $1 < p < \infty$. Then there exists $C > 0$ s.t.*

$$\int_{\mathbf{R}^n} \Big(Mf(x)\Big)^p w(x)\,\mathrm{d}x \le C \int_{\mathbf{R}^n} |f(x)|^p\, w(x)\,\mathrm{d}x$$

if and only if $w \in A_p$.

Proof. "$\Rightarrow$" has already been proven.

"$\Leftarrow$" We know that $w > 0$ a.e. so that

$$0 = \mu(E) = \int_E w(x)\,\mathrm{d}x \quad \Leftrightarrow \quad m(E) = 0.$$

and thus

$$\begin{aligned} ||f||_{L^\infty(\mu)} &\overset{\text{def}}{=} \inf\{\lambda \,:\, \mu(\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}) = 0\} \\ &= \inf\{\lambda \,:\, m(\{x \in \mathbf{R}^n \,:\, |f(x)| > \lambda\}) = 0\} \\ &= ||f||_\infty . \end{aligned}$$

Then

$$||Mf||_{L^\infty(\mu)} = ||Mf||_\infty \overset{\text{Lemma 2.8}}{\le} ||f||_\infty = ||f||_{L^\infty(\mu)}$$

so that M is of a weighted strong type (∞, ∞). On the other hand, by Theorem 4.27 implies that M is of weak type (p, p). Moreover, the Marcinkiewicz interpolation theorem Theorem 2.21 holds for all the measures. Thus M is of strong type (q, q) with $q > p$

$$||Mf||_{L^q(\mu)} \le C\, ||f||_{L^q(\mu)} .$$

By the previous theorem $w \in A_p$ implies that $w \in A_{p-\varepsilon}$. Thus we can repeat the above argument starting with $p - \varepsilon$ to see that

$$||Mf||_{L^p(\mu)} \le C\, ||f||_{L^p(\mu)}$$

with the original p. □

7.10.2010

5. Fourier transform

5.1. **On rapidly decreasing functions.** We define a Fourier transform of $f \in L^1(\mathbf{R})$ as

$$F(f) = \hat{f}(\xi) = \int_{\mathbf{R}} f(x) e^{-2\pi i x\xi}\,\mathrm{d}x. \tag{5.1}$$

Remark 5.2. (i) $e^{-2\pi i x\xi} = \cos(2\pi x\xi) - i\sin(2\pi x\xi)$, (even part in real, and odd in imaginary).
(ii) Theory generalizes to $\mathbf{R}^n$ (then $\mathbf{x} \cdot \xi = \sum_{i=1}^n x_i\xi_i$ and $e^{-2\pi i \mathbf{x}\cdot\xi}$).

Example 5.3 (Warning)**.** *The Fourier transform is well defined for $f \in L^1(\mathbf{R})$ because*

$$\left|f(x)e^{-2\pi i x\xi}\right| = |f(x)|$$

which is integrable. However, nothing guarantees that $\hat{f}(\xi)$ would be in $L^1(\mathbf{R})$. Indeed let $f : \mathbf{R} \to \mathbf{R}$, $f(x) = \chi_{\{-1/2,1/2\}}(x)$, which is in $L^1(\mathbf{R})$. Then for $\xi \neq 0$,

$$\begin{aligned}
\hat{f}(\xi) &= \int_{\mathbf{R}} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x \\
&= \int_{-1/2}^{1/2} e^{-2\pi i x\xi}\,\mathrm{d}x \\
&= \int_{-1/2}^{1/2} \cos(2\pi x\xi)\,\mathrm{d}x - i\underbrace{\int_{-1/2}^{1/2} \sin(2\pi x\xi)\,\mathrm{d}x}_{=0} \\
&= \Big/_{-1/2}^{1/2} \frac{\sin(2\pi x\xi)}{2\pi\xi} \\
&= \frac{2\sin(\pi\xi)}{2\pi\xi} = \frac{\sin(\pi\xi)}{\pi\xi},
\end{aligned}$$

but $\frac{\sin(\pi\xi)}{\pi\xi}$ is not integrable (the integral of the positive part $= \infty$ and the integral over the negative part $= -\infty$ over any interval $(a, \infty]$). Later, we would like to write

$$F^{-1}\hat{f}(\xi) = \int_{\mathbf{R}} \hat{f}(x)e^{2\pi i x\xi}\,\mathrm{d}x$$

for the inverse Fourier transform, which however makes no sense as such for the function that is not integrable.

The problem described in the example above does not appear for the functions that are smooth and decay rapidly at the infinity, the so called Schwartz class. Later we use the functions on the Schwartz class to define Fourier transform in L^2 and further in L^p.

Definition 5.4. A function f is in the Schwartz class $S(R)$ if

(i) $f \in C^\infty(\mathbf{R})$
(ii)

$$\sup_{x\in\mathbf{R}} |x|^k \left|\frac{d^l f(x)}{dx^l}\right| < \infty, \quad \text{for every} \quad k, l \geq 0.$$

In other words, every derivative decays at least as fast as any power of $|x|$.

Example 5.5. *The standard mollifier (as well as all of $C_0^\infty(\mathbf{R})$)*

$$\varphi = \begin{cases} \exp\left(\frac{1}{|x|^2-1}\right), & x \in (-1,1) \\ 0, & \textit{else.} \end{cases}$$

is in $S(\mathbf{R})$. Also for the Gaussian

$$f(x) = e^{-x^2} \in S(\mathbf{R}).$$

Indeed,

$$\frac{df(x)}{dx} = -2xe^{-x^2} = -2xf(x)$$

and so forth so that all the derivatives will be of the form

$$\textit{polynomial} \cdot f(x)$$

and

$$|x|^k \, |\textit{polynomial} \cdot f(x)| \le |\textit{polynomial}| \, |f(x)| \, .$$

Thus as e^{-x^2} decays faster than any polynomial, we see that $e^{-x^2} \in S(\mathbf{R})$.

Lemma 5.6. *Suppose that $f \in S(\mathbf{R})$. Then*

(i) $\widehat{(\alpha f + \beta g)} = \alpha \hat{f} + \beta \hat{g}$.

(ii) $\widehat{\left(\frac{df}{dx}\right)}(\xi) = 2\pi i \xi \hat{f}(\xi)$.

(iii) $\frac{d\hat{f}}{d\xi}(\xi) = (\widehat{-2\pi i x f})(\xi)$,

(iv) *$\hat{f}$ is continuous,*

(v) $||\hat{f}||_\infty \le ||f||_1$,

(vi) $\widehat{f(\varepsilon x)} = \frac{1}{\varepsilon}\hat{f}(\frac{\xi}{\varepsilon}) = \hat{f}_\varepsilon(\xi), \varepsilon > 0$,

(vii) $\widehat{f(x+h)} = \hat{f}(\xi)e^{2\pi i h \xi}$,

(viii) $\widehat{f(x)e^{2\pi i h x}} = \hat{f}(\xi - h)$,

Proof. (i) Integral is linear.

(ii)

$$\begin{aligned} \widehat{\left(\frac{df}{dx}\right)}(\xi) &= \int_{\mathbf{R}} \left(\frac{df}{dx}\right) e^{-2\pi i x \xi} \, \mathrm{d}x \\ &\overset{\text{integrate by parts}}{=} - \int_{\mathbf{R}} f(x) \frac{d}{dx} e^{-2\pi i x \xi} \, \mathrm{d}x \\ &= 2\pi i \xi \int_{\mathbf{R}} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x = 2\pi i \xi \hat{f}(\xi). \end{aligned}$$

(iii)

$$\begin{aligned}\frac{d\hat{f}}{d\xi}(\xi) &= \frac{d}{d\xi}\int_{\mathbf{R}} f(x)e^{-2\pi ix\xi}\,\mathrm{d}x \\ &= \int_{\mathbf{R}} f(x)\frac{d}{d\xi}e^{-2\pi ix\xi}\,\mathrm{d}x \\ &= -\int_{\mathbf{R}} f(x)2\pi ixe^{-2\pi ix\xi}\,\mathrm{d}x \\ &= (\widehat{-2\pi ixf})(\xi).\end{aligned}$$

The interchange of the derivative and integral is ok as $f \in S(\mathbf{R})$: in the detailed proof one can write down the difference quotient and estimate it by definition of $S(\mathbf{R})$.

(iv)

$$\begin{aligned}\lim_{h\to 0}\hat{f}(\xi+h) &= \lim_{h\to 0}\int_{\mathbf{R}} f(x)e^{-2\pi ix(\xi+h)}\,\mathrm{d}x \\ &\overset{\text{DOM, } \left|f(x)e^{-2\pi ix(xi+h)}\right|\le|f(x)|}{=} \int_{\mathbf{R}} f(x)\lim_{h\to 0}e^{-2\pi ix(\xi+h)}\,\mathrm{d}x = \hat{f}(\xi).\end{aligned}$$

(v)

$$\left|\int_{\mathbf{R}} f(x)e^{-2\pi ix\xi}\,\mathrm{d}x\right| \le \int_{\mathbf{R}} |f(x)|\underbrace{\left|e^{-2\pi ix\xi}\right|}_{=1}\,\mathrm{d}x.$$

(vi)

$$\begin{aligned}\widehat{f(\varepsilon x)} &= \int_{\mathbf{R}} f(\varepsilon x)e^{-2\pi ix\xi}\,\mathrm{d}x \\ &\overset{y=\varepsilon x,\,\mathrm{d}y=\varepsilon\,\mathrm{d}x}{=} \frac{1}{\varepsilon}\int_{\mathbf{R}} f(y)e^{(-2\pi iy\xi)/\varepsilon}\,\mathrm{d}y = \frac{1}{\varepsilon}\hat{f}(\frac{\xi}{\varepsilon}).\end{aligned}$$

(vii)

$$\begin{aligned}f\widehat{(x+h)} &= \int_{\mathbf{R}} f(x+h)e^{-2\pi ix\xi}\,\mathrm{d}x \\ &\overset{y=x+h,\,\mathrm{d}y=\mathrm{d}x}{=} \int_{\mathbf{R}} f(y)e^{-2\pi i(y-h)\xi}\,\mathrm{d}y = \hat{f}(\xi)e^{2\pi ih\xi}.\end{aligned}$$

(viii)

$$\begin{aligned}f\widehat{(x)e^{2\pi ihx}} &= \int_{\mathbf{R}} f(x)e^{2\pi ihx}e^{-2\pi ix\xi}\,\mathrm{d}x \\ &= \int_{\mathbf{R}} f(x)e^{-2\pi ix(\xi-h)}\,\mathrm{d}x = \hat{f}(\xi-h).\end{aligned}$$

□

Example 5.7. *If*

$$f(x) = e^{-\pi x^2}$$

then its Fourier transform is

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

By using complex integration around a rectangle and recalling that $e^{-\pi z^2}$ *is analytic function, we could calculate* $\int_{\mathbf{R}} e^{-\pi x^2} e^{-2\pi i x\xi} \,\mathrm{d}x$ *directly by using complex integration. We however follow a strategy that does not require complex integration and observe that* $f(x) = e^{-\pi x^2}$ *solves the differential equation*

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

By taking Fourier transform of $f' + 2\pi x f = 0$ *and using Lemma 5.6, we obtain*

$$0 = F(f' + 2\pi x f) = \widehat{f'} + \widehat{2\pi x f} = 2\pi i \xi \hat{f} - \frac{\hat{f}'}{i} = i(2\pi\xi\hat{f} + \hat{f}').$$

And

$$\hat{f}(0) = \int_{\mathbf{R}} e^{-\pi x^2} \,\mathrm{d}x = 1$$

because

$$\begin{aligned} \Big(\int_{\mathbf{R}} e^{-\pi x^2} \,\mathrm{d}x \Big)^2 &= \int_{\mathbf{R}} \int_{\mathbf{R}} e^{-\pi x^2} e^{-\pi x^2} \,\mathrm{d}x \,\mathrm{d}y \\ &= \int_0^\infty \int_{\partial B(0,r)} e^{-\pi r^2} \,\mathrm{d}r \,\mathrm{d}S \\ &= \int_0^\infty 2\pi r e^{-\pi r^2} \,\mathrm{d}r \\ &= - \Big/_0^\infty e^{-\pi r^2} = 1. \end{aligned}$$

Thus $\hat{f}$ *satisfies the same differential equation and the uniqueness of such a solution implies the claim.*

Theorem 5.8. *If* $f \in S(\mathbf{R})$*, then*

(i) $\hat{f} \in S(\mathbf{R})$ *(similar result does not hold in* L^1*),*

(ii)

$$F^{-1}(f) := \int_{\mathbf{R}} f(\xi) e^{2\pi i x\xi} \,\mathrm{d}\xi \in S(\mathbf{R})$$

whenever $f \in S(\mathbf{R})$*.*

Proof. (i) Recall that by Lemma 5.6, $\hat{f}$ is continuous and for any pair of integers k, l

$$\begin{aligned} F\left(\frac{1}{(2\pi i)^k}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k (-2\pi i x)^l f(x)\right) &= \frac{1}{(2\pi i)^k} F\left(\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k (-2\pi i x)^l f(x)\right) \\ &= \frac{1}{(2\pi i)^k}(2\pi i \xi)^k F\left((-2\pi i x)^l f(x)\right) \\ &= \frac{1}{(2\pi i)^k}(2\pi i \xi)^k \left(\frac{\mathrm{d}}{\mathrm{d}\xi}\right)^l \hat{f}(\xi) \\ &= \xi^k \left(\frac{\mathrm{d}}{\mathrm{d}\xi}\right)^l \hat{f}(\xi). \end{aligned}$$

Therefore

$$\begin{aligned} |\xi|^k \left|\left(\frac{\mathrm{d}}{\mathrm{d}\xi}\right)^l \hat{f}(\xi)\right| &= \left|\xi^k \left(\frac{\mathrm{d}}{\mathrm{d}\xi}\right)^l \hat{f}(\xi)\right| \\ &= \left|F\left(\frac{1}{(2\pi i)^k}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k (-2\pi i x)^l f(x)\right)\right| \\ &\overset{\text{Lemma 5.6}}{\leq} \left\|\frac{1}{(2\pi i)^k}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k (-2\pi i x)^l f(x)\right\|_1 < \infty \end{aligned}$$

so that $\hat{f} \in S(\mathbf{R})$.

(ii) This follows from the previous by a change of variable.

□

Lemma 5.9. *If $f, g \in S(\mathbf{R})$, then*

$$\int_{\mathbf{R}} \hat{f}(x) g(x)\,\mathrm{d}x = \int_{\mathbf{R}} f(x)\hat{g}(x)\,\mathrm{d}x$$

Proof.

$$\begin{aligned} \int_{\mathbf{R}} \hat{f}(y) g(y)\,\mathrm{d}y &= \int_{\mathbf{R}}\int_{\mathbf{R}} f(x) e^{-2\pi i x y}\,\mathrm{d}x\ g(y)\,\mathrm{d}y \\ &\overset{\text{Fubini}}{=} \int_{\mathbf{R}} f(x) \int_{\mathbf{R}} e^{-2\pi i x y} g(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{\mathbf{R}} f(x)\hat{g}(x)\,\mathrm{d}x. \end{aligned}$$

□

Next one of the main results of the section: inversion formula for the rapidly decreasing functions:

Theorem 5.10 (Fourier inversion)**.** *If $f \in S(\mathbf{R})$, then*

$$f(x) = \int_{\mathbf{R}} \hat{f}(y) e^{2\pi i x \xi}\,\mathrm{d}\xi,$$

or with the other notation $f(x) = F^{-1}(F(f)) = F^{-1}(\hat{f})$.

Proof. First we show that

$$f(0) = \int_{\mathbf{R}} \hat{f}(y)\,\mathrm{d}y. \tag{5.11}$$

To see this let $\phi \in S(\mathbf{R})$ and define $h(y) = f(-y)$. Then $\hat{\phi} \in S(\mathbf{R})$ and by the convergence result Theorem 3.12 (and the remark after the theorem)

$$\lim_{\varepsilon\to 0}\int_{\mathbf{R}} h(-y)\hat{\phi}_\varepsilon(y)\,\mathrm{d}y = \lim_{\varepsilon\to 0}(h * \hat{\phi}_\varepsilon)(0) = h(0) = f(0).$$

On the other hand, by Lemma 5.6 and the previous lemma

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_{\mathbf{R}} h(-y)\hat{\phi}_\varepsilon(y)\,\mathrm{d}y &= \lim_{\varepsilon\to 0}\int_{\mathbf{R}} \widehat{h(-y)}\phi(\varepsilon y)\,\mathrm{d}y \\ &\overset{h(-y)=f(y)}{=} \lim_{\varepsilon\to 0}\int_{\mathbf{R}} \hat{f}(y)\phi(\varepsilon y)\,\mathrm{d}y.\end{aligned}$$

Let $\phi(x) = e^{-\pi x^2}$, then

$$\lim_{\varepsilon\to 0}\phi(\varepsilon x) = 1, \qquad \left|\hat{f}(y)\phi(\varepsilon y)\right| \le \left|\hat{f}(\xi)\right|.$$

It follows that

$$\lim_{\varepsilon\to 0}\int_{\mathbf{R}} \hat{f}(y)\phi(\varepsilon y)\,\mathrm{d}y \overset{\text{DOM}}{=} \int_{\mathbf{R}} \hat{f}(y)\underbrace{\lim_{\varepsilon\to 0}\phi(\varepsilon y)}_{=1}\,\mathrm{d}y$$

proving (5.11). Then defining $g(x) := f(x+h)$ and using from Lemma 5.6 the fact that $\hat{g}(y) = \widehat{f(x+h)} = \hat{f}(y)e^{2\pi hy}$ and observing $g(0) = f(h)$, the equation (5.11) implies

$$f(h) = \int_{\mathbf{R}} \hat{f}(y)e^{2\pi ihy}\,\mathrm{d}y,$$

which proves the claim. □ 12.10.2010

Corollary 5.12. *Let $f \in S(\mathbf{R})$. Then by taking consecutive Fourier transforms, we obtain*

$$f(x) \xrightarrow{F} \hat{f}(\xi) \xrightarrow{F} f(-x) \xrightarrow{F} \hat{f}(-\xi) \xrightarrow{F} f(x).$$

In particular, $F^{-1}(\hat{f}) = F(F(F(\hat{f})))$.

Proof. The second arrow:

$$\begin{aligned}\int_{\mathbf{R}} \hat{f}(\xi)e^{-2\pi ix\xi}\,\mathrm{d}\xi &\overset{\xi=-\zeta}{=} \int_{\mathbf{R}} \hat{f}(-\zeta)e^{2\pi ix\zeta}\,\mathrm{d}\zeta \\ &= \int_{\mathbf{R}}\int_{\mathbf{R}} f(y)e^{-2\pi iy(-\zeta)}\,\mathrm{d}y\; e^{2\pi ix\zeta}\,\mathrm{d}\zeta \\ &\overset{y=-z}{=} \int_{\mathbf{R}}\int_{\mathbf{R}} f(-z)e^{-2\pi iz\zeta}\,\mathrm{d}z\; e^{2\pi ix\zeta}\,\mathrm{d}\zeta = f(-x).\end{aligned}$$

The other arrows are easier. □

Lemma 5.13. *If $f, g \in S(\mathbf{R})$, then*

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

Proof. The proof is based on Fubini's theorem. To this end, observe that by the proof of Young's inequality for convolution, Theorem 3.2, we have

$$\int_{\mathbf{R}} \int_{\mathbf{R}} \left| f(y)g(x-y)\ e^{-2\pi i x\xi} \right| \,\mathrm{d}y \,\mathrm{d}x = \int_{\mathbf{R}} |f(y)| \int_{\mathbf{R}} |g(x-y)| \,\mathrm{d}x \,\mathrm{d}y < \infty.$$

Now we can calculate

$$\begin{aligned} \widehat{f * g} &= \int_{\mathbf{R}} \int_{\mathbf{R}} f(y)g(x-y) \,\mathrm{d}y\ e^{-2\pi i x\xi} \,\mathrm{d}x \\ &\overset{\text{Fubini}}{=} \int_{\mathbf{R}} f(y) \int_{\mathbf{R}} g(x-y) e^{-2\pi i x\xi} \,\mathrm{d}x \,\mathrm{d}y \\ &\overset{x-y=z,\, \mathrm{d}x=\mathrm{d}z}{=} \int_{\mathbf{R}} f(y) \int_{\mathbf{R}} g(z) e^{-2\pi i (z+y)\xi} \,\mathrm{d}z \,\mathrm{d}y \\ &= \int_{\mathbf{R}} f(y) e^{-2\pi i y\xi} \,\mathrm{d}y \int_{\mathbf{R}} g(z) e^{-2\pi i z\xi} \,\mathrm{d}z = \hat{f}\hat{g}. \end{aligned}$$

□

Next we prove Plancherel's theorem. The theorem plays a central role, when extending the definition of the Fourier transform to the L^2-functions. It will also be needed in connection to singular integrals.

Theorem 5.14 (Plancherel)**.** *If $f \in S(\mathbf{R})$, then*

$$||f||_2 = ||\hat{f}||_2. \tag{5.15}$$

Proof. Set $g = \overline{\hat{f}}$. Then $\hat{g} = \overline{f}$. To see this, we first calculate

$$\begin{aligned} g = \overline{\hat{f}} &= \overline{\int_{\mathbf{R}} f(x) e^{-2\pi i x\xi} \,\mathrm{d}x} \\ &= \int_{\mathbf{R}} \overline{f}(x) e^{2\pi i x\xi} \,\mathrm{d}x \\ &= \int_{\mathbf{R}} \overline{f}(x) e^{-2\pi i x(-\xi)} \,\mathrm{d}x = \widehat{\overline{f}}(-\xi) \end{aligned}$$

and thus by Corollary 5.12

$$\hat{g}(x) = F(\widehat{\overline{f}}(-\xi))(x) = \overline{f}(x).$$

Utilizing this and Lemma 5.9, we have

$$\begin{aligned} ||f||_2 &= \int_{\mathbf{R}} f(x)\overline{f}(x) \,\mathrm{d}x = \int_{\mathbf{R}} f(x)\hat{g}(x) \,\mathrm{d}x \\ &\overset{\textit{Lemma } 5.9}{=} \int_{\mathbf{R}} \hat{f}(x) g(x) \,\mathrm{d}x = \int_{\mathbf{R}} \hat{f}(x)\overline{\hat{f}}(x) \,\mathrm{d}x = ||\hat{f}||_2. \end{aligned}$$

□

5.2. **On L^1.** As stated above for $f \in L^1(\mathbf{R})$, the Fourier transform $\hat{f}(\xi) = \int_{\mathbf{R}} f(x)e^{-2\pi ix\xi}\,\mathrm{d}x$ is well defined but it might well be that $\hat{f} \notin L^1(\mathbf{R})$.

Question: Then how do we obtain f from $\hat{f}$ in this case as $\int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi ix\xi}\,\mathrm{d}\xi$ might not be well defined?

The answer is that we can make sure that the inversion formula makes sense by multiplying by a bumb function which makes sure that the integrand gets small enough values far away, and then pass to a limit.

Theorem 5.16. *Let $\phi \in L^1(\mathbf{R})$, be bounded and continuous with $\hat{\phi} \in L^1(\mathbf{R})$, $||\hat{\phi}||_1 = 1$. Then*

$$\lim_{\varepsilon\to 0}\left|\left|\int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi ix\xi}\phi(-\varepsilon\xi)\,\mathrm{d}\xi - f(x)\right|\right|_1 = 0.$$

A suitable ϕ in the theorem above is for example $\phi(x) = e^{-\pi x^2}$, see Example 5.7.

Proof. First, we show that

$$\int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi ix\xi}\phi(-\varepsilon\xi)\,\mathrm{d}\xi = (f * \hat{\phi}_\varepsilon)(x).$$

To this end, recall that $\widehat{\phi(-\varepsilon x)} = \hat{\phi}_\varepsilon(-\xi)$ and $\widehat{f(x)e^{2\pi ihx}} = \hat{f}(\xi - h)$ by Lemma 5.6. Observe that these results hold also for L^1 functions. Since ϕ is bounded also the proof of Lemma 5.9 holds. Thus

$$\begin{aligned}
\int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi ix\xi}\phi(-\varepsilon\xi)\,\mathrm{d}\xi &= \int_{\mathbf{R}}\int_{\mathbf{R}} f(y)e^{-2\pi iy\xi}\,\mathrm{d}y\; e^{2\pi ix\xi}\phi(-\varepsilon\xi)\,\mathrm{d}\xi \\
&\overset{\text{Lemma 5.9}}{=} \int_{\mathbf{R}} f(y)\int_{\mathbf{R}} \big(e^{2\pi ix\xi}\phi(-\varepsilon\xi)\big)e^{-2\pi iy\xi}\,\mathrm{d}\xi\,\mathrm{d}y \\
&= \int_{\mathbf{R}} f(y)F\big(e^{2\pi ix\xi}\phi(-\varepsilon\xi)\big)(y)\,\mathrm{d}y \qquad (5.17)\\
&\overset{\text{Lemma 5.6:(vi),(viii)}}{=} \int_{\mathbf{R}} f(y)\,\hat{\phi}_\varepsilon(x-y)\,\mathrm{d}y \\
&= (f * \hat{\phi}_\varepsilon)(x).
\end{aligned}$$

When dealing with convolutions, we showed in Theorem 3.7 that

$$(f * \hat{\phi}_\varepsilon)(x) \to f(x) \quad \text{in } L^1(R).$$ □

If $\hat{f} \in L^1(\mathbf{R})$, then the inversion formula $f(x) = \int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi ix\xi}\,\mathrm{d}\xi$ works as such. This can be seen by adding a condition $\phi(0) = 1$ for the bumb function and passing to limit in (5.17) using Lebesgue's dominated convergence on the left.

5.3. **On L^2.**

Theorem 5.18. *Let $f \in L^2(\mathbf{R}^n)$, and $\phi_j \in S(\mathbf{R})$, $j = 1, 2, \dots$ such that*

$$\lim_{j\to\infty} ||\phi_j - f||_2 = 0.$$

Then there exists a limit which we denote by $\hat{f}$ such that

$$\lim_{j\to\infty} ||\hat{\phi}_j - \hat{f}||_2 = 0.$$

The function $\hat{f}$ is called a Fourier transform of $f \in L^2(\mathbf{R})$.

Proof. First of all, there exists a sequence $\phi_j \in S(\mathbf{R})$, $j = 1, 2, \dots$ such that

$$\lim_{j\to\infty} ||\phi_j - f||_2 = 0$$

because $S(\mathbf{R})$ is dense in $L^2(\mathbf{R})$: We have already seen that $C_0(\mathbf{R})$ is dense in $L^2(\mathbf{R})$. On the other hand, if $f \in C_0(\mathbf{R})$ then $C_0^\infty(\mathbf{R}) \ni f * \phi_\varepsilon \to f$ in $L^2(\mathbf{R})$, where ϕ_ε is a standard mollifier, and we see that $C_0^\infty(\mathbf{R})$ is dense in $L^2(\mathbf{R})$, which is contained in $S(\mathbf{R})$.

Then by Plancherel's theorem

$$||\hat{\phi}_j - \hat{\phi}_k||_2 = ||\phi_j - \phi_k||_2 \to 0$$

as $j, k \to \infty$ and thus $\hat{\phi}_j$, $j = 1, 2, \dots$ is a Cauchy sequence. Since $L^2(\mathbf{R})$ is complete, $\hat{\phi}_j$ converges to a limit, which we denote by $\hat{f}$.

Next we show that the limit is independent of the approximating sequence. Let φ_j be another sequence such that

$$\varphi_j \to f \quad \text{in} \quad L^2(\mathbf{R})$$

and let $g \in L^2(\mathbf{R})$ be the limit

$$\hat{\varphi}_j \to g \quad \text{in} \quad L^2(\mathbf{R}).$$

Then

$$0 \overset{\phi_j, \varphi_j \to f}{=} \lim_{j\to 0} ||\varphi_j - \phi_j||_2 \overset{\text{Plancherel}}{=} \lim_{j\to 0} ||\hat{\varphi}_j - \hat{\phi}_j||_2 = ||g - \hat{f}||_2. \qquad \square$$

Similarly we obtain a unique inverse Fourier transform of any L^2-function.

We state separately a result from the previous proof.

Corollary 5.19 (Plancerel in L^2)**.** *If $f \in L^2(\mathbf{R})$, then*

$$||f||_2 = ||\hat{f}||_2.$$

Proof.

$$||f||_2 = \lim_{j\to\infty} ||\phi_j||_2 = \lim_{j\to\infty} ||\hat{\phi}_j||_2 = ||\hat{f}||_2.$$

$\square$

We also obtain formulas for calculating the Fourier transform and the inverse Fourier transform for L^2-functions. Observe that in the corollary below, $\chi_{B(0,R)} f \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ by Hölder's inequality since $⨍_B |f| \,\mathrm{d}x \leq \big(⨍_B |f|^2 \,\mathrm{d}x \big)^{1/2}$.

Corollary 5.20. *If $f \in L^2(\mathbf{R})$, then*

$$\lim_{R\to\infty} \left|\left| \int_{\{|x|<R\}} f(x) e^{-2\pi i x\xi} \,\mathrm{d}x - \hat{f} \right|\right|_2 = 0,$$

and

$$\lim_{R\to\infty} \left|\left| \int_{\{|\xi|<R\}} \hat{f}(\xi) e^{2\pi i x\xi} \,\mathrm{d}\xi - f(x) \right|\right|_2 = 0.$$

Proof. Recall that if $f \in L^2(\mathbf{R})$, then $\chi_{B(0,R)} f \to f$ in $L^2(\mathbf{R})$ by Lebesgue's monotone/dominated convergence theorem. Let us denote

$$\lim_{R\to\infty} \int_{\{|x|<R\}} f(x) e^{-2\pi i x\xi} \,\mathrm{d}x = \lim_{R\to\infty} F(f\chi_{B(0,R)}).$$

The convergence $F(f\chi_{B(0,R)}) \to \hat{f}$ follows from the Plancherel's theorem, because the right hand side of

$$\left|\left| F(f\chi_{B(0,R)}) - \hat{f} \right|\right|_2 = \left|\left| f\chi_{B(0,R)} - f \right|\right|_2$$

can be made as small as we please by choosing R large enough. The proof of the inversion formula is similar. □

5.4. **On L^p, $1 < p < 2$.** Fourier transform is a linear operator and thus for $f \in L^p(\mathbf{R})$, $1 < p < 2$, we have

$$f = f_1 + f_2 = f\chi_{\{|f|>\lambda\}} + f\chi_{\{|f|\leq\lambda\}} \in L^1 + L^2.$$

we have $\hat{f} = \hat{f}_1 + \hat{f}_2 \in L^\infty + L^2$ and

$$\lim_{R\to\infty} \int_{\{|x|<R\}} f(x) e^{-2\pi i x\xi} \,\mathrm{d}x,$$

can also be utilized here. However by a special case of the *Riesz-Thorin interpolation theorem* we obtain even better. We omit the proof.

Theorem 5.21 (Riesz-Thorin interpolation)**.** *Let T be a linear operator*

$$T : L^1(\mathbf{R}) + L^2(\mathbf{R}) \to L^\infty(\mathbf{R}) + L^2(\mathbf{R})$$

such that

$$||Tf_1||_\infty \leq C_1 \,||f_1||_1$$

for every $f_1 \in L^1(\mathbf{R})$, and

$$||Tf_2||_2 \leq C_2 \,||f_2||_2 \,,$$

for every $f_2 \in L^2(\mathbf{R})$. *Then*

$$||Tf||_{p'} \leq C_1^{1-2/p'} C_2^{2/p'} ||f||_p \,,$$

where $1/p + 1/p' = 1$.

Corollary 5.22 (Hausdorff-Young inequality)**.** *If* $f \in L^p(\mathbf{R})$, $1 \leq p \leq 2$, *then* $\hat{f} \in L^{p'}(\mathbf{R})$ *and*

$$||\hat{f}||_{p'} \leq ||f||_p \,.$$

Proof. By Lemma 5.6, we have $\left|\left|\hat{f}\right|\right|_\infty \leq ||f||_1$ and by Plancherel's theorem $\left|\left|\hat{f}\right|\right|_2 = ||f||_2$. Thus we can use Riesz-Thorin interpolation. □

Observe however that obtaining f from $\hat{f}$ by using

$$\hat{f}(\xi) = \lim_{R \to \infty} \int_{\{|x|<R\}} f(x) e^{-2\pi i x \xi} \, \mathrm{d}x,$$

is a nontrivial problem. For example in the case $p = 1$ the Fourier transform of $\chi_{B(0,R)}$ is not in L^1 as shown in Example 5.3, it does not satisfy the assumptions of Theorem 5.16, and thus our results do not imply the convergence. In higher dimensions there is no, in general,
14.10.2010 the convergence in L^p, $p \neq 2$, as $R \to \infty$.

6. Singular integrals

In this section we consider integral operators of type

$$Tf(x) = \int_{\mathbf{R}^n} \frac{K(x,y)}{|x-y|^n} f(y)\, dx.$$

If $K(x,y) = K(x-y)$, we say that the operator is of convolution type.

Let us motivate the study of such operators. Indeed, integral operators naturally arise in the analysis of partial differential equations. Consider, for example, the Poisson equation

$$-\Delta u = f \qquad \text{in } \mathbf{R}^n, \qquad n \geq 3,$$

where, for simplicity, we assume that $f : \mathbf{R}^n \mapsto \mathbf{R}$ is smooth compactly supported function. Then the global solution is obtained via convolution

$$u(x) = c_n \int_{\mathbf{R}^n} \frac{f(y)}{|x-y|^{n-2}}\, dy =: I_0(f).$$

Here c_n is a constant depending on the dimension n. Proceeding formally, we may take the k^{th} partial derivative and obtain

$$\frac{\partial u}{\partial x_k}(x) = c_n \int_{\mathbf{R}^n} f(y) \frac{\partial}{\partial x_k} |x-y|^{2-n}\, dy =: I_k(f)$$

and

$$\frac{\partial^2 u}{\partial x_k \partial x_m}(x) = c_n \int_{\mathbf{R}^n} f(y) \frac{\partial^2}{\partial x_k \partial x_m} |x-y|^{2-n}\, dy =: I_{km}(f)(x).$$

A direct calculation gives

$$\frac{\partial}{\partial x_k} |x-y|^{2-n} = -(n-2) \frac{x_k - y_k}{|x-y|^n}$$

and

$$\frac{\partial^2}{\partial x_k \partial x_m} |x-y|^{2-n} = -(n-2) \frac{1}{|x-y|^n} \delta_{km} + n(n-2) \frac{(x_k - y_k)(x_m - y_m)}{|x-y|^{n+2}},$$

where δ_{km} stands for the Kronecker delta function. Observe that the kernel of I_{ij} is not in $L^1(\mathbf{R}^n)$ and we need to carefully define in what sense the operator takes values.

Now, when considering the regularity of u, i.e. properties of first and second derivatives, we are led to the analysis of mapping properties of singular integral operators I_i and I_{ij}. For example, a relevant question is that if f above belongs to $L^p(\mathbf{R}^n)$, do second derivatives of u belong to $L^p(\mathbf{R}^n)$ as well?

We start our journey to the fascinating world of singular integral operators considering two model cases: Hilbert and Riesz transforms.

6.1. **Hilbert transform.** The Hilbert transform of $f \in L^p(\mathbf{R})$, $1 \le p < \infty$, is defined as

$$(Hf)(x) = \text{p.v.}\frac{1}{\pi}\int_{\mathbf{R}} \frac{f(t)}{x-t}\,dt := \lim_{\varepsilon \downarrow 0} \frac{1}{\pi}\int_{|x-t|>\varepsilon} \frac{f(t)}{x-t}\,dt.$$

Denote

$$(H^{(\varepsilon)}f)(x) := \frac{1}{\pi}\int_{|x-t|>\varepsilon} \frac{f(t)}{x-t}\,dt.$$

Observe immediately that since $1/t$ does not belong to L^1, we are not in position to apply Theorem 3.12 to obtain the existence of the limit above. However, when $f \in \mathcal{S}(\mathbf{R})$, then the limit exists.

Lemma 6.1. *Let $f \in S(\mathbf{R})$. Then the limit*

$$(Hf)(x) = \lim_{\varepsilon \downarrow 0}(H^{(\varepsilon)}f)(x)$$

exists for all $x \in \mathbf{R}$.

Proof. Since $(x-t)^{-1}\chi_{\{\varepsilon<|x-t|<\delta\}}$ is an odd function, we have

$$\frac{1}{\pi}\int_{\varepsilon<|x-t|<\delta} \frac{f(x)}{x-t}\,dt = \frac{f(x)}{\pi}\int_{\varepsilon<|x-t|<\delta} \frac{1}{x-t}\,dt = 0$$

for all $0 < \varepsilon < \delta$. Thus also

$$(H^{(\varepsilon)}f)(x) = \frac{1}{\pi}\int_{\varepsilon<|x-t|<\delta} \frac{f(t)-f(x)}{x-t}\,dt + \frac{1}{\pi}\int_{|x-t|\ge\delta} \frac{f(t)}{x-t}\,dt$$

holds for all $0 < \varepsilon < \delta$ and now the limit on the right, as $\varepsilon \downarrow 0$, exists by Lebesgue's dominated convergence since

$$\begin{aligned}\left|\chi_{\{\varepsilon<|x-t|<\delta\}}\frac{f(t)-f(x)}{x-t}\right| &\le \chi_{\{|x-t|<\delta\}} \sup_{(x,t)\in\mathbf{R}\times\mathbf{R}} \frac{|f(x)-f(t)|}{|x-t|} \\ &\le \chi_{\{|x-t|<\delta\}}\|f'\|_\infty\end{aligned}$$

and the second integral converges for fixed $\delta > 0$ by the decay of f, i.e. that there is a constant c_x such that $\sup_{t\in\mathbf{R}}|(x-t)f(t)| \le c_x$. $\square$

The proof shows that we also have a uniform bound

$$|(H^{(\varepsilon)}f)(x)| \le \frac{\delta}{\pi}\|f'\|_\infty + \frac{2}{\pi\delta}\|f\|_1 \tag{6.2}$$

for all $\delta > 0$ and $x \in \mathbf{R}$. It also follows that H operating on $S(\mathbf{R})$ is a linear operator.

Our primary goal is to extend the existence of the limit for all L^p-functions and deduce that H is a linear operator from $L^p(\mathbf{R})$ to $L^p(\mathbf{R})$ for every $1 < p < \infty$. The plan to do this is as follows: first we study the behavior of H operating on $S(\mathbf{R})$. We find that for every $p = 2$ there is an extension for H from $S(\mathbf{R})$ to the whole of $L^2(\mathbf{R})$. Using this we show that H satisfies weak $(1,1)$ and strong (p,p) when restricted to $S(\mathbf{R})$. Then we will find suitable bounds for a maximal type operator

$H^{(*)}f := \sup_{\varepsilon>0} |H^{(\varepsilon)}f|$ and then, using obtained bounds, infer the existence of the limit almost everywhere.

6.1.1. *Fourier transform of the Hilbert transform.* We start by studying the Fourier transform of Hf. At this stage, the most important implication of the next theorem is that H is strong $(2,2)$ when restricting functions in $S(\mathbf{R})$ allowing for an extension to the whole $L^2(\mathbf{R})$.

Theorem 6.3. *Let $f \in S(\mathbf{R})$. Then we have*

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

for every $\xi \in \mathbf{R}$ and hence $\|Hf\|_2 = \|f\|_2$. Here

$$\mathrm{sgn}\,\xi := \begin{cases} 1, & \xi > 0, \\ 0, & \xi = 0, \\ -1, & \xi < 0. \end{cases}$$

As a corollary *we may define the extension $\widetilde{H}$ of H* to the whole $L^2(\mathbf{R})$ via the inverse Fourier transform

$$(\widetilde{H}f)(x) = -i\int_{\mathbf{R}} \mathrm{sgn}(\xi)\hat{f}(\xi)e^{2\pi i x\xi}\,\mathrm{d}\xi.$$

Note carefully, however, that this extension does not yet tell anything about the existence of the pointwise limit $\lim_{\varepsilon\downarrow 0} H^{(\varepsilon)}f$ when $f \in L^2(\mathbf{R})$. Collecting facts:

Corollary 6.4. *The Hilbert transform H allows for a unique extension $\widetilde{H}$ from $S(\mathbf{R})$ to $L^2(\mathbf{R})$ such that $\widetilde{H}$ is a (linear) operator from $L^2(\mathbf{R})$ to $L^2(\mathbf{R})$, $\widetilde{H}f = Hf$ for all $f \in S(\mathbf{R})$ and if $f \in L^2(\mathbf{R})$, then*

$$\|\widetilde{H}f\|_2 = \|f\|_2, \qquad \widehat{(\widetilde{H}f)}(\xi) = -i\,\mathrm{sgn}(\xi)\widehat{f}(\xi).$$

Proof of Theorem 6.3. To begin with, define the approximating operator

$$\widetilde{f}_{\varepsilon,\omega}(x) := (H^{(\varepsilon,\omega)}f)(x) = \frac{1}{\pi}\int_{\varepsilon<|x-t|<\omega} \frac{f(t)}{x-t}\,dt = (f * K_{\varepsilon,\omega})(x),$$

where $0 < \varepsilon < 1$ and

$$K_{\varepsilon,\omega}(x) = \frac{1}{\pi}\frac{\chi_{\varepsilon<|x|<\omega}}{x}.$$

Clearly $K_{\varepsilon,\omega} \in L^p(\mathbf{R})$ for all $1 \le p \le \infty$. Therefore

$$\widehat{\widetilde{f}_{\varepsilon,\omega}} = \widehat{K_{\varepsilon,\omega}}\widehat{f}$$

holds by Lemma 5.13, in view of Theorem 5.18, almost everywhere. A straightforward calculation gives

$$\begin{aligned}
\widehat{K_{\varepsilon,\omega}}(\xi) =& \frac{1}{\pi}\int_{\varepsilon<|x|<\omega} \frac{e^{-2\pi i\xi x}}{x}\,dx \\
=& \frac{1}{\pi}\int_{\varepsilon<|x|<\omega} \frac{1}{x}\left(\cos(2\pi x\xi) - i\sin(2\pi x\xi)\right)\,dx \\
=& -\frac{2i}{\pi}\int_{\varepsilon}^{\omega} \frac{1}{x}\sin(2\pi x\xi)\,dx,
\end{aligned}$$

where we used the fact that $\cos(ax)/x$ and $\sin(ax)/x$ are odd and even functions, respectively, for all $a \in \mathbf{R}$. Changing the variables as $t = 2\pi x\xi$, we obtain

$$\widehat{K_{\varepsilon,\omega}}(\xi) = -\frac{2i}{\pi}\operatorname{sgn}(\xi)\int_{2\pi|\xi|\varepsilon}^{2\pi|\xi|\omega} \frac{\sin(t)}{t}\,dt.$$

The integral on the right has a limit as the following lemma reveals.

Lemma 6.5.

$$\lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\int_{\varepsilon}^{\omega}\frac{\sin(t)}{t}\,dt = \frac{\pi}{2}, \qquad \left|\int_{\varepsilon}^{\omega}\frac{\sin(t)}{t}\,dt\right| \le c \quad \forall 0<\varepsilon<\omega$$

for a constant c independent of ε and ω.

We postpone the proof. Deduce using the lemma that

$$\lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\widehat{K_{\varepsilon,\omega}}(\xi) = -i\operatorname{sgn}(\xi)$$

and

$$\sup_{0<\varepsilon<\omega\le\infty}\left|\widehat{K_{\varepsilon,\omega}}(\xi)\right| \le c. \tag{6.6}$$

Consequently also

$$\lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\|\widehat{\widetilde{f}_{\varepsilon,\omega}} - (-i\operatorname{sgn}(\xi)\widehat{f})\|_2 = \lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\|(\widehat{K_{\varepsilon,\omega}} + i\operatorname{sgn}(\xi))\widehat{f}\|_2 = 0$$

holds by the Lebesgue's dominated convergence. Let then $g \in L^2(\mathbf{R})$ be the inverse Fourier transform of $i\operatorname{sgn}(\xi))\widehat{f}$, which certainly is well-defined by Corollary 5.20 since $i\operatorname{sgn}(\xi)\widehat{f} \in L^2(\mathbf{R})$. Plancerel's identity, i.e. Corollary 5.19, implies that

$$\lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\|\widetilde{f}_{\varepsilon,\omega} - g\|_2 = \lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\|\widehat{\widetilde{f}_{\varepsilon,\omega}} - \widehat{g}\|_2 = 0$$

and

$$\|g\|_2 = \|f\|_2.$$

Let us now calculate an integrable upper bound for $\widetilde{f}^2_{\varepsilon,\omega}$ independent of ω. Suppose first that $|x| > 2\varepsilon$. Then

$$\begin{aligned}|\widetilde{f}_{\varepsilon,\omega}(x)| \leq & \frac{1}{\pi}\int_{\varepsilon<|x-t|<|x|/2} \frac{|f(t)|}{|x-t|}\,dt + \frac{1}{\pi}\int_{|x-t|\geq|x|/2} \frac{|f(t)|}{|x-t|}\,dt \\ \leq & \frac{1}{\pi}\int_{\varepsilon<|x-t|<|x|/2} \frac{|f(t)|}{|x-t|}\,dt + \frac{2}{\pi|x|}\|f\|_1\end{aligned}$$

If $|x-t| < |x|/2$, then $|x| \leq |t| \leq (3/2)|x|$ and therefore

$$\begin{aligned}\int_{\varepsilon<|x-t|<|x|/2} \frac{|f(t)|}{|x-t|}\,dt = & \int_{\varepsilon<|x-t|<|x|/2} \frac{|f(t)|}{|x-t|}\frac{|t|}{|x|}\frac{|x|}{|t|}\,dt \\ \leq & \frac{1}{\varepsilon|x|}\sup_{t\in\mathbf{R}}|tf(t)|\end{aligned}$$

holds. Since $f \in S(\mathbf{R})$, $\sup_{t\in\mathbf{R}}|tf(t)| < \infty$. If $|x| < 2\varepsilon$, then estimate simply as $|\widetilde{f}_{\varepsilon,\omega}(x)| \leq \|f\|_1/\varepsilon$. In all cases we have

$$|\widetilde{f}_{\varepsilon,\omega}(x)| \leq \frac{c}{1+|x|}$$

for a constant c depending on f and ε, but independent of ω. It follows that

$$(\widetilde{f}_{\varepsilon,\omega}(x) - g(x))^2 \leq \frac{2c^2}{(1+|x|)^2} + 2g(x)^2$$

and the right hand side is integrable. Thus Lebesgue's dominated convergence gives

$$\lim_{\omega\uparrow\infty}\|\widetilde{f}_{\varepsilon,\omega} - g\|_2 = \|\lim_{\omega\uparrow\infty}\widetilde{f}_{\varepsilon,\omega} - g\|_2$$

On the other hand, Lebesgue's dominated convergence[2] implies

$$\begin{aligned}\lim_{\omega\uparrow\infty}\widetilde{f}_{\varepsilon,\omega}(x) = & \lim_{\omega\uparrow\infty}\frac{1}{\pi}\int_{\varepsilon<|x-t|<\omega}\frac{f(t)}{x-t}\,dt \\ = & \frac{1}{\pi}\int_{\varepsilon<|x-t|}\frac{f(t)}{x-t}\,dt = H^{(\varepsilon)}f(x).\end{aligned}$$

Thus we conclude with

$$\lim_{\varepsilon\downarrow 0}\|H^{(\varepsilon)}f - g\|_2 = 0.$$

It follows that *there is a subsequence* (ε_k) such that $\varepsilon_k \downarrow 0$ and $H^{(\varepsilon_k)}f \to g$ almost everywhere as $k \to \infty$. But now we know by Lemma 6.1 that the limit Hf exists and thus

$$Hf = \lim_{\varepsilon\downarrow 0}H^{(\varepsilon)}f = \lim_{k\uparrow\infty}H^{(\varepsilon_k)}f = g$$

almost everywhere. $\square$

[2] $|\chi_{\varepsilon<|x-t|<\omega}f(t)/(x-t)| \leq f(t)/\varepsilon \in L^1(\mathbf{R})$

Remark 6.7. While proving the last theorem, we also obtained (6.6) implying

$$|(\widehat{H^{(\varepsilon)}f})(\xi)| \le c|\widehat{f}(\xi)| \qquad \forall \xi \in \mathbf{R}$$

for a constant c independent of ε and f. In particular, $H^{(\varepsilon)}$ is strong $(2,2)$ by Plancerel's equality for a constant independent of ε.

Proof of Lemma 6.5. Observe first that $|\sin(t)/t| \le e$ and hence

$$\left|\int_0^{\varepsilon} \frac{\sin(t)}{t}\,dt\right| \le e\varepsilon$$

for all $\varepsilon > 0$. On the other hand, rewriting

$$\int_\omega^\infty \frac{\sin(t)}{t}\,dt = \sum_{k=0}^\infty \int_{\omega+2\pi k}^{\omega+2\pi(k+1)} \frac{\sin(t)}{t}\,dt$$

and then changing variables as

$$\int_{\omega+\pi(2k+1)}^{\omega+2\pi(k+1)} \frac{\sin(t)}{t}\,dt = -\int_{\omega+2\pi k}^{\omega+\pi(2k+1)} \frac{\sin(t)}{t+\pi}\,dt,$$

we obtain

$$\int_{\omega+2\pi k}^{\omega+2\pi(k+1)} \frac{\sin(t)}{t}\,dt = \pi\int_{\omega+2\pi k}^{\omega+\pi(2k+1)} \frac{\sin(t)}{t(t+\pi)}\,dt.$$

Therefore

$$\left|\int_\omega^\infty \frac{\sin(t)}{t}\,dt\right| \le \pi\int_\omega^\infty \frac{1}{t(t+\pi)}\,dt \le \frac{\pi}{\omega}$$

and the term on the right tends to zero as $\omega \uparrow \infty$. Thus we have

$$\lim_{\varepsilon\downarrow 0}\lim_{\omega\uparrow\infty}\int_\varepsilon^\omega \frac{\sin(t)}{t}\,dt = \int_0^\infty \frac{\sin(t)}{t}\,dt.$$

We now proceed in calculating the value of this integral.

To this end, define

$$I(a) := \int_0^\infty \frac{\sin(t)}{t}e^{-at}\,dt, \qquad a > 0.$$

Differentiate with respect to a to get

$$I'(a) = -\int_0^\infty \sin(t)e^{-at}\,dt.$$

One can easily make differentiation rigorous by considering differential quotients and using the fact $a > 0$. Integration by parts gives

$$\int_0^\infty \sin(t)e^{-at}\,dt = -\frac{1}{a}\int_0^\infty \cos(t)e^{-at}\,dt = \frac{1}{a^2} - \frac{1}{a^2}\int_0^\infty \sin(t)e^{-at}\,dt$$

and hence

$$I'(a) = -\frac{1}{1+a^2} \quad \iff \quad I(a) = -\arctan(a) + c, \qquad a > 0,\ c \in \mathbf{R}.$$

Since $I(a) \to 0$ as $a \to \infty$, we have $c = \pi/2$. If we now can show that I is continuous on $[0, \infty)$, we may deduce that

$$\int_0^\infty \frac{\sin(t)}{t}\, dt = I(0) = \frac{\pi}{2}.$$

The only point we have to check is $a = 0$. Let $0 < a < 10^{-2010}$. Rewrite the difference as

$$\begin{aligned} I(a) - I(0) &= \int_0^\infty \frac{\sin(t)}{t}\left(e^{-at} - 1\right)\, dt \\ &= \int_0^{2\pi N} \frac{\sin(t)}{t}\left(e^{-at} - 1\right)\, dt + \sum_{k=N}^\infty \int_{2\pi k}^{2\pi(k+1)} \frac{\sin(t)}{t}\left(e^{-at} - 1\right)\, dt. \\ &=: J_1 + J_2 \end{aligned}$$

for some $N \in \mathbf{N}$. Note that N is at our disposal. We will estimate J_1 and J_2 separately by means of a.

First, using the Taylor expansion

$$\frac{1}{t}\left(e^{-at} - 1\right) = \frac{1}{t}\sum_{k=1}^\infty \frac{(-at)^k}{k!},$$

we obtain by integration by parts that

$$\begin{aligned} b_j :=& -\frac{a}{j!}\int_0^{2\pi N} \sin(t)(-at)^{j-1}\, dt \\ =& a\frac{(-2\pi N a)^{j-1}}{j!} + \frac{(-a)^{j-1}}{j(j-2)!}\int_0^{2\pi N} \cos(t) t^{j-2}\, dt \\ =& a\frac{(-2\pi N a)^{j-1}}{j!} - \frac{(-a)^{j-1}}{j(j-3)!}\int_0^{2\pi N a} \sin(t/a) t^{j-3}\, dt \\ =& -\frac{1}{2\pi N}\frac{(-2\pi N a)^j}{j!} - \frac{(-a)^{j-1}(j-2)}{j} b_{j-2} \end{aligned} \tag{6.8}$$

for any $j > 2$. Similarly,

$$b_1 = 0, \qquad b_2 = -\frac{2\pi N a^2}{2}.$$

In particular, we have

$$J_1 = \sum_{j=2}^\infty b_j.$$

Choose now N be the smallest integer larger than $1/(8\pi a)$. By denoting $\beta_j := |b_j|$ we deduce by (6.8) that

$$\beta_j \le a\frac{2^{2-j}}{j!} + a\beta_{j-2}$$

for all $j > 2$. Sum the estimate over j to obtain

$$\sum_{j=3}^{\infty} \beta_j \leq a4e^{1/2} + a\sum_{j=1}^{\infty} \beta_j,$$

readily implying

$$|J_1| = \left|\sum_{j=2}^{\infty} b_j\right| \leq \sum_{j=2}^{\infty} \beta_j \leq \frac{a}{1-a}\left(4e^{1/2} + \beta_2\right) + \beta_2 \leq ca$$

for some constant c independent of a.

Second, by changing variables as in the beginning of the proof we obtain

$$\begin{aligned}\int_{(2k+1)\pi}^{(2k+2)\pi} \frac{\sin(t)}{t}\left(e^{-at} - 1\right)\, dt &= \int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t+\pi)}{t+\pi}\left(e^{-a(t+\pi)} - 1\right)\, dt \\ &= -\int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t)}{t+\pi}\left(e^{-a(t+\pi)} - 1\right)\, dt\end{aligned}$$

for any $k \geq N$. Applying further the identity

$$\frac{e^{-at} - 1}{t} - \frac{e^{-a(t-\pi)} - 1}{t+\pi} = \frac{e^{-at}}{t+\pi}\left(1 - e^{a\pi}\right) - \frac{\pi}{t(t+\pi)},$$

we arrive at

$$\begin{aligned}\int_{2k\pi}^{2(k+1)\pi} \frac{\sin(t)}{t}\left(e^{-at} - 1\right)\, dt = (1 - e^{a\pi}) &\int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t)}{t+\pi} e^{-at}\, dt \\ - \pi &\int_{2k\pi}^{(2k+1)\pi} \frac{\sin(t)}{t(t+\pi)}\, dt.\end{aligned}$$

Using then the fact that $k \geq N$ this leads to

$$\begin{aligned}&\left|\int_{2k\pi}^{2(k+1)\pi} \frac{\sin(t)}{t}\left(e^{-at} - 1\right)\, dt\right| \\ &\qquad \leq \frac{e^{a\pi} - 1}{2\pi N} \int_{2k\pi}^{2(k+1)\pi} e^{-at}\, dt + \frac{\pi}{\sqrt{2\pi N}} \int_{2k\pi}^{2(k+1)\pi} \frac{1}{\sqrt{t}(t+\pi)}\, dt\end{aligned}$$

and we conclude with

$$|J_2| \leq \frac{e^{a\pi} - 1}{2\pi N} \int_1^{\infty} e^{-at}\, dt + \frac{\pi}{\sqrt{2\pi N}} \int_1^{\infty} \frac{1}{\sqrt{t}(t+\pi)}\, dt \leq c\sqrt{a},$$

again with a constant c independent of a. Estimates for J_1 and J_2 establishes the continuity of I on the interval $[0, \infty)$ and finishes the proof. $\square$

6.1.2. *Dual operator of the Hilbert transform.* We next briefly comment few aspects concerning the dual space $L^p(\mathbf{R})'$ of $L^p(\mathbf{R})$, i.e. collection of all bounded linear functionals acting on $L^p(\mathbf{R})$, and their connection to the Hilbert transform. A well-known fact (see e.g. Theorem 6.16 in Rudin, Real and Complex Analysis) is that for every $\Psi \in L^p(\mathbf{R})'$ there is $g \in L^{p'}(\mathbf{R})$ such that

$$\Psi(f) = \langle g, f\rangle := \int_{\mathbf{R}} g(x)f(x)\,dx \qquad \forall f \in L^p(\mathbf{R}).$$

Therefore $L^p(\mathbf{R})'$ is isometrically isomorphic to $L^{p'}(\mathbf{R})$ whenever $1 \leq p < \infty$.

Consider now a linear convolution operator $Tf = \int_{\mathbf{R}} h(x-y)f(y)\,dy$, where $h \in L^{p'}(\mathbf{R}) \cap L^\infty(\mathbf{R})$ and $f \in L^p(\mathbf{R})$, $1 < p < \infty$. Then, by Fubini's theorem[3],

$$\Psi(Tf) = \int_{\mathbf{R}} g(x)(Tf)(x)\,dx = \int_{\mathbf{R}}\int_{\mathbf{R}} g(x)h(x-y)f(y)\,dy\,dx = \langle T'g, f\rangle,$$

where

$$(T'g)(x) = \int_{\mathbf{R}} h(y-x)g(y)\,dy.$$

On the other hand, we have by the reflexivity of $L^p(\mathbf{R})$ that

$$\|\eta\|_{p'} = \sup\{|\langle \eta, \phi\rangle| \ : \ \phi \in L^p(\mathbf{R}),\ \|\phi\|_p = 1\} \tag{6.9}$$

for all $\eta \in L^{p'}(\mathbf{R})$, readily implying that

$$\|T'g\|_{p'} \leq \sup\{\|Tf\|_p \ : \ \|f\|_p \leq 1\}\,\|g\|_{p'}. \tag{6.10}$$

The reflexivity of L^p, $1 < p < \infty$, gives that $(T')' = T$ and hence we conclude with

$$\|T'\|_{p'} = \|T\|_p,$$

where we have denoted $\|T\|_p := \sup\{\|Tf\|_p \ : \ \|f\|_p \leq 1\}$.

In fact, since $S(\mathbf{R})$ is dense in $L^p(\mathbf{R})$, we can, without losing the generality, assume above that $f, g \in S(\mathbf{R})$. Indeed, if a functional $\Psi \in L^p(\mathbf{R})'$ corresponding $g \in L^{p'}(\mathbf{R})$ via the inner product above, is bounded on a dense subset of $L^p(\mathbf{R})$, then Hahn-Banach theorem allows for an extension to the whole $L^p(\mathbf{R})$ and the norm is preserved. By the reflexivity, roles of f and g may be changed, leading to a similar conclusion. Therefore, if we are able to verify

$$\|T'\eta\|_{p'} \leq \sup\{\|T\phi\|_p \ : \ \|\phi\|_p \leq 1,\ \phi \in S(\mathbf{R})\}\,\|\eta\|_{p'}$$

for all $\eta \in S(\mathbf{R})$, also (6.10) follows in the case $1 < p < \infty$ for the extensions of T and T' from $S(\mathbf{R})$ to $L^p(\mathbf{R})$.

[3] Fubini's theorem is indeed at our disposal since $\psi_1 : (x,y) \mapsto x$, $\psi_2 : (x,y) \mapsto x-y$, and $\psi_3 : (x,y) \mapsto y$ are Borel functions and hence $g(\psi_1(x,y))$, $h(\psi_2(x,y))$, and $f(\psi_1(x,y))$ are all product measurable, as is also the product of them. Moreover, by the fact that $h \in L^\infty(\mathbf{R})$, $g(\psi_1(x,y))h(\psi_2(x,y))f(\psi_1(x,y)) \in L^1(\mathbf{R}\times\mathbf{R})$.

Next, if h above is an odd function, then clearly $T' = -T$. Let us consider now the special case

$$h(x) = K_\varepsilon(x) := \frac{\chi_{\varepsilon<|x|}}{x},$$

which clearly belongs to $L^p(\mathbf{R})$ for all $1 < p \leq \infty$ and is an odd function. Recall that K_ε is the kernel of the operator $H^{(\varepsilon)}$. By above reasoning, we thus have

$$\left\langle g, H^{(\varepsilon)} f\right\rangle = -\left\langle H^{(\varepsilon)} g, f\right\rangle$$

whenever $f, g \in S(\mathbf{R})$. Recall then the bound (6.2) from the beginning of the section implying

$$|g(H^{(\varepsilon)} f)| \leq \frac{2}{\pi}\left(\|f'\|_\infty + \|f\|_1\right)|g| \in L^1(\mathbf{R}).$$

A similar bound holds for $(H^{(\varepsilon)} g)f$ as well. Thus Lebesgue's dominated convergence implies that

$$\langle g, Hf\rangle = -\langle Hg, f\rangle$$

and hence $H' = -H$ in the class $S(\mathbf{R})$. Using this result, it is immediate that if $\|Hf\|_p \leq c_p\|f\|_p$ for $p \geq 2$ or $1 < p < 2$ and for all $f \in S(\mathbf{R})$, then $\|Hf\|_{p'} \leq c_p\|f\|_{p'}$ for all $f \in S(\mathbf{R})$.

6.1.3. *L^p-boundedness of the Hilbert Transform.* After establishing the L^2-boundedness of H and characterizing H' in $S(\mathbf{R})$, one can actually prove the L^p-boundedness in $S(\mathbf{R})$ as well. We sketch two proofs for strong (p, p). The first one is the original due to Riesz and the second one relies on Calderon-Zygmund decomposition. Both proofs go in several steps and the details of the first one are left as exercises.

Theorem 6.11. *Let $f \in S(\mathbf{R})$ and $1 < p < \infty$. Then there is a constant c_p depending only on p such that*

$$\|Hf\|_p \leq c_p\|f\|_p.$$

Moreover, weak $(1,1)$

$$|\{x \in \mathbf{R} \ : \ |Hf(x)| > \lambda\}| \leq \frac{16}{\lambda}\|f\|_1$$

holds for all $\lambda > 0$.

Proof number 1 for strong (p, p). Step 1. Let $f \in S(\mathbf{R})$. Then $\widetilde{H}f = Hf$. Taking Fourier transform of $f^2 + 2\widetilde{H}(f(\widetilde{H}f))$, one can show that

$$f^2 + 2\widetilde{H}(f(\widetilde{H}f)) = (\widetilde{H}f)^2. \tag{6.12}$$

(Recall that $\widetilde{H}f$ is defined via $(\widehat{\widetilde{H}f})(\xi) = -i\,\mathrm{sgn}(\xi)\widehat{f}(\xi)$!) To obtain (6.12) use the fact $\widehat{f^2} = \widehat{f} * \widehat{f}$ and then show

$$\begin{aligned}&\widehat{f^2}(\xi) + 2[\widetilde{H}(f(\widetilde{H}f))]^\frown(\xi)\\ &\qquad = \int_R \widehat{f}(\eta)\widehat{f}(\xi-\eta)\,[1 + m(\xi)(m(\eta)+m(\xi-\eta))]\,d\eta,\end{aligned}$$

where $m(\xi) = -i\,\mathrm{sgn}(\xi)$. By verifying the identity

$$m(\eta)m(\xi-\eta) = 1 + m(\xi)m(\eta) + m(\xi)m(\xi-\eta),$$

the claimed equality (6.12) follows using

$$\widehat{(\widetilde{H}f)^2} = \widehat{\widetilde{H}f} * \widehat{\widetilde{H}f}.$$

Step 2. Since L^p bound is true when $p = 2$, we may make an induction assumption that for $p = 2^k$, $k \in \mathbf{N}$, there is c_p such that $\|\widetilde{H}f\|_p \le c_p\|f\|_p$. Using (6.12) as

$$\|\widetilde{H}f\|_{2p}^2 = \|f^2 + 2\widetilde{H}(f(\widetilde{H}f))\|_p$$

and then the induction assumption to deduce

$$\|2\widetilde{H}(f(\widetilde{H}f))\|_p \le 2c_p\|f(\widetilde{H}f)\|_p \le 2c_p\|f\|_{2p}\|\widetilde{H}f\|_{2p},$$

one can show also the L^{2p}-boundedness with a new constant c_{2p}. Note here that $\|\widetilde{H}f\|_q = \|Hf\|_q < \infty$ for all $q \ge 2$ by (6.2).

Step 3. By using a Marcinkiewicz interpolation theorem and a duality argument, the result follows for all $1 < p < \infty$, see Step 5 in the proof number 2. □

Proof number 2. The goal is to first establish weak $(1,1)$ estimate and then interpolate using already obtained strong $(2,2)$.

Step 1. Application of Calderón-Zygmund decomposition. Suppose that $f \in S(\mathbf{R})$ and let $\lambda > 0$. Calderón-Zygmund decomposition gives us disjoint intervals $\{I_j\}$ such that

$$\begin{aligned}&|f(x)| \le \lambda \qquad \text{for a.e. } x \notin \Omega := \bigcup_j I_j,\\ &|\Omega| \le \frac{1}{\lambda}\|f\|_1,\\ &\lambda < (|f|)_j \le 2\lambda, \qquad (|f|)_j := \fint_{I_j} |f(t)|\,dt.\end{aligned}$$

Denote by c_j the center of I_j and by $2I_j$ the interval centered at c_j and with the length $2|I_j|$. Let $2\Omega = \cup_j 2I_j$.

Split f into the "good part" g and the "bad part" b as follows:

$$g(x) = \begin{cases} f(x), & x \notin \Omega,\\ (f)_j, & x \in I_j,\end{cases}$$

and

$$b = \sum_j b_j, \qquad b_j = (f - (f)_j)\, \chi_{I_j}.$$

Then $f = g + b$, $|g| \le 2\lambda$ almost everywhere and $\int_{\mathbf{R}} b_j(t)\, dt = 0$.

Step 2. Hilbert transforms of g and b. The immediate problem is that neither g or b does not belong to $S(\mathbf{R})$, in general. Nonetheless, Corollary 6.4 gives the linear extension $\widetilde{H}$ to $L^2(\mathbf{R})$ and therefore both $\widetilde{H}g$ and $\widetilde{H}b$ exists provided $g, b \in L^2(\mathbf{R})$ and

$$Hf = \widetilde{H}f = \widetilde{H}g + \widetilde{H}b,$$

because $f \in S(\mathbf{R})$. Since $|\Omega| < \infty$ and $f \in S(\mathbf{R}) \subset L^2(\mathbf{R})$, it is easy to see that $g \in L^2(\mathbf{R})$. Also $b \in L^2(\mathbf{R})$, because using the fact that $\{I_j\}$ is a disjoint collection of intervals, we obtain

$$\begin{aligned}\|b\|_2^2 =& \sum_j \int_{I_j} (f - (f)_j)^2\, dt \\ \le& 2 \sum_j \int_{I_j} (Mf)^2(t)\, dt \le 2 \int_{\mathbf{R}} (Mf)^2(t)\, dt \le c\|f\|_2^2.\end{aligned}$$

Here Mf is the Hardy-Littlewood maximal function of f and M is strong $(2, 2)$ by Theorem 2.19. Furthermore, it is easy to check that when $x \notin 2I_j$, then

$$(H^{(\varepsilon)} b_j)(x) = \int_{I_j \cap \{|x-t| > \varepsilon\}} \frac{b_j(t)}{x - t}\, dt = \int_{I_j} \frac{b_j(t)}{x - t}\, dt = (Hb_j)(x)$$

for all $\varepsilon < |I_j|$. By approximating b_j with functions from $C_0^\infty(\frac{3}{2}I_j)$, it is not hard to show that $\widetilde{H}b_j = Hb_j$ in $\mathbf{R} \setminus 2I_j$ almost everywhere. Thus the linearity of $\widetilde{H}$ in L^2 implies that

$$\widetilde{H}b = \sum_j \widetilde{H}b_j = \sum_j Hb_j$$

almost everywhere in $\mathbf{R} \setminus 2\Omega$.

Step 3. Bound for $\sum_j \int_{\mathbf{R} \setminus 2I_j} |(\widetilde{H}b_j)(t)|\, dt$. The first observation is that if $x \ne c_j$, then

$$\int_{I_j} \frac{b_j(t)}{x - c_j}\, dt = \frac{1}{x - c_j} \int_{I_j} b_j(t)\, dt = \frac{|I_j|}{x - c_j} ((f)_j - (f)_j) = 0.$$

Thus, whenever $x \notin 2I_j$,

$$(Hb_j)(x) = \int_{I_j} \frac{b_j(t)}{x - t}\, dt = \int_{I_j} \frac{b_j(t)(t - c_j)}{(x - t)(x - c_j)}\, dt.$$

Therefore,

$$\int_{\mathbf{R} \setminus 2I_j} |Hb_j(x)|\, dx \le \int_{\mathbf{R}} \int_{\mathbf{R}} \chi_{\mathbf{R} \setminus 2I_j}(x) \chi_{I_j}(t) |b_j(t)| \frac{|t - c_j|}{|x - t||x - c_j|}\, dt\, dx$$

holds. Integrand is product measurable and also in $L^1(\mathbf{R}\times\mathbf{R})$. By Fubini's theorem it follows that

$$\begin{aligned}\int_{\mathbf{R}\setminus 2I_j}|Hb_j(x)|\,dx &\le \int_{I_j}|b_j(t)|\int_{\mathbf{R}\setminus 2I_j}\frac{|t-c_j|}{|x-t||x-c_j|}\,dx\,dt\\ &\le \int_{I_j}|b_j(t)|\int_{\mathbf{R}\setminus 2I_j}\frac{\frac{1}{2}|I_j|}{\frac{1}{2}|x-c_j|^2}\,dx\,dt\\ &= \int_{I_j}|b_j(t)|\,dt\;2|I_j|\int_{|I_j|}^{\infty}s^{-2}\,ds\\ &\le 4\int_{I_j}|f(t)|\,dt.\end{aligned}$$

The estimate implies

$$\sum_j\int_{\mathbf{R}\setminus 2I_j}|\widetilde{H}b_j(x)|\,dx=\sum_j\int_{\mathbf{R}\setminus 2I_j}|Hb_j(x)|\,dx\le 4\|f\|_1.$$

Step 4. Weak (1,1) estimate. Since $Hf=\widetilde{H}f=\widetilde{H}g+\widetilde{H}b$ ($\widetilde{H}$ linear in $L^2(\mathbf{R})$), we have

$$\begin{aligned}&|\{x\in\mathbf{R}\,:\,|Hf(x)|>\lambda\}|\\ &\quad\le\left|\{x\in\mathbf{R}\,:\,|\widetilde{H}g(x)|>\lambda/2\}\right|+\left|\{x\in\mathbf{R}\,:\,|\widetilde{H}b(x)|>\lambda\}\right|.\end{aligned}$$

Corollary 6.4, together with $0\le|g|\le 2\lambda$ almost everywhere, implies

$$\begin{aligned}\left|\{x\in\mathbf{R}\,:\,|\widetilde{H}g(x)|>\lambda\}\right|&\le\frac{1}{\lambda^2}\int_{\mathbf{R}}|\widetilde{H}g(x)|^2\,dx=\frac{1}{\lambda^2}\int_{\mathbf{R}}|g|^2\,dx\\ &\le\frac{2}{\lambda}\int_{\mathbf{R}}|g|\,dx=\frac{2}{\lambda}\int_{\mathbf{R}\setminus\Omega}|f|\,dx+\frac{2}{\lambda}\sum_j|I_j||(f_j)|\le\frac{2}{\lambda}\|f\|_1.\end{aligned}$$

Next, we have

$$\left|\{x\in\mathbf{R}\,:\,|\widetilde{H}b(x)|>\lambda\}\right|\le|2\Omega|+\frac{1}{\lambda}\int_{\mathbf{R}\setminus 2\Omega}|\widetilde{H}b(x)|\,dx\le\frac{6}{\lambda}\|f\|_1$$

by Step 3 and by the fact that $|2\Omega|\le 2|\Omega|\le 2\lambda^{-1}\|f\|_1$ from the Calderón-Zygmund decomposition. Combining estimates gives

$$|\{x\in\mathbf{R}\,:\,|Hf(x)|>\lambda\}|\le\frac{16}{\lambda}\|f\|_1,$$

as asserted.

Step 5. Strong (p,p) in $S(\mathbf{R})$ via interpolation. In Step 4 we have established that H is weak $(1,1)$ in $S(\mathbf{R})$. Theorem 6.3 gives that H is also strong $(2,2)$ in $S(\mathbf{R})$. Going back to the proof of Marcinkiewicz interpolation theorem 2.21, we can actually show that H is also strong (p,p) in $S(\mathbf{R})$ for all $1<p<2$ with a constant c_p.[4] The constant,

[4] In the proof of Theorem 2.21, taking the splitting of f into f_1 and f_2, both $\widetilde{H}f_1$ and $\widetilde{H}f_2$ are well-defined, because f_1 and f_2 belong to $L^2(\mathbf{R})$ and $Hf=\widetilde{H}f_1+\widetilde{H}f_2$.

however, "blows up" as $p \uparrow 2$. We thus use the identity $H' = -H$ and obtain that, for $p = 3/2$ and $p' = 3$, H is strong $(3,3)$. Appealing again to Marcinkiewicz interpolation theorem, we get that H is strong (p,p) for all $1 < p \leq 2$ with a "stable constant" as $p \uparrow 2$. Now the result follows by the duality $H = -H'$. $\square$

Using uniform strong $(2,2)$ in Remark 6.7, it is straightforward to check that in the above proof one can replace H by $H^{(\varepsilon)}$ and obtain weak $(1,1)$ and strong (p,p) uniform in ε (in fact, the proof is much simpler in this case). Furthermore, $H^{(\varepsilon)}$ is bounded on $L^p(\mathbf{R})$, $1 < p < \infty$, and thus we have the following:

Lemma 6.13. *Let $f \in L^p(\mathbf{R})$, $1 < p < \infty$, and $\varepsilon > 0$. Then there is a constant c_p depending only on p such that*

$$\|H^{(\varepsilon)} f\|_p \leq c_p \|f\|_p.$$

Moreover, weak $(1,1)$

$$\left|\{x \in \mathbf{R} \ : \ |H^{(\varepsilon)} f(x)| > \lambda\}\right| \leq \frac{16}{\lambda}\|f\|_1$$

holds for all $\lambda > 0$.

6.1.4. *Existence of the principal value in L^p.* After establishing weak $(1,1)$ and strong (p,p), we attack the existence of the pointwise limit. For this, define the "maximal operator" $H^{(*)}$ as follows:

$$(H^{(*)})f(x) := \sup_{\varepsilon > 0} |H^{(\varepsilon)} f(x)|$$

for all $f \in L^p(\mathbf{R})$, $1 < p < \infty$. We first show a pointwise upper bound for $H^{(*)} f$ whenever $f \in S(\mathbf{R})$.

Lemma 6.14. *If $f \in S(\mathbf{R})$, then*

$$(H^{(*)} f)(x) \leq M(Hf)(x) + cMf(x)$$

for a constant c independent of f.

Proof. Let ϕ stand for the standard mollifier supported in $(-1,1)$ and $\int_{\mathbf{R}} \phi(t)\,dt = 1$. Denote $\phi_\varepsilon(t) = 2\varepsilon^{-1}\phi(2t/\varepsilon)$, which is supported in $(-\varepsilon/2, \varepsilon/2)$. We rewrite the kernel of $H^{(\varepsilon)}$ as

$$\chi_{|x|>\varepsilon}\frac{1}{x} = (H\phi_\varepsilon)(x) + \left(\chi_{|x|>\varepsilon}\frac{1}{x} - (H\phi_\varepsilon)(x)\right).$$

Let us first estimate the second term on the right. If $|x| > \varepsilon$, then

$$\begin{aligned}
\left|\chi_{|x|>\varepsilon}\frac{1}{x} - (H\phi_\varepsilon)(x)\right| &= \left|\frac{1}{x}\int_{\mathbf{R}} \phi_\varepsilon(t)\,dt - (H\phi_\varepsilon)(x)\right| \\
&= \left|\int_{|t|<\varepsilon/2} \phi_\varepsilon(t)\left(\frac{1}{x} - \frac{1}{x-t}\right)\,dt\right| \leq \int_{|t|<\varepsilon/2} \phi_\varepsilon(t)\frac{|t|}{|x||x-t|}\,dt \\
&\leq \frac{\varepsilon}{|x|^2}.
\end{aligned}$$

On the other hand, if $|x| < \varepsilon$, then

$$\left|\chi_{|x|>\varepsilon}\frac{1}{x} - (H\phi_\varepsilon)(x)\right| = |(H\phi_\varepsilon)(x)| = \left|\int_{|t|<\varepsilon/2} \frac{\phi_\varepsilon(t)-\phi_\varepsilon(x)}{x-t}\,dt\right|.$$

Denoting $\tilde{x} = 2x/\varepsilon$ and making the change of variable $\tilde{t} = 2t/\varepsilon$, we have

$$\left|\int_{|t|<\varepsilon/2} \frac{\phi_\varepsilon(t)-\phi_\varepsilon(x)}{x-t}\,dt\right| \le \frac{2}{\varepsilon}\sup_{-1<\tilde{x},\tilde{t}<1}\left|\frac{\phi(\tilde{x})-\phi(\tilde{t})}{\tilde{x}-\tilde{t}}\right| \le \frac{c}{\varepsilon}$$

and thus

$$\left|\chi_{|x|>\varepsilon}\frac{1}{x} - (H\phi_\varepsilon)(x)\right| \le \frac{c\varepsilon}{\varepsilon^2+|x|^2} =: \Psi_\varepsilon(x).$$

follows. This readily implies that

$$\left|\int_{\mathbf{R}} \left(\chi_{|t|>\varepsilon}\frac{1}{t} - (H\phi_\varepsilon)(t)\right) f(x-t)\,dt\right| \le (\Psi_\varepsilon * |f|)(x)$$

Since $\int_{\mathbf{R}} \Psi_\varepsilon(x)\,dx = \pi c$ and Ψ is a radially decreasing function, we obtain by Theorem 3.10 that

$$\left|\int_{\mathbf{R}} \left(\chi_{|t|>\varepsilon}\frac{1}{t} - (H\phi_\varepsilon)(t)\right) f(x-t)\,dt\right| \le (\Psi_\varepsilon * |f|)(x) \le cMf(x)$$

for a new constant c independent of ε.

Next, we have by Fubini's theorem[5] that

$$((H\phi_\varepsilon) * f)(x) = \int_{\mathbf{R}}\int_{\mathbf{R}} \frac{\phi_\varepsilon(t)}{x-y-t}\,dt f(y)\,dy = (\phi_\varepsilon * (Hf))(x)$$

and again by Theorem 3.10 (when $n = 1$, then $C(n,\phi) = \|\phi\|_1$),

$$|(\phi_\varepsilon * (Hf))(x)| \le M(Hf)(x)$$

follows. In conclusion, we have

$$|H^{(\varepsilon)}f| \le M(Hf)(x) + cMf(x)$$

for a constant c independent of ε. This finishes the proof. □

Corollary 6.15. *The maximal operator $H^{(*)}$ is strong (p,p) in $S(\mathbf{R})$ for $1 < p < \infty$, i.e. there is a constant c_p such that*

$$\|H^{(*)}f\|_p \le c_p\|f\|_p$$

for all $f \in S(\mathbf{R})$.

Proof. Both H and M are strong (p,p) and thus, by the previous lemma,

$$\|H^{(*)}f\|_p \le \|M(Hf) + cMf\|_p \le c_1\|Hf\|_p + c_1\|f\|_p \le c_2\|f\|_p.$$

□

[5]the function $\phi_\varepsilon(t)f(y)/(x-y-t)$ is product measurable and in $L^1(\mathbf{R}\times\mathbf{R})$ since ϕ_ε and f are in $S(\mathbf{R})$

To characterize the pointwise limit, we yet need another result concerning families of linear operators.

Theorem 6.16. *Suppose that $\{T_t\}_{t>0}$ is a family of linear operators. If the maximal operator*

$$(T^*f)(x) := \sup_{t>0} |(T_t f)(x)|$$

is weak (p,q), i.e. there is a constant c such that

$$|\{x \in \mathbf{R} \ : (T^*f)(x) > \lambda\}| \leq c\left(\frac{\|f\|_p}{\lambda}\right)^q$$

for some $p, q > 0$, then the set

$$\Psi := \left\{ f \in L^p(\mathbf{R}) \ : \ \lim_{t\downarrow 0} (T_t f)(x) \text{ exists for a.e. } x \in \mathbf{R} \right\}$$

is closed in $L^p(\mathbf{R})$.

Proof. If Ψ is empty, then it is trivially closed. We thus assume that it is nonempty. Let then $(f_j)_{j\in\mathbf{N}} \subset \Psi$ be a sequence in $L^p(\mathbf{R})$ converging to f in $L^p(\mathbf{R})$. We will show that $f \in \Psi$. Denote $g_j = \lim_{t\downarrow 0} T_t f_j$, which exists for almost every $x \in \mathbf{R}$ since $f_j \ni \Psi$. By the linearity of T_t, we have

$$\begin{aligned}
\limsup_{t\downarrow 0} (T_t f)(x) &= \limsup_{t\downarrow 0} (T_t(f - f_j) + T_t f_j)(x) \\
&\leq \limsup_{t\downarrow 0} (T_t(f - f_j))(x) + g_j(x)
\end{aligned}$$

and similarly

$$\liminf_{t\downarrow 0} (T_t f)(x) \geq \liminf_{t\downarrow 0} (T_t(f - f_j))(x) + g_j(x)$$

Therefore

$$\begin{aligned}
&\limsup_{t\downarrow 0} (T_t f)(x) - \liminf_{t\downarrow 0} (T_t f)(x) \\
&\quad \leq \limsup_{t\downarrow 0} (T_t(f - f_j))(x) - \liminf_{t\downarrow 0} (T_t(f - f_j))(x) \\
&\quad \leq 2(T^*(f - f_j))(x)
\end{aligned}$$

holds, readily implying

$$\begin{aligned}
&\left|\left\{ x \in \mathbf{R} \ : \ \limsup_{t\downarrow 0} (T_t f)(x) - \liminf_{t\downarrow 0} (T_t f)(x) > \lambda \right\}\right| \\
&\quad \leq |\{x \in \mathbf{R} \ : \ 2(T^*(f - f_j))(x) > \lambda\}|
\end{aligned}$$

for all $\lambda > 0$. Since T^* is weak (p,q), we obtain

$$|\{x \in \mathbf{R} \ : \ 2(T^*(f - f_j))(x) > \lambda\}| \leq \left(2c\frac{\|f - f_j\|_p}{\lambda}\right)^q \to 0$$

as $j \to \infty$. It follows that

$$\left|\left\{x \in \mathbf{R} \;:\; \limsup_{t\downarrow 0}(T_t f)(x) - \liminf_{t\downarrow 0}(T_t f)(x) > 0\right\}\right| = 0$$

and thus $f \in \Psi$, showing that Ψ is sequently closed in $L^p(\mathbf{R})$, and therefore also topologically closed. $\square$

We are now ready to prove the existence of the pointwise limit.

Theorem 6.17. *Let $f \in L^p(\mathbf{R})$, $1 < p < \infty$. Then the limit*

$$(Hf)(x) = \lim_{\varepsilon\downarrow 0}(H^{(\varepsilon)} f)(x)$$

exists for almost every $x \in \mathbf{R}$ and there is a constant c_p such that

$$\|Hf\|_p \le c_p\|f\|_p.$$

Proof. By the previous theorem,

$$\Psi := \left\{f \in L^p(\mathbf{R}) \;:\; \lim_{\varepsilon\downarrow 0}(H^{(\varepsilon)} f)(x) \text{ exists for a.e. } x \in \mathbf{R}\right\}$$

is closed in $L^p(\mathbf{R})$. Let $(\phi_j)_{j\in\mathbf{N}} \subset S(\mathbf{R})$ be a sequence converging to f in $L^p(\mathbf{R})$. Since $H\phi_j = \lim_{\varepsilon\downarrow 0} H^{(\varepsilon)}\phi_j$ for every $j \in \mathbf{N}$, $(\phi_j) \subset \Psi$. But now Ψ is closed in $L^p(\mathbf{R})$ and consequently $f \in \Psi$. Thus the limit in the statement exists.

To show the strong (p, p), infer first by Theorem 6.11 that

$$\|H(\phi_j - \phi_k)\| \le c_p\|\phi_j - \phi_k\|_p.$$

It follows that $(H\phi_j)_{j\in\mathbf{N}}$ is a Cauchy sequence in $L^p(\mathbf{R})$ and, by the completeness, there is $g \in L^p(\mathbf{R})$ such that $\|H\phi_j - g\|_p \to 0$ as $j \to \infty$. We proceed in showing that $Hf = g$ almost everywhere.

By Fatou's lemma, we have

$$\begin{aligned}\|Hf - g\|_p &= \left(\int_{\mathbf{R}} \liminf_{\varepsilon\downarrow 0} |(H^{(\varepsilon)} f)(t) - g(t)|^p\, dt\right)^{1/p} \\ &\le \liminf_{\varepsilon\downarrow 0} \|H^{(\varepsilon)} f - g\|_p.\end{aligned}$$

Fix any $\delta > 0$ and take $j \in \mathbf{N}$ to be so large that

$$\|H^{(\varepsilon)}(\phi_j - f)\|_p + \|H\phi_j - g\|_p < \delta$$

uniformly in ε. This is possible by Lemma 6.13 and since both $\|\phi_j - f\|_p$ and $\|H\phi_j - g\|_p \to 0$ as $j \to \infty$. It follows that

$$\begin{aligned}\|H^{(\varepsilon)} f - g\|_p &\le \|H^{(\varepsilon)}(\phi_j - f)\|_p + \|H\phi_j - H^{(\varepsilon)}\phi_j\|_p + \|H\phi_j - g\|_p \\ &\le \delta + \|H\phi_j - H^{(\varepsilon)}\phi_j\|_p\end{aligned}$$

and therefore

$$\|Hf - g\|_p \le \liminf_{\varepsilon\downarrow 0} \|H^{(\varepsilon)} f - g\|_p < \delta.$$

This holds for all $\delta > 0$ and thus $\|Hf - g\|_p = 0$. Finally, since

$$\|g\|_p = \lim_{j\to\infty} \|H\phi_j\|_p \le c_p \lim_{j\to\infty} \|\phi_j\|_p = c_p\|f\|_p$$

by Theorem 6.11, we obtain that H is strong (p, p). □

We finally proceed in showing that the limit exists also when $f \in L^1(\mathbf{R})$. The details are left as exercises, but we outline the proof.

Lemma 6.18. *The maximal operator $H^{(*)}$ is weak $(1, 1)$ in $S(\mathbf{R})$.*

Proof. The proof follows of the second proof of Theorem 6.11.

Step 1. Calderón-Zygmund decomposition. Form the Calderón-Zygmund decomposition as in the proof of Theorem 6.11. Then reason that if both $H^{(*)}g$ and $H^{(*)}b$ satisfy weak $(1, 1)$, also $H^{(*)}f$ satisfy the same.

Step 2. Weak (1,1) for $H^{(*)}g$. Use the fact that $H^{(*)}$ is strong $(2, 2)$ to obtain the weak $(1, 1)$ bound for $H^{(*)}g$.

Step 3. Weak (1,1) for $H^{(*)}b$. In order to show that

$$\left|\{x \notin 2\Omega \ : \ (H^{(*)}b)(x) > \lambda\}\right| \le \frac{c}{\lambda}\|b_1\|$$

for some constant $c > 0$, fix $x \notin 2\Omega$, $\varepsilon > 0$, and b_j with the support I_j. Treat separately cases

$$\begin{aligned} &(1) \qquad (x-\varepsilon, x+\varepsilon) \cap I_j = I_j,\\ &(2) \qquad (x-\varepsilon, x+\varepsilon) \cap I_j = \emptyset,\\ &(3) \qquad x-\varepsilon \in I_j \text{ or } x+\varepsilon \in I_j. \end{aligned}$$

In the first case show that $H^{(\varepsilon)}b_j(x) = 0$. In the second apply the fact $(H^{(\varepsilon)}b_j)(x) = (Hb_j)(x)$ and then show that

$$|(H^{(\varepsilon)}b_j)(x)| \le \frac{|2I_j|}{|x - c_j|^2}\|b_j\|_1.$$

In the third case third prove that

$$|(H^{(\varepsilon)}b_j)(x)| \le \frac{3}{\varepsilon}\int_{x-3\varepsilon}^{x+3\varepsilon} |b_j(t)|\,dt \le (Mb_j)(x) \le (Mb)(x).$$

From these facts weak (1,1) follows easily. □

We are ready to prove:

Theorem 6.19. *Let $f \in L^1(\mathbf{R})$. Then the limit*

$$(Hf)(x) = \lim_{\varepsilon\downarrow 0}(H^{(\varepsilon)}f)(x)$$

exists for almost every $x \in \mathbf{R}$ and there is a constant c such that

$$|\{x \in \mathbf{R} \ : \ |(Hf)(x)| > \lambda\}| \le \frac{c}{\lambda}\|f\|_1$$

for all $\lambda > 0$.

Proof. Using the above lemma, reproduce the proof of Theorem 6.17 in $L^1(\mathbf{R})$. □

6.1.5. *Connection to analytic functions.* Hilbert transform appears naturally in the theory of analytic functions. Suppose that $f \in L^p(\mathbf{R})$, $1 \le p < \infty$, and let

$$F(z) = \frac{1}{i\pi} \int_{\mathbf{R}} \frac{f(t)}{t-z}\, dt,$$

where $z = x + iy \in \mathbf{R}^2_+$. The first observation is that when $y > 0$, then Hölder's inequality yields

$$\begin{aligned}\int_{\mathbf{R}} \left|\frac{f(t)}{t-z}\right| dt \le & \left(\int_{\mathbf{R}} |f(t)|^p\, dt\right)^{1/p} \left(\int_{\mathbf{R}} |t-z|^{-p'}\, dt\right)^{1/p'} \\ \le & \|f\|_p \left(\int_{\mathbf{R}} \left(y^2 + (t-x)^2\right)^{-p'/2}\, dt\right)^{1/p'} < \|f\|_p C(y,p),\end{aligned}$$

because $p' = p/(p-1) > 1$. Next, the function

$$F_N(z) := \frac{1}{i\pi} \int_{-N}^{N} \frac{f(t)}{t-z}\, dt,$$

is analytic in $\mathbf{R}^2_+$, because

$$\begin{aligned}\frac{F_N(z+h) - F_N(z)}{h} &= \frac{1}{i\pi h} \int_{-N}^{N} \left(\frac{f(t)}{t-z-h} - \frac{f(t)}{t-z}\right) dt \\ &= \frac{1}{i\pi} \int_{-N}^{N} \frac{f(t)}{(t-z-h)(t-z)}\, dt \to \frac{1}{i\pi} \int_{-N}^{N} \frac{f(t)}{(t-z)^2}\, dt\end{aligned}$$

as $h \to 0$ and the integral on the right exists for all $z \in \mathbf{R}^2_+$. Moreover, for any fixed $y_0 > 0$ and $y \ge y_0$ we have

$$F_N(z) \to F(z)$$

uniformly. Therefore also $F(z)$ is analytic in $\mathbf{R}^2_+$. Decomposing

$$\frac{1}{t-z} = \frac{1}{t-x-iy} = \frac{t-x}{(t-x)^2+y^2} + i\frac{y}{(t-x)^2+y^2}$$

we arrive at

$$\begin{aligned}F(z) =& \frac{1}{\pi} \int_{\mathbf{R}} f(t) \frac{y}{(t-x)^2+y^2}\, dt + \frac{i}{\pi} \int_{\mathbf{R}} f(t) \frac{x-t}{(t-x)^2+y^2}\, dt \\ =:& (f * P_y)(x) + i(f * Q_y)(x).\end{aligned}$$

Here P_y is the Poisson kernel, see Example 3.18, and

$$Q_y(x) := \frac{1}{\pi} \frac{x}{x^2+y^2}$$

so-called conjugate Poisson kernel. The function f is real-valued and hence $f * P_y$ is the real and $f * Q_y$ is the imaginary part of F.

Theorem 3.12 implies that $f * P_y \to f$ almost everywhere as $y \downarrow 0$. (The convergence is even nontangential, see Definition 3.19, by

Corollary 3.21.) But what about the convergence of $f * Q_y$? We would like to infer

$$\lim_{y\downarrow 0}(f * Q_y)(x) = \lim_{y\downarrow 0}\frac{1}{\pi}\int_{\mathbf{R}} f(t)\frac{x-t}{(t-x)^2+y^2}\,dt.$$

If such a limit exists, this would readily imply that also the limit $\lim_{y\downarrow 0} F(x+iy)$ exists. The next theorem says that the limit exists and it is precisely the Hilbert transform of f.

Theorem 6.20. *Let $f \in L^p(\mathbf{R})$, $1 \le p < \infty$. Then*

$$\lim_{y\downarrow 0}(f * Q_y)(x) = (Hf)(x)$$

for almost every $x \in \mathbf{R}$.

Proof. Let us write the difference of the kernels of $f * Q_y$ and $H^{(y)}f$:

$$\begin{aligned} Q_y(t) - K^{(y)}(t) =& \frac{t}{t^2+y^2} - \chi_{|t|>y}\frac{1}{t} \\ =& \frac{1}{t(t^2+y^2)}\left(t^2 - \chi_{|t|>y}(t^2+y^2)\right) \\ =& \frac{t}{t^2+y^2}\chi_{|t|\le y} - \frac{y^2}{t(t^2+y^2)}\chi_{|t|>y} =: \phi_y(t). \end{aligned}$$

Denote

$$\phi(t) = y\phi_y(yt) = \frac{t}{t^2+1}\chi_{|t|\le 1} - \frac{1}{t(t^2+1)}\chi_{|t|>1}.$$

Since ϕ is an odd function, we have

$$\int_{\mathbf{R}} \phi(t)\,dt = 0.$$

Moreover, it is not hard to see that ϕ has a radially decreasing integrable majorant

$$\Psi(t) = \begin{cases} \dfrac{1}{2}, & |t| \le 1, \\ \dfrac{1}{|t|(t^2+1)}, & |t| > 1. \end{cases}$$

Then Theorem 3.12 (see also the preceding remark) implies that

$$\phi_y * f \to 0$$

almost everywhere as $y \to 0$. But since $(H^{(y)}f)(x) \to (Hf)(x)$ for almost every $x \in \mathbf{R}$, we obtain that $(f * Q_y)(x) \to (Hf)(x)$ for almost every $x \in \mathbf{R}$, concluding the proof. $\square$

As a corollary F defined above attains "boundary values" as $y \downarrow 0$.

Corollary 6.21. *Let $f \in L^p(\mathbf{R})$, $1 \le p < \infty$. Then*

$$F(x+iy) = \frac{1}{i\pi}\int_{\mathbf{R}} \frac{f(t)}{t-x-iy}\,dt \to f(x) + i(Hf)(x)$$

as $y \downarrow 0$ for almost every $x \in \mathbf{R}$.

6.2. **Riesz transform.** We continue our studies of singular integral operators with the Riesz transform, which is a natural generalization of the Hilbert transform to higher dimensions. It is defined in sense of principal values, i.e. as

$$(R_j f)(x) := \lim_{\varepsilon \downarrow 0} (R_j^{(\varepsilon)} f) = \lim_{\varepsilon \downarrow 0} c_n \int_{|y|>\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)\, dy,$$

whenever the limit exists. Here

$$c_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{(n+1)/2}} = \frac{2}{\omega_n}.$$

In the case of the Hilbert transform, we gave a rather complete treatment of the L^p-theory. For the Riesz transform we will be somewhat informal and, for example, will prove strong (p,p) later in context of more general convolution type operators. Our point of view here will be in the study of the connection between the Riesz transform and the theory of harmonic functions and Fourier analysis.

As in the case of the Hilbert transform, we first give a simple condition to guarantee the existence of the limit.

Lemma 6.22. *Let $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$. Suppose that at the point $x \in \mathbf{R}^n$,*

$$|f(y) - f(x)| \le C|x-y|^\alpha$$

for all $y \in B(x,\delta)$, for some $C > 0$ and $\alpha, \delta > 0$. Then the limit $(R_j f)(x)$ exists. Moreover, under the same condition on f,

$$\lim_{t \downarrow 0} (Q_{j,t} * f)(x) = (R_j f)(x), \qquad Q_{j,t}(y) := \frac{y_j}{(t^2 + |y|^2)^{(n+1)/2}}.$$

Proof. Exercise. □

In particular, if $f \in S(\mathbf{R}^n)$, limits above exist for all $x \in \mathbf{R}^n$. Furthermore, in $S(\mathbf{R}^n)$, there is a powerful pointwise bound for the Riesz transform, which belongs to $L^p(\mathbf{R}^n)$ for all $p > 1$:

Lemma 6.23. *For all $f \in S(\mathbf{R}^n)$ there is a constant c depending on f and n such that*

$$\sup_{0<\varepsilon<1/2} \left| \int_{|y|>\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)\, dy \right| \le \frac{c}{(1+|x|)^n}.$$

Proof. Let $0 < \varepsilon < 1/2$. Rewrite the integral as

$$\begin{aligned} &\int_{|y|>\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)\, dy \\ &\quad = \int_{1/2>|y|>\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)\, dy + \int_{|y|\ge 1/2} \frac{y_j}{|y|^{n+1}} f(x-y)\, dy \\ &\quad =: I_1(x) + I_2(x). \end{aligned}$$

Since the kernel $\chi_{1/2>|y|\geq\varepsilon} y_j/|y|^{n+1}$ is an odd function, we have that

$$I_1(x) = \int_{1/2>|y|\geq\varepsilon} \frac{y_j}{|y|^{n+1}}(f(x-y)-f(x))\,dy.$$

By defining $g(s) = f(x+s(z-x))$, it is not hard to prove the estimate

$$\frac{|f(z)-f(x)|}{|z-x|} \leq 2^n(1+|x|)^{-n} \sup_{y\in\mathbf{R}^n}(1+|y|)^n|\nabla f(y)|$$

provided that $|z-x| \leq |x|/2$. Therefore,

$$|I_1(x)| \leq 2^n(1+|x|)^{-n} \sup_{y\in\mathbf{R}^n}(1+|y|)^n|\nabla f(y)| \int_{1/2>|y|\geq\varepsilon} |y|^{1-n}\,dy$$

follows. The integral is bounded uniformly in t and ε and thus

$$|I_1(x)| \leq \frac{c}{(1+|x|)^n}.$$

Next, if $|x| \leq 1$, then

$$|I_2(x)| \leq 2^n \int_{\mathbf{R}^n} |f(x-y)|\,dy = 2^n\|f\|_1 \leq \frac{c}{(1+|x|)^n}.$$

If $|x| > 1$, split the integral into two parts:

$$\begin{aligned} I_2(x) = &\int_{1/2\leq|y|\leq|x|/2} \frac{y_j}{|y|^{n+1}} f(x-y)\,dy \\ &+ \int_{|y|>|x|/2} \frac{y_j}{|y|^{n+1}} f(x-y)\,dy =: I_{2,1}(x) + I_{2,2}(x). \end{aligned}$$

The last term we estimate simply as

$$|I_{2,2}(x)| \leq 2^n|x|^{-n} \int_{|y|>|x|/2} |f(x-y)|\,dy \leq 2^n|x|^{-n}\|f\|_1 \leq \frac{c}{(1+|x|)^n}.$$

For the first term, observe that the condition $|y| \leq |x|/2$ implies $|x|/2 \leq |x|-|y| \leq |x-y|$ and therefore

$$\begin{aligned} |I_{2,1}(x)| \leq &2^n \int_{1/2\leq|y|\leq|x|/2} |x-y|^{-2n}|x-y|^{2n}|f(x-y)|\,dy \\ \leq &4^n\omega_n|x|^{-n} \sup_{z\in\mathbf{R}^n} |z|^{2n}|f(z)| \leq \frac{c}{(1+|x|)^n}. \end{aligned}$$

The result follows. $\square$

A similar argument provides us information about convolutions with Poisson and conjugate Poisson kernels

$$P_t(y) := c_n \frac{t}{(t^2+|y|^2)^{(n+1)/2}}, \qquad Q_{k,t}(y) := c_n \frac{y_k}{(t^2+|y|^2)^{(n+1)/2}},$$

respectively.

Lemma 6.24. *For all $f \in S(\mathbf{R}^n)$ there is a constant c depending on f and n such that*

$$|P_t * f(x)| \le \frac{c(1+t)}{(1+t^2+|x|^2)^{(n+1)/2}}$$

and

$$|Q_{k,t} * f(x)| \le \frac{c}{(1+t^2+|x|^2)^{n/2}}$$

Proof. The details are left as an exercise. □

Combining results of Lemmata 6.22–6.25 gives us the following:

Lemma 6.25. *Let $f \in S(\mathbf{R}^n)$. Then*

$$\lim_{t\downarrow 0} \|Q_{j,t} * f - R_j f\|_p = 0, \qquad \lim_{t\downarrow 0} \|P_t * f - f\|_p = 0,$$

for all $p > 1$.

Proof. In view of the Theorem 3.12,

$$\lim_{t\downarrow 0}(P_t * f - f) = 0$$

pointwise (almost) everywhere and, on the other hand, Lemma 6.22 gives

$$\lim_{t\downarrow 0}(Q_{j,t} * f - R_j f) = 0.$$

But now by Lemmata 6.31 and 6.25.

$$|(P_t * f - f)(x)| \le |f| + \frac{c}{(1+|x|^2)^{(n+1)/2}}$$

and

$$|(Q_{j,t} * f - R_j f)(x)| \le \frac{c}{(1+|x|^2)^{n/2}}\,.$$

for all $0 < t < 1$ with a constant c depending only on f and n, we obtain the desired result by the dominated convergence. □

6.2.1. *Fourier transform of the Riesz transform.* Following the path we took in the analysis of the Hilbert transform, we calculate the Fourier transform of the Riesz transform of Schwartz functions. Although we have defined Fourier transform only in $\mathbf{R}$, an analogous theory holds in $\mathbf{R}^n$. Fourier transform of $f \in S(\mathbf{R}^n)$ is defined as

$$\widehat{f}(\xi) = \int_{\mathbf{R}^n} f(x) e^{-i2\pi x\cdot\xi}\, d\xi.$$

It has the following basic properties.

Lemma 6.26. *Suppose that $f \in S(\mathbf{R}^n)$. Then*

(i) $\widehat{(\alpha f + \beta g)} = \alpha\hat{f} + \beta\hat{g}$,

(ii) $\widehat{\left(\frac{\partial f}{\partial x_j}\right)}(\xi) = 2\pi i \xi_j \hat{f}(\xi)$, $j = 1, \dots, n$,

(iii) $\frac{\partial \hat{f}}{\partial \xi_j}(\xi) = (\widehat{-2\pi i x_j f})(\xi)$, $j = 1, \dots, n$,

(iv) $\hat{f}$ *is continuous,*
(v) $||\hat{f}||_\infty \leq ||f||_1$,
(vi) $\widehat{f(\varepsilon x)} = \frac{1}{\varepsilon}\hat{f}(\frac{\xi}{\varepsilon}) = \hat{f}_\varepsilon(\xi), \varepsilon > 0$,
(vii) $\widehat{fg} = \hat{f} * \hat{g}$,
(viii) $\widehat{f * g} = \hat{f}\hat{g}$.

The proof is similar to the one-dimensional case and we leave it as an exercise.

We now *take from the literature* the Fourier transform of the Poisson kernel. The proof is somewhat hideous, but standard.

Lemma 6.27.

$$\widehat{P}(\xi, t) = e^{-2\pi t|\xi|}.$$

The preceding lemma lets us to calculate

Theorem 6.28. *Let* $f \in S(\mathbf{R}^n)$*. Then*

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

for $j = 1, \ldots, n$ *and for all* $\xi \in \mathbf{R}^n \setminus \{0\}$.

Proof. We will pursue the fact from the previous section that

$$f * Q_{j,t} \to R_j f.$$

in $L^2(\mathbf{R}^n)$ as $t \downarrow 0$. Since

$$Q_{j,t}(x) = \frac{x_j}{t}P_t(x),$$

we have

$$\widehat{Q}_{j,t}(\xi) = (\widehat{\frac{x_j}{t}P_t(x)}) = \frac{1}{i2\pi t}(\widehat{i2\pi x_j P_t}(x)) = \frac{1}{i2\pi t}\left(-\frac{\partial}{\partial \xi_j}\widehat{P}_t(\xi)\right).$$

Lemma 6.27 implies that

$$\frac{\partial}{\partial \xi_j}\widehat{P}_t(\xi) = -2\pi t\frac{\xi_j}{|\xi|}\widehat{P}_t(\xi),$$

an therefore

$$\widehat{Q}_{j,t}(\xi) = -i\frac{\xi_j}{|\xi|}\widehat{P}_t(\xi).$$

Then

$$\widehat{Q}_{j,t}(\xi)\widehat{f}(\xi) + i\frac{\xi_j}{|\xi|}\widehat{f}(\xi) = -i\frac{\xi_j}{|\xi|}\left(\widehat{P}_t(\xi)\widehat{f}(\xi) - \widehat{f}(\xi)\right)$$

and, in particular, Plancerel's formula implies

$$\| -i\xi_j|\xi|^{-1}\widehat{f}(\xi)(\widehat{P}_t(\xi) - 1)\|_2 \leq \|\widehat{P}_t(\xi)\widehat{f}(\xi) - \widehat{f}(\xi)\|_2 = \|P_t * f - f\|_2.$$

Using two above formulaes, together with the triangle inequality and again Plancerel's formula, leads to

$$\begin{aligned}
&\|\widehat{R_j f} + i\xi_j|\xi|^{-1}\widehat{f}(\xi)\|_2 \\
&\qquad \le \|\widehat{Q}_{j,t}(\xi)\widehat{f}(\xi) + i\xi_j|\xi|^{-1}\widehat{f}(\xi)\|_2 + \|\widehat{Q}_{j,t}(\xi)\widehat{f}(\xi) - \widehat{R_j f}\|_2 \\
&\qquad \le \|P_t * f - f\|_2 + \|Q_{j,t} * f - R_j f\|_2
\end{aligned}$$

and letting $t \downarrow 0$, we obtain by Lemma 6.25 that the two norms on the right tend to zero and consequently,

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

concluding the proof. □

As in the case of the Hilbert transform, we get an extension for the Riesz transform to whole $L^2(\mathbf{R}^n)$.

Corollary 6.29. *The Riesz transform R_j allows for an extension $\widetilde{R}_j$ to $L^2(\mathbf{R})$ via the Fourier transform*

$$\widehat{(\widetilde{R}_j f)}(\xi) = -i\frac{\xi_j}{|\xi|}\widehat{f}(\xi), \qquad \xi \in \mathbf{R}^n \setminus \{0\},$$

for all $f \in L^2(\mathbf{R}^n)$.

Finally, using the Fourier transformation of the Riesz transform, one obtains one of the cornerstones of Harmonic Analysis.

Theorem 6.30. *Let $f \in S(\mathbf{R}^n)$. Then*

$$\left\| \frac{\partial^2 f}{\partial x_i \partial x_j} \right\|_2 \le \|\Delta f\|_2.$$

6.2.2. *Conjugate harmonic functions in $\mathbf{R}^{n+1}_+$.* Our first task is to find the connection between the Riesz transform and the Poisson integral. For this, let u be a harmonic function in $\mathbf{R}^{n+1}_+$. Define

$$U(x) = (u_1(x), \dots, u_{n+1}(x)) = \nabla u(x) = \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_{n+1}} \right).$$

Each component of U is a harmonic function and they satisfy a *generalized Cauchy-Riemann system*

$$\begin{cases} \dfrac{\partial u_k}{\partial x_j} = \dfrac{\partial u_j}{\partial x_k}, \quad 1 \le j, k \le n+1, \\ \displaystyle\sum_{j=1}^{n+1} \frac{\partial u_j}{\partial x_j} = 0. \end{cases}$$

In other words, when defining the higher dimensional curl of a vector field V via

$$\operatorname{curl}(V)_{jk} = \frac{\partial V_k}{\partial x_j} - \frac{\partial V_j}{\partial x_k},$$

we have

$$\operatorname{curl}(U) = 0, \qquad \operatorname{div}(U) = 0, \qquad \text{in } \mathbf{R}^{n+1}_+.$$

Conversely, if $\operatorname{curl}(U) = 0$ in $\mathbf{R}^{n+1}_+$, then there is a potential H such that $U = \nabla H$. The existence of such a potential follows by the generalized Stokes theorem and the fact that $\mathbf{R}^{n+1}_+$ is simply connected. If, in addition, $\operatorname{div}(U) = 0$ in $\mathbf{R}^{n+1}_+$, then $\Delta H = 0$ in $\mathbf{R}^{n+1}_+$. Consequently U is the gradient of a harmonic function H, implying, in particular, that each component of U is harmonic as well.

Let now $f \in S(\mathbf{R}^n)$ and define the Poisson integral of f as

$$u(x,t) := (f * P_t)(x), \qquad P_t(x) = c_n \frac{t}{(t^2 + |x|^2)^{(n+1)/2}},$$

where $c_n = 2/\omega_n$. Then u is a harmonic function, see Example 3.18. A natural question is that whether u is a component of a system of conjugate functions in $\mathbf{R}^{n+1}_+$, i.e. is there a vector field $U = (u_1, \dots, u_n, u)$ satisfying the generalized Cauchy-Riemann system? Note first that by defining

$$H(x,t) := \frac{c_n}{1-n} \int_{\mathbf{R}^n} f(y) \frac{1}{(t^2 + |x-y|^2)^{(n-1)/2}} \, dy,$$

we have

$$u(x,t) = \frac{\partial H}{\partial t}(x,t).$$

Here we may indeed take the derivative inside the integral since $f \in S(\mathbf{R}^n)$. Define

$$u_k(x,t) := \frac{\partial H}{\partial x_k}(x,t), \qquad k = 1, \dots, n.$$

Differentiating under the integral gives

$$u_k(x,t) = (f * Q_{k,t})(x,t), \qquad k = 1, \dots, n,$$

$Q_{k,t}$ being the conjugate Poisson kernel. It is easy to verify that $U = (u_1, \dots, u_n, u)$ satisfies the generalized Cauchy-Riemann system. But what happens when $t \downarrow 0$? As we showed before,

$$\lim_{t \downarrow 0} u_k(x,t) = \lim_{t \downarrow 0} (f * Q_k)(x,t) = (R_k f)(x)$$

whenever $f \in S(\mathbf{R}^n)$. Thus Riesz transforms are "boundary values" of particular solutions to the generalized Cauchy-Riemann system.

6.3. **Singular integrals of convolution type.** We will now set our sails towards more general theory of integral operators. We confine ourselves to study only integral operators of convolution type, while different methods would provide results for even more general type of kernels. Nonetheless, many arguments we are using here work as such in the more general case.

Define the integral operator T of convolution type as

$$(Tf)(x) = \lim_{\varepsilon\downarrow}(T^{(\varepsilon)}f)(x) = \lim_{\varepsilon\downarrow}\int_{|y|>\varepsilon}\frac{\Omega(y)}{|y|^n}f(x-y)\,dy,$$

where the kernel function Ω has the following properties:

(i) Ω is homogeneous, i.e. $\Omega(rx) = \Omega(x)$ for all $r > 0$,

(ii) Ω is bounded, i.e. $|\Omega(x)| \le A_1$, $A_1 \ge 0$, for all $x \in S^{n-1} := \{x \in \mathbf{R}^n \ : \ |x| = 1\}$,

(iii) Ω is Dini continuous, that is, the modulus of continuity

$$\omega(r) := \sup_{x,y\in S^{n-1},|x-y|\le r}|\Omega(x)-\Omega(y)|$$

satisfies

$$\int_0^1 \omega(r)\,\frac{dr}{r} \le A_2, \qquad A_2 \ge 0,$$

(iv) The average of Ω on S^{n-1} is zero, i.e.

$$\int_{S^{n-1}}\Omega(x)\,d\mathcal{H}^{n-1} = 0.$$

Here $\mathcal{H}^{n-1}$ is the codimension one Hausdorff-measure or, in other words, the surface measure on S^{n-1}.

In general, the last condition is necessary for the principal value to exist. We leave this as an exercise. Moreover, it will allow us to calculate the Fourier transform of $T^{(\varepsilon)}f$ and Tf.

The structure of this section is suggested by the section considering the Hilbert transform. Indeed, we first calculate the Fourier transform of $T^{(\varepsilon)}f$ and Tf and using them, we deduce strong $(2,2)$ in $S(\mathbf{R}^n)$. This, in particular, allows us to extend operators in the whole $L^2(\mathbf{R}^n)$. We then show strong (p,p) using rather modern approach, which in this context is nowadays often cited as *Nonlinear Calderón-Zygmund theory.* This technique, for instance, turns out to be extremely useful in the study of nonlinear PDEs. After establishing strong (p,p), we study the maximal operator $T^{(*)}$ and show a pointwise integrable bound for it and strong (p,p) as well. These results will guarantee the existence of principal values in $L^p(\mathbf{R}^n)$ as in the case of the Hilbert transform.

To begin with, analogous arguments used to prove Lemmata 6.22 and 6.31 give

Lemma 6.31. *For all $f \in S(\mathbf{R}^n)$ the limit $(Tf)(x) = \lim_{\varepsilon\downarrow}(T^{(\varepsilon)}f)(x)$ exists and there is a constant c depending on f and n such that*

$$\sup_{\varepsilon>0}\left|\int_{|y|>\varepsilon}\frac{\Omega(y)}{|y|^n}f(x-y)\,dy\right| \le \frac{c}{(1+|x|)^n}.$$

The lemma, in particular, shows that whenever $f \in S(\mathbf{R}^n)$, we also have the converge in $L^p(\mathbf{R}^n)$ and that $\sup_{\varepsilon>0}|(T^{(\varepsilon)}f)(x)|$ belongs to $L^p(\mathbf{R}^n)$ for all $1 < p \le \infty$.

6.3.1. *Fourier transform of the operator.* The next theorem is the starting point for our study:

Theorem 6.32. *Let $f \in S(\mathbf{R}^n)$. Then there is a constant c depending only on n and A_2, but independent of ε, such that*

$$|\widehat{T^{(\varepsilon)}f}(\xi)| \le c|\widehat{f}(\xi)|, \qquad |\widehat{Tf}(\xi)| \le c|\widehat{f}(\xi)|$$

for $\xi \in \mathbf{R}^n \setminus \{0\}$. In particular, both $T^{(\varepsilon)}$ and T allow for extensions $\widetilde{T}^{(\varepsilon)}$ and $\widetilde{T}$, respectively, to $L^2(\mathbf{R}^n)$.

Proof. Write first

$$(T^{(\varepsilon,k)}f)(x) = \int_{\varepsilon<|y|<k/|\xi|} \frac{\Omega(y)}{|y|^n} f(x-y)\,dy,$$

$k > \max\{1, \varepsilon|\xi|\}$. Then we have

$$\widehat{T^{(\varepsilon,k)}f}(\xi) = \widehat{T^{(\varepsilon,k)}}(\xi)\widehat{f}(\xi),$$

where

$$\widehat{T^{(\varepsilon,k)}}(\xi) = \int_{\varepsilon<|y|<k/|\xi|} \frac{\Omega(y)}{|y|^n} e^{-i2\pi\xi\cdot y}\,dy.$$

Fix then $\xi \in \mathbf{R}^n \setminus \{0\}$. Make first the change of variables $z = y/|\xi|$, $dy = |\xi|^n dz$, we have

$$\widehat{T^{(\varepsilon,k)}}(\xi) = \int_{|\xi|\varepsilon<|y|<k} \frac{\Omega(y)}{|y|^n} e^{-i2\pi\xi'\cdot y}\,dy,$$

where we have written $\xi' = \xi/|\xi|$. In polar coordinates $y = ru$, $u \in S^{n-1}$, this takes the form

$$\begin{aligned}\widehat{T^{(\varepsilon,k)}}(\xi) &= \int_{|\xi|\varepsilon}^{k}\int_{S^{n-1}} \frac{\Omega(ru)}{|ru|^n} e^{-i2\pi r\xi'\cdot u}\,d\mathcal{H}^{n-1}(u)\,r^{n-1}\,dr \\ &= \int_{|\xi|\varepsilon}^{k}\int_{S^{n-1}} \Omega(u) e^{-i2\pi r\xi'\cdot u}\,d\mathcal{H}^{n-1}(u)\,\frac{dr}{r}.\end{aligned}$$

Applying Fubini's theorem, we obtain

$$\widehat{T^{(\varepsilon,k)}}(\xi) = \int_{S^{n-1}} \Omega(u)\left(\int_{|\xi|\varepsilon}^{k} e^{-i2\pi r\xi'\cdot u}\,\frac{dr}{r}\right) d\mathcal{H}^{n-1}(u).$$

Since $\int_{S^{n-1}} \Omega(u)\,d\mathcal{H}^{n-1}(u) = 0$, we may add a constant to the integrand and hence

$$\begin{aligned}\widehat{T^{(\varepsilon,k)}}(\xi) =& \int_{S^{n-1}} \Omega(u)\left(\int_{\min\{|\xi|\varepsilon,1\}}^{1} \left(e^{-i2\pi r\xi'\cdot u} - 1\right)\frac{dr}{r}\right) d\mathcal{H}^{n-1}(u) \\ &+ \int_{S^{n-1}} \Omega(u)\left(\int_{\max\{|\xi|\varepsilon,1\}}^{k} e^{-i2\pi r\xi'\cdot u}\,\frac{dr}{r}\right) d\mathcal{H}^{n-1}(u) \\ =& I_1(\xi) - iI_2(\xi),\end{aligned}$$

where

$$I_1^{\varepsilon,k}(\xi) = \int_{S^{n-1}} \Omega(u) \left(\int_{\min\{|\xi|\varepsilon,1\}}^1 (\cos(2\pi r\xi' \cdot u) - 1) \, \frac{dr}{r} \right) d\mathcal{H}^{n-1}(u)$$
$$+ \int_{S^{n-1}} \Omega(u) \left(\int_{\max\{|\xi|\varepsilon,1\}}^k \cos(2\pi r\xi' \cdot u) \, \frac{dr}{r} \right) d\mathcal{H}^{n-1}(u)$$

and

$$I_2^{\varepsilon,k}(\xi) = \int_{S^{n-1}} \Omega(u) \left(\int_{|\xi|\varepsilon}^k \sin(2\pi r\xi' \cdot u) \, \frac{dr}{r} \right) d\mathcal{H}^{n-1}(u)$$

Performing the change of variables $s = r|\xi' \cdot u|$ in the inner integrals (the cone $\{u \in S^{n-1} \ : |\xi' \cdot u| << 1\}$ requires some care but can be nevertheless handled), we arrive at the following identities:

$$\begin{aligned} J_1^{\varepsilon,k}(\xi, u) &:= \int_{\min\{|\xi|\varepsilon,1\}}^1 (\cos(2\pi r\xi \cdot u) - 1) \, \frac{dr}{r} \\ &= \int_{|\xi' \cdot u| \min\{|\xi|\varepsilon,1\}}^{|\xi' \cdot u|} (\cos(2\pi s) - 1) \, \frac{ds}{s}, \end{aligned}$$

$$J_2^{\varepsilon,k}(\xi, u) := \int_{\max\{|\xi|\varepsilon,1\}}^k \cos(2\pi r\xi' \cdot u) \, \frac{dr}{r} = \int_{|\xi' \cdot u| \max\{|\xi|\varepsilon,1\}}^{|\xi' \cdot u|k} \cos(2\pi s) \, \frac{ds}{s}$$

$$J_3^{\varepsilon,k}(\xi, u) := \int_{|\xi|\varepsilon}^k \sin(2\pi r\xi \cdot u) \, \frac{dr}{r} = \operatorname{sgn}(\xi \cdot u) \int_{|\xi \cdot u|\varepsilon}^{|\xi' \cdot u|k} \sin(2\pi s) \, \frac{ds}{s}$$

We now distinguish ourselves into two cases. First, assume that

$$|\xi' \cdot u|k \leq 1 \qquad \Longleftrightarrow \qquad k \leq \frac{1}{|\xi' \cdot u|}.$$

Then

$$\begin{aligned} \left| J_1^{\varepsilon,k}(\xi, u) + J_2^{\varepsilon,k}(\xi, u) \right| &\leq \Bigg(\int_{|\xi' \cdot u| \min\{|\xi|\varepsilon,1\}}^{|\xi' \cdot u|k} |\cos(2\pi s) - 1| \, \frac{ds}{s} \\ &\qquad + \int_{|\xi' \cdot u| \max\{|\xi|\varepsilon,1\}}^{|\xi' \cdot u|k} \frac{ds}{s} \Bigg) \chi_{|\xi' \cdot u|k \leq 1} \\ &\leq \left(\int_0^1 |\cos(2\pi s) - 1| \, \frac{ds}{s} + \log(k) \right) \chi_{|\xi' \cdot u|k \leq 1} \\ &\leq \Big(c - \log|u \cdot \xi'| \Big) \chi_{|\xi' \cdot u|k \leq 1}. \end{aligned}$$

If $|\xi' \cdot u|k > 1$, rewrite

$$
\begin{aligned}
J_1^{\varepsilon,k}(\xi,u) = &\int_{\min\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\}}^{1} (\cos(2\pi s)-1)\,\frac{ds}{s} \\
&+ \int_{\min\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\}}^{1} \frac{ds}{s} \\
&+ \int_{\max\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\}}^{|\xi'\cdot u|k} \cos(2\pi s)\,\frac{ds}{s}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
J_1^{\varepsilon,k}(\xi,u) + J_2^{\varepsilon,k}(\xi,u) = &\int_{\min\{1,|\xi\cdot u|\varepsilon\}}^{1} (\cos(2\pi s)-1)\,\frac{ds}{s} \\
&+ \int_{\min\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\}}^{1} \frac{ds}{s} \\
&+ \int_{\max\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\}}^{|\xi'\cdot u|k} \cos(2\pi s)\,\frac{ds}{s}.
\end{aligned}
$$

The second integral on the right is

$$
\int_{\min\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\}}^{1} \frac{ds}{s} = -\log(\min\{1,|\xi'\cdot u|\max\{|\xi|\varepsilon,1\}\})
$$

Now, Lemma 6.5, and an analogous version, where sin is replaced with cos, proves that there is a universal constant c, independent of ξ, ε, and k, such that

$$
|J_j^{(\varepsilon,k)}(\xi,u)| \le c\,(1-\log|u\cdot\xi'|)\,, \qquad j=1,2,3.
$$

But since $|\Omega|$ is bounded by the constant A_2 and $\log|u\cdot\xi'| \in L^1(S^{n-1})$, we arrive at the result with a constant depending only on n and A_2 after letting $k\to\infty$. $\square$

In fact, analyzing further the above proof, we obtain:

Theorem 6.33. *Let $f \in S(\mathbf{R}^n)$. Then*

$$
\widehat{Tf}(\xi) = \int_{S^{n-1}} \Omega(u)\left(\frac{1}{\log|u\cdot\xi/|\xi||} - i\frac{\pi}{2}\operatorname{sgn}(u\cdot\xi)\right) d\mathcal{H}^{n-1}(u)\widehat{f}(\xi)
$$

for all $\xi \in \mathbf{R}^n\setminus\{0\}$.

Proof. Exercise. $\square$

This theorem would allow us to relax the condition on the boundedness of Ω in connection to strong $(2,2)$. Define

$$
\Omega_o(u) = \frac{1}{2}\left(\Omega(u)-\Omega(-u)\right), \qquad \Omega_e(u) = \frac{1}{2}\left(\Omega(u)+\Omega(-u)\right).
$$

Then clearly Ω_o is an odd function and Ω_e is an even function.

Theorem 6.34. *Let $f \in S(\mathbf{R}^n)$. Suppose that $\Omega_o \in L^1(S^{n-1})$ and $\Omega_e \in (L\log(L))(S^{n-1})$, i.e.*

$$|\Omega_e| \max\left(\log(|\Omega_e|), 0\right) \in L^1(S^{n-1}).$$

Then T is strong $(2,2)$ in $S(\mathbf{R}^n)$.

Proof. Exercise. The proof goes in two steps. First, show that

$$AB \le A\log(A) + e^B, \qquad A \ge 1,\ B \ge 0.$$

Second, using the above estimate and Theorem 6.33, conclude the proof. □

6.3.2. *Strong (p,p).* The next step in our approach is to show strong (p,p). For this we apply methods from the Nonlinear Calderón-Zygmund Theory.

Theorem 6.35. *Let $1 < p < \infty$ and $f \in S(\mathbf{R}^n)$. Then there is a constant $c_p = c_p(n, p, c_2, A_1, A_2)$ such that*

$$\|Tf\|_p \le c_p\|f\|_p, \qquad \|T^{(\varepsilon)}f\|_p \le c_p\|f\|_p,$$

for all $\varepsilon > 0$.

Proof. Step 1: Basic notation. To stylize the notation, we will henceforth use the abbreviation $T = \widetilde{T}$, when using the extension from $S(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n)$, or $T = T^{(\varepsilon)}$, when considering the case $\varepsilon > 0$. The extension of T is used when we are for instance splitting f into two parts, which are not necessarily in $S(\mathbf{R}^n)$ anymore.

To begin with, set

$$\lambda_0^2 = \int_{\mathbf{R}^n} |Tf|^2\,dx + \delta^{-2}\int_{\mathbf{R}^n} |f|^2\,dx$$

for a small constant $\delta \in (0,1)$, which will be fixed on the course of the proof. Define

$$f_\lambda := \frac{f}{\lambda\lambda_0}, \qquad \lambda > 0,$$

and

$$J_\lambda(U) := \fint_U |Tf_\lambda|^2\,dx + \delta^{-2}\fint_U |f|^2\,dx$$

for any Borel set U in $\mathbf{R}^n$ with $|U| > 0$. Recall that $\fint_U = |U|^{-1}\int$. Moreover, we denote the level set as

$$E_\lambda = E_\lambda(1) := \{x \in \mathbf{R}^n \ :\ |Tf_\lambda(x)| > 1\}\ .$$

Step 2: Decomposition of the level set. In Step 2 we will prove the following lemma:

Lemma 6.36. *For any $\lambda > 0$ there exists a numerable family of disjoint balls $\{B(x_i, \varrho_i)\}_i$, $x_i \in E_\lambda$ and $\varrho_i = \varrho_i(x_i, \lambda) > 0$, such that*

$$J_\lambda(B(x_i, \varrho_i)) = 1, \qquad J_\lambda(B(x_i, \varrho_i)) \leq 1 \quad \forall \varrho > \varrho_i, \tag{6.37}$$

and

$$E_\lambda \subset \bigcup_i B(x_i, 5\varrho_i). \tag{6.38}$$

Moreover, for any $\theta > 1$, we have

$$\begin{aligned} \sum_i \int_{B(x_i, \theta\varrho_i)} |f|^2 \, dx \leq & 2\delta^2\theta^n \int_{\{|Tf_\lambda| > 1/2\}} |Tf_\lambda|^2 \, dx \\ & + 2\theta^n \int_{|f_\lambda| > \delta/2} |f_\lambda|^2 \, dx. \end{aligned} \tag{6.39}$$

Proof. Estimate first

$$\begin{aligned} J_\lambda(B(x, \varrho)) = & \fint_{B(x,\varrho)} |Tf_\lambda|^2 \, dx + \delta^{-2} \fint_{B(x,\varrho)} |f_\lambda|^2 \, dx \\ = & \frac{1}{B(x, \varrho)} \frac{1}{(\lambda\lambda_0)^2} \left(\int_{B(x,\varrho)} |Tf|^2 \, dx + \delta^{-2} \int_{B(x,\varrho)} |f|^2 \, dx \right) \\ \leq & \frac{1}{B(x, \varrho)} \frac{1}{(\lambda\lambda_0)^2} \left(\int_{\mathbf{R}^n} |Tf|^2 \, dx + \delta^{-2} \int_{\mathbf{R}^n} |f|^2 \, dx \right) \\ = & \frac{1}{\lambda^2 B(x, \varrho)}. \end{aligned}$$

Thus, for $r = r(\lambda) > 0$ such that $\lambda^2 |B_r| = 1$, we have

$$\sup_{x \in \mathbf{R}^n, \varrho \geq r} J_\lambda(B(x, \varrho)) \leq 1$$

for all $\varrho \geq r$. On the other hand, Lebesgue's differentiation theorem gives that

$$\lim_{\varrho \to 0} J_\lambda(B(x, \varrho)) > 1$$

for almost every $x \in E_\lambda(1)$, implying, in particular, that for almost every $x \in E_\lambda(1)$ there is $\varrho_x \in (0, r]$ such that

$$J_\lambda(B(x, \varrho_x)) = 1, \qquad J_\lambda(B(x, \varrho)) \leq 1, \quad \varrho > \varrho_x.$$

Indeed, for a fixed x, it is easy to see that $J_\lambda(B(x, \cdot)) : \mathbf{R}^+ \mapsto \mathbf{R}^+$ is continuous. Since this can be done for almost every $x \in E_\lambda(1)$, there is a dense subset $\{\widetilde{x}_i\} \subset E_\lambda(1)$ for which each member $\widetilde{x}_i$ satisfy the above condition and $E_\lambda(1) \subset \cup_i B(\widetilde{x}_i, \varrho_{\widetilde{x}_i})$. Appealing then to Vitali's covering theorem, Theorem 2.13, we find a (countable) subset $\{x_i\}$ of $\{\widetilde{x}_i\}$ such that $\{B(x_i, \varrho_i)\}$, $\varrho_i = \varrho_{x_i}$, is a disjoint family of balls satisfying (6.37) and (6.38).

Next, denote in short $B_i := B(x_i, \varrho_i)$. Since $J_\lambda(B_i) = 1$, we have

$$|B_i| = \int_{B_i} |Tf_\lambda|^2 \, dx + \delta^{-2} \int_{B_i} |f_\lambda|^2 \, dx.$$

Using the elementary inequality

$$\int_{B_i} |Tf_\lambda|^2 \, dx \le \int_{B_i \cap \{|Tf_\lambda| > 1/2\}} |Tf_\lambda|^2 \, dx + \frac{1}{4}|B_i|$$

and similarly for the term containing f_λ, we obtain

$$|B_i| \le 2 \int_{B_i \cap \{|Tf_\lambda| > 1/2\}} |Tf_\lambda|^2 \, dx + \frac{2}{\delta^2} \int_{B_i \cap \{|f_\lambda| > \delta/2\}} |f_\lambda|^2 \, dx.$$

Moreover, for any $\theta \ge 1$,

$$\frac{1}{\delta^2} \int_{B(x_i, \theta \varrho_i)} |f_\lambda|^2 \, dx \le \theta^n |B_i| J_\lambda(B(x_i, \theta \varrho_i)) \le \theta^n |B_i|.$$

The last two estimates give (6.39), since $\{B_i\}_i$ is a disjoint family of balls. $\square$

Step 3: Decomposition of f. Fix $i \ge 1$ and let

$$f_\lambda^1(x) := \begin{cases} f_\lambda(x) & \text{in } B(x_i, 25\varrho_i), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$f_\lambda^2(x) = f_\lambda(x) - f_\lambda^1(x)\,.$$

We will now show that Tf_λ^2 is bounded in $B(x_i, 5\varrho_i)$.

Lemma 6.40. *There exists N depending only on n, A_1, A_2 such that*

$$\sup_{x \in B(x_i, 5\varrho_i)} |Tf_\lambda^2(x)| \le N.$$

Proof. Let us use the short-hand notation $\theta B_i := B(x_i, \theta \varrho_i)$, $\theta > 0$. Let $z \in 5B_i$. Rewrite

$$Tf_\lambda^2(z) = \fint_{5B_i} Tf_\lambda^2 \, dx + \fint_{5B_i} (Tf_\lambda^2(z) - Tf_\lambda^2) \, dx =: I_1 + I_2.$$

We first estimate I_1. For this, estimate simply as

$$|I_1| \le \fint_{5B_i} |Tf_\lambda^2| \, dx \le \fint_{5B_i} |Tf_\lambda^1| \, dx + \fint_{5B_i} |Tf_\lambda| \, dx.$$

Hölder's inequality, together with (6.37), implies

$$\fint_{5B_i} |Tf_\lambda| \, dx \le \left(\fint_{5B_i} |Tf_\lambda|^2 \, dx \right)^{1/2} \le (J_\lambda(5B_i))^{1/2} \le 1.$$

Strong $(2,2)$ with the constant $c_2 = c_2(n, A_2)$, on the other hand, gives

$$\begin{aligned}
\fint_{5B_i} |Tf_\lambda|\,dx &\le \left(\fint_{5B_i} |Tf_\lambda^1|^2\,dx \right)^{1/2} \\
&\le \left(\frac{c_2}{|5B_i|} \int_{\mathbf{R}^n} |f_\lambda^1|^2\,dx \right)^{1/2} \\
&= \left(c_2 5^n \fint_{25B_i} |f_\lambda^1|^2\,dx \right)^{1/2} \le \delta\sqrt{c_2 5^n}.
\end{aligned}$$

Therefore,

$$|I_1| \le 1 + \sqrt{c_2 5^n} \le \frac{N}{5}$$

for suitably large N.

We then estimate I_2. Note first that for any $x \in 5B_i$,

$$\begin{aligned}
&|Tf_\lambda^2(z) - Tf_\lambda^2(x)| \\
&\quad \le \int_{\mathbf{R}^n \setminus 25B_i} \left| \frac{\Omega(x-y)}{|x-y|^n} \chi_{|x-y|>\varepsilon} - \frac{\Omega(z-y)}{|z-y|^n} \chi_{|z-y|>\varepsilon} \right| |f_\lambda^2(y)|\,dy.
\end{aligned}$$

Rewrite the kernel as

$$\begin{aligned}
&\left| \frac{\Omega(x-y)}{|x-y|^n} \chi_{|x-y|>\varepsilon} - \frac{\Omega(z-y)}{|z-y|^n} \chi_{|z-y|>\varepsilon} \right| \\
&\qquad = \left| \frac{\Omega(x-y)}{|x-y|^n} - \frac{\Omega(z-y)}{|z-y|^n} \right| \chi_{|x-y|,|z-y|>\varepsilon} \\
&\qquad\quad + \left| \frac{\Omega(x-y)}{|x-y|^n} \right| \chi_{|x-y|>\varepsilon, |z-y|\le\varepsilon} \\
&\qquad\quad + \left| \frac{\Omega(z-y)}{|z-y|^n} \right| \chi_{|z-y|>\varepsilon, |x-y|\le\varepsilon}.
\end{aligned} \tag{6.41}$$

Let us first handle the last two terms on the right. Since $f_\lambda^2(y) = 0$ in $25B_i$, we may assume that $y \in \mathbf{R}^n \setminus 25B_i$. Moreover, since $z, x \in 5B_i$, we have

$$20\varrho_i \le \min\{|z-y|, |x-y|\}, \qquad |z-x| \le 10\varrho_i$$

In particular, if $|z-y| > \varepsilon$ and $|x-y| \le \varepsilon$, then

$$\varepsilon \ge |x-y| \ge 20\varrho_i, \quad \varepsilon < |z-y| \le |x-y| + |z-x| \le \varepsilon + 10\varrho_i \le \frac{3}{2}\varepsilon.$$

It follows that

$$|y - x_i| \le |y-z| + |z-x_i| \ge \varepsilon + 5\varrho_i \le \frac{3}{2}\varepsilon + \frac{1}{4}\varepsilon < 2\varepsilon.$$

Therefore,

$$\begin{aligned}&\left|\frac{\Omega(z-y)}{|z-y|^n}\right|\chi_{|z-y|>\varepsilon,|x-y|\le\varepsilon}|f_\lambda^2(y)|\\&\qquad\le 2A_1\varepsilon^{-n}|f_\lambda(y)|\chi_{|y-x_i|<2\varepsilon}\le 2^{n+1}A_1\frac{\chi_{B(x_i,2\varepsilon)}(y)}{|B(x_i,2\varepsilon)|}|f_\lambda(y)|\end{aligned}$$

and consequently also

$$\int_{\mathbf{R}^n}\left|\frac{\Omega(z-y)}{|z-y|^n}\right|\chi_{|z-y|>\varepsilon,|x-y|\le\varepsilon}|f_\lambda^2(y)|\,dy\le\delta 2^{n+1}A_1\le\frac{N}{5}$$

holds by Hölder's inequality and the fact that $J_\lambda(B(x_i,2\varepsilon))\le 1$, $\varepsilon>\varrho_i$. Similarly, we obtain

$$\int_{\mathbf{R}^n}\left|\frac{\Omega(x-y)}{|x-y|^n}\right|\chi_{|x-y|>\varepsilon,|z-y|\le\varepsilon}|f_\lambda^2(y)|\,dy\le\delta 2^{n+1}A_1\le\frac{N}{5}.$$

We will then estimate the remaining term in (6.41) preliminary as

$$\begin{aligned}&\left|\frac{\Omega(x-y)}{|x-y|^n}-\frac{\Omega(z-y)}{|z-y|^n}\right|\\&\qquad\le A_1\left|\frac{1}{|x-y|^n}-\frac{1}{|z-y|^n}\right|\\&\qquad\quad+\frac{1}{|x-y|^n}\,|\Omega(x-y)-\Omega(z-y)|\end{aligned}\tag{6.42}$$

Thus we need upper bounds for

$$I_3:=\int_{\mathbf{R}^n\setminus 25B_i}\left|\frac{1}{|x-y|^n}-\frac{1}{|z-y|^n}\right||f_\lambda^2(y)|\,dy$$

and

$$I_4:=\int_{\mathbf{R}^n\setminus 25B_i}\frac{1}{|x-y|^n}\,|\Omega(x-y)-\Omega(z-y)|\,|f_\lambda^2(y)|\,dy$$

Recall here that $x,z\in 5B_i$ and $y\in\mathbf{R}^n\setminus 25B_i$ so that

$$|x-z|\le 10\varrho_i,\qquad |y-z|\ge 20\varrho_i\ge 2|x-z|$$

By considering the smooth function $g:[0,1]\mapsto\mathbf{R}$,

$$g(t)=\frac{1}{|z-y+t(x-z)|^n},$$

we find by the mean value theorem $t_0\in[0,1]$ such that

$$\begin{aligned}&\left|\frac{1}{|x-y|^n}-\frac{1}{|z-y|^n}\right|\\&\qquad=|g(1)-g(0)|=|g'(t_0)|=\frac{n|x-z|}{|z-y+t_0(x-z)|^{n+1}}\\&\qquad\le 10n2^{n+1}\frac{\varrho_i}{|z-y|^{n+1}}\le 10n4^{n+1}\frac{\varrho_i}{|x_i-y|^{n+1}}.\end{aligned}$$

We obtain

$$\begin{aligned}
I_3 =& \int_{\mathbf{R}^n \setminus 25B_i} \left| \frac{1}{|x-y|^n} - \frac{1}{|z-y|^n} \right| |f_\lambda^2(y)|\, dy \\
\leq& 10n4^{n+1} \sum_{j=2}^{\infty} \int_{5^{j+1}B_i \setminus 5^j B_i} \frac{\varrho_i}{|x_i - y|^{n+1}} |f_\lambda^2(y)|\, dy \\
\leq& 10n4^{n+1} \sum_{j=2}^{\infty} 5^{-j} \fint_{5^{j+1}B_i} |f_\lambda^2(y)|\, dy
\end{aligned}$$

But now Hölder's inequality gives

$$\begin{aligned}
\fint_{5^{j+1}B_i} |f_\lambda^2(y)|\, dy \leq& 2 \fint_{5^{j+1}B_i} |f_\lambda(y)|\, dy \\
\leq& 2 \left(\fint_{5^{j+1}B_i} |f_\lambda(y)|^2\, dy \right)^{1/2} \leq 2\delta \sqrt{J_\lambda(5^{j+1}B_i)} \leq 2\delta
\end{aligned} \tag{6.43}$$

and thus

$$A_1 I_3 \leq 20n4^{n+1}\delta A_1 \leq \frac{N}{5}$$

follows.

We then estimate I_4. The homogeneity of Ω first gives

$$|\Omega(x-y) - \Omega(z-y)| = \left| \Omega\left(\frac{x-y}{|x-y|} \right) - \Omega\left(\frac{z-y}{|z-y|} \right) \right|.$$

Since

$$\left| \frac{x-y}{|x-y|} - \frac{z-y}{|z-y|} \right| \leq \frac{|x-z|}{|x-y|} + |z-y| \left| \frac{1}{|x-y|} - \frac{1}{|z-y|} \right|,$$

and by considering the smooth function $g : [0,1] \mapsto \mathbf{R}$,

$$g(t) = \frac{1}{|x-y-t(x-z)|},$$

the mean value theorem gives $t_1 \in [0,1]$ such that

$$\left| \frac{1}{|x-y|} - \frac{1}{|z-y|} \right| = \frac{|x-z|}{|x-y-t_1(x-z)|},$$

we obtain similarly as before that

$$\left| \frac{x-y}{|x-y|} - \frac{z-y}{|z-y|} \right| \leq \frac{40\varrho_i}{|x_i - y|}.$$

Therefore,

$$\left| \Omega\left(\frac{x-y}{|x-y|} \right) - \Omega\left(\frac{z-y}{|z-y|} \right) \right| \leq \omega\left(\frac{40\varrho_i}{|x_i - y|} \right).$$

In an annuli $y \in 5^{j+1}B_i \setminus 5^j B_i$ we thus obtain

$$\omega\left(\frac{40\varrho_i}{|x_i - y|} \right) \leq \omega(5^{3-j}).$$

Estimate then

$$
\begin{aligned}
I_4 =& \sum_{j=2}^{\infty} \int_{5^{j+1}B_i \setminus 5^j B_i} \frac{1}{|x-y|^n} \, |\Omega(x-y) - \Omega(z-y)| \, |f_\lambda^2(y)| \, dy \\
\leq & 10^n \sum_{j=2}^{\infty} \omega(5^{3-j}) \fint_{5^{j+1}B_i} |f_\lambda^2(y)| \, dy \\
\leq & 2\delta 10^n \sum_{j=2}^{\infty} \omega(5^{3-j}),
\end{aligned}
$$

where we have also applied (6.43). The sum on the right has an upper bound

$$
\begin{aligned}
\sum_{j=2}^{\infty} \omega(5^{3-j}) =& \sum_{j=2}^{4} \omega(5^{3-j}) + \sum_{j=4}^{\infty} \omega(5^{3-j}) \\
\leq & 6A_1 + \sum_{j=1}^{\infty} \omega(5^{-j}) \, (\log(5))^{-1} \int_{5^{-j}}^{5^{1-j}} \frac{dr}{r} \\
\leq & 6A_1 + (\log(5))^{-1} \int_0^1 \omega(r) \, \frac{dr}{r} \\
\leq & 6A_1 + (\log(5))^{-1} \, A_2 \leq (2\delta 10^n)^{-1} \frac{N}{5}.
\end{aligned}
$$

The result follows. □

Step 3: Final conclusion. The details are left as an exercise. First show using Lemma 6.40 that

$$
|\{x \in 5B_i \ : \ |Tf_\lambda| > 2N\}| \leq c \int_{5B_i} |f_\lambda|^2 \, dx.
$$

Prove then the general principle in the measure theory: If $g \in L^p(\mathbf{R}^n)$, then

$$
\int_{\mathbf{R}^n} |g|^p \, dx = (p-2) \int_0^\infty \mu^{p-3} \left(\int_{\mathbf{R}^n \cap \{|g| > \mu\}} |g|^2 \, dx \right) d\mu.
$$

Finally, prove applying Lemma 6.36 and previous two estimates

$$
\begin{aligned}
\int_{\mathbf{R}^n} |Tf|^p \, dx \leq & c(n,p,A_1,A_2)\delta^2 \int_{\mathbf{R}^n} |Tf|^p \, dx \\
& + c(n,p,A_1,A_2,\delta) \int_{\mathbf{R}^n} |f|^p \, dx,
\end{aligned}
$$

and make the final conclusion from this. □

We end up the section by noting that modifying calculations in Step 2. above, it is possible to prove the following two lemmata:

Lemma 6.44. *There is a constant $c = c(n, A_1, A_2)$ such that*

$$\int_{\mathbf{R}^n \cap \{|y|>2|x|\}} \left| \frac{\Omega(y-x)}{|y-x|^n} - \frac{\Omega(y)}{|y|^n} \right| dy \le c$$

for all $x \in \mathbf{R}^n$.

Proof. Exercise. Use already proved estimates from Step 2 above. □

Lemma 6.45. *Let $\varepsilon > 0$ and suppose that $f \in S(\mathbf{R}^n)$. Let $x_0 \in \mathbf{R}^n$, $\varrho > 0$, and $g = (1 - \chi_{B(x_0, 25\varrho)})f$. Then there is a constant $c = c(n, A_1, A_2)$ such that*

$$\left|T^{(\varepsilon)}g(z) - T^{(\varepsilon)}g(x)\right| \le cMf(x), \qquad |Tg(z) - Tg(x)| \le cMf(x)$$

for all $x, z \in B(x, 5\varrho)$.

Proof. Exercise. Use already proved estimates from Step 2 above. □

6.3.3. *Weak (1,1).* Weak (1,1) is attainable via modification of the second proof of Theorem 6.11. This gives naturally an alternative way to prove strong (p,p) via Marcinkiewicz interpolation theorem using duality and strong $(2,2)$. Since the proof is essentially the same, we only sketch the needed modifications.

Theorem 6.46. *Let $f \in S(\mathbf{R}^n)$. Then there is a constant c depending only on n, A_1, A_2 such that weak* $(1,1)$

$$|\{x \in \mathbf{R}^n \ : \ |Tf(x)| > \lambda\}| \le \frac{c}{\lambda}\|f\|_1$$

holds for all $\lambda > 0$. The same estimate holds if T is replaced with $T^{(\varepsilon)}$.

Proof. Suppose that $f \in S(\mathbf{R})$ and let $\lambda > 0$.

Step 1. Application of Calderón-Zygmund decomposition. Calderón-Zygmund decomposition gives us disjoint dyadic cubes $\{Q_j\}$ such that

$$\begin{aligned}
&|f(x)| \le \lambda \qquad \text{for a.e. } x \notin \Xi := \bigcup_j Q_j,\\
&|\Xi| \le \frac{1}{\lambda}\|f\|_1,\\
&\lambda < (|f|)_j \le 2^n\lambda, \qquad (|f|)_j := \fint_{Q_j} |f(t)|\, dt.
\end{aligned}$$

Denote by c_j the center of Q_j and by θQ_j the cube centered at c_j and with θ times longer sides. Let $\theta\Xi = \cup_j \theta Q_j$.

Split f as follows:

$$g(x) = \begin{cases} f(x), & x \notin \Xi, \\ (f)_j, & x \in Q_j, \end{cases}$$

and

$$b = \sum_j b_j, \qquad b_j = (f - (f)_j)\, \chi_{Q_j}.$$

Then $f = g + b$, $|g| \le 2\lambda$ almost everywhere and $\int_{\mathbf{R}^n} b_j(t)\,dt = 0$.

Step 2. Bound for $\int_{\mathbf{R}^n\setminus\sqrt{4n}Q_j} \widetilde{T}b_j(x)\,dx$. Using the extension of T (or $T^{(\varepsilon)}$) to $L^2(\mathbf{R}^n)$, both $\widetilde{T}g$ and $\widetilde{T}b$ exist. Since b_j has zero integral, we have formally

$$\widetilde{T}b_j(x) = \int_{Q_j} K(x-y)b_j(y)\,dy = \int_{Q_j} (K(x-y) - K(x-c_j))\,b_j(y)\,dy.$$

for almost every $x \in \mathbf{R}^n$. Here we have denoted

$$K(x) = \frac{\Omega(x)}{|x|^n}.$$

(To make the calculation rigorous, one should justify $\widetilde{T}b_j = Tb_j$ almost everywhere. This goes along the lines made for the Hilbert transform.) Fubini's theorem now gives

$$\begin{aligned}
&\left|\int_{\mathbf{R}^n\setminus\sqrt{4n}Q_j} \widetilde{T}b_j(x)\,dx\right| \\
&\qquad \le \int_{Q_j} |b_j(y)| \left(\int_{\mathbf{R}^n\setminus\sqrt{4n}Q_j} |K(x-y) - K(x-c_j)|\,dx\right) dy.
\end{aligned}$$

The inner integral is bounded by Lemma 6.44, since

$$\mathbf{R}^n \setminus \sqrt{4n}Q_j \subset \{x \in \mathbf{R}^n \;:\; |x-c_j| > 2|y-c_j|\},$$

and it follows that

$$\left|\int_{\mathbf{R}^n\setminus\sqrt{4n}Q_j} \widetilde{T}b_j(x)\,dx\right| \le c\int_{Q_j} |b_j(y)|\,dy.$$

Step 3. Weak (1,1) estimate. Since $Tf = \widetilde{T}f = \widetilde{T}g + \widetilde{T}b$ ($\widetilde{T}$ linear in $L^2(\mathbf{R}^n)$), we have

$$\begin{aligned}
&|\{x \in \mathbf{R}^n \;:\; |Hf(x)| > \lambda\}| \\
&\qquad \le \left|\{x \in \mathbf{R}^n \;:\; |\widetilde{T}g(x)| > \lambda/2\}\right| + \left|\{x \in \mathbf{R}^n \;:\; |\widetilde{T}b(x)| > \lambda\}\right|.
\end{aligned}$$

Since $\widetilde{T}$ is strong (2,2), together with the fact that $0 \le |g| \le 2^n\lambda$ almost everywhere, implies

$$\begin{aligned}
\left|\{x \in \mathbf{R}^n \;:\; |\widetilde{T}g(x)| > \lambda\}\right| &\le \frac{1}{\lambda^2}\int_{\mathbf{R}^n} |\widetilde{T}g(x)|^2\,dx = \frac{c_2}{\lambda^2}\int_{\mathbf{R}^n} |g|^2\,dx \\
&\le \frac{c}{\lambda}\int_{\mathbf{R}^n} |g|\,dx = \frac{c}{\lambda}\int_{\mathbf{R}^n\setminus\Xi} |f|\,dx
\end{aligned}$$

Next, we have

$$\left|\{x \in \mathbf{R}^n \;:\; |\widetilde{T}b(x)| > \lambda\}\right| \le |\sqrt{4n}\Xi| + \frac{1}{\lambda}\int_{\mathbf{R}^n\setminus\sqrt{4n}\Xi} |\widetilde{T}b(x)|\,dx \le \frac{c}{\lambda}\|f\|_1$$

by Step 2 and by the fact that $|\sqrt{4n}\Xi| \le (4n)^{n/2}|\Xi| \le c\lambda^{-1}\|f\|_1$ from the Calderón-Zygmund decomposition. Combining estimates gives

$$|\{x \in \mathbf{R}^n \ : \ |Tf(x)| > \lambda\}| \le \frac{c}{\lambda}\|f\|_1,$$

concluding the proof. □

6.3.4. *Existence of limits.* In this section we establish the existence of the limit Tf whenever $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$. The approach has been already paved in the study of the Hilbert transform. The necessary object to study is the maximal operator

$$T^{(*)}f(x) := \sup_{\varepsilon>0} |T^{(\varepsilon)f}(x)|.$$

We will first show a pointwise bound with the aid of the Hardy-Littlewood maximal function. For this, we need a lemma.

Lemma 6.47. *Let S be weak $(1,1)$ with the constant c_1 and let $\nu \in (0,1)$. Then there exists $c = c(\nu, c_1)$ such that*

$$\int_E |Sf(y)|\,dy \le c|E|^{1-\nu}\|f\|_1^\nu$$

for any set E with a finite measure.

Proof. Weak (1,1), together with Cavalieri's principle, implies

$$\begin{aligned}
\int_E |Sf(y)|^\nu\,dy =&\nu \int_0^\infty \lambda^{\nu-1}|\{x \in E \ : \ |Sf(x)| > \lambda\}\,d\lambda \\
\le&\nu \int_0^\infty \lambda^{\nu-1}\min\{|E|, c_1\lambda^{-1}\|f\|_1\}\,d\lambda \\
=&\nu|E| \int_0^{c_1\|f\|_1/|E|} \lambda^{\nu-1}\,d\lambda + \nu c_1\|f\|_1 \int_{c_1\|f\|_1/|E|}^\infty \lambda^{\nu-2}\,d\lambda \\
=&c_1^\nu|E|^{1-\nu}\|f\|_1^\nu + c_1^\nu\frac{\nu}{1-\nu}|E|^{1-\nu}\|f\|_1^\nu \\
=&\frac{c_1^\nu}{1-\nu}|E|^{1-\nu}\|f\|_1^\nu.
\end{aligned}$$

□

We now prove the pointwise bound commonly cited as Cotlar's inequality.

Lemma 6.48. *Let $f \in S(\mathbf{R}^n)$. Then, for all $\nu \in (0,1]$, there is a constant $c = c(n, A_1, A_2, \nu)$ such that*

$$(T^{(\varepsilon)}f)(x) \le c((M(|Tf|^\nu)(x)^{1/\nu} + Mf(x)). \tag{6.49}$$

Proof. Fix $f \in S(\mathbf{R}^n)$ and $\varepsilon > 0$. Denote $\theta B = B(x, 5^{-1}\theta\varepsilon)$, $\theta > 0$ and $B = 1B$. Let $f_1 = f\chi_{5B}$ and $f_2 = f - f_1$. Observe that

$$(T^{(\varepsilon)}f)(x) = (T^{(\varepsilon)}f_2)(x) = (Tf_2)(x)$$

Lemma 6.45 implies that

$$|Tf_2(x) - Tf_2(z)| \le cMf(x)$$

whenever $z \in B$. It readily follows that

$$|(T^{(\varepsilon)}f)(x)| \le cMf(x) + |(Tf_1)(z)| + |(Tf)(z)|, \qquad z \in B.$$

Assume now that $|(T^{(\varepsilon)}f)(x)| > 0$, for if it is not, there is nothing to prove. Take any λ, $0 < \lambda < |(T^{(\varepsilon)}f)(x)|$ and define

$$B_1 := \{z \in B \ : \ |(Tf_1)(z)| > \lambda/3\}, \quad B_2 := \{z \in B \ : \ |(Tf)(z)| > \lambda/3\},$$

and

$$B_3 := \begin{cases} \emptyset, & \text{if } cMf(x) \le \lambda/3, \\ B, & \text{otherwise.} \end{cases}$$

Then $B = B_1 \cup B_2 \cup B_3$ and consequently

$$|B| \le |B_1| + |B_2| + |B_3|.$$

The weak $(1,1)$ for T implies (to be accurate, we are using here $\widetilde{T}$, but in B, $\widetilde{T}f = Tf$)

$$|B_1| = |\{z \in B \ : \ |(Tf_1)(z)| > \lambda/3\}| \le \frac{c}{\lambda}\int_{5B} |f(y)|\, dy \le \frac{c}{\lambda}|B| Mf(x)$$

For $|B_2|$ we have the estimate

$$|\{z \in B \ : \ |(Tf)(z)| > \lambda/3\}| \le \frac{3}{\lambda}\int_B |Tf(z)|\, dz \le \frac{c}{\lambda}|B| M(Tf)(x).$$

If now $B_3 = \emptyset$, i.e. $|B| \le |B_1| + |B_2|$, then

$$1 \le \frac{c}{\lambda} Mf(x) + \frac{c}{\lambda} M(Tf)(x).$$

If $B_3 = B$, then

$$\lambda \le 3cMf(x)$$

In any case the result follows and for $\nu = 1$.

Let now $0 < \nu < 1$. Then we actually have

$$|(T^{(\varepsilon)}f)(x)|^\nu \le cMf(x)^\nu + |(Tf_1)(z)|^\nu + |(Tf)(z)|^\nu, \qquad z \in B.$$

Averaging this over B and then taking the power $1/\nu$, we obtain

$$|(T^{(\varepsilon)}f)(x)| \le cMf(x) + c\left(⨍_B |(Tf_1)(z)|^\nu\, dz\right)^{1/\nu} + cM(|Tf|^\nu)^{1/\nu}$$

The previous lemma implies that

$$\left(⨍_B |(Tf_1)(z)|^\nu\, dz\right)^{1/\nu} \le c|B|^{-1}\|f_1\|_1 \le cMf(x),$$

concluding the proof. $\square$

The pointwise bound allows us to prove strong (p,p) and weak $(1,1)$ for the maximal operator $T^{(*)}$. For weak (1,1), however, we need yet another lemma concerning the dyadic maximal function $M_d f$.

Lemma 6.50. *Let $f \in L^1(\mathbf{R}^n)$ be nonnegative. Then*

$$|\{x \in \mathbf{R}^n : \ Mf(x) > 4^n\lambda\}| \leq 3^n \, |\{x \in \mathbf{R}^n : \ M_d f(x) > \lambda\}| \, .$$

Proof. In the fifth exercise round, Exercise 5.3, it was shown that taking the Calderón-Zygmund decomposition of f at the level λ with disjoint dyadic cubes $\{Q_j\}_j$, then

$$\{x \in \mathbf{R}^n : \ M_d f(x) > \lambda\} = \bigcup_j Q_j.$$

Let $3Q_j$ be the cube with the same center as Q_j, whose sides are three times longer. Let $x \notin \cup_j 3Q_j$ and let Q be any cube containing x. Let k be an integer such that the side length of Q, $\ell(Q)$, satisfies $2^k < \ell(Q) \leq 2^{k+1}$. Then there are at most $m \leq 2^n$ of dyadic cubes, say $R_1, \dots, R_m$, with side lengths 2^k intersecting Q. None of these cubes are contained in $\{Q_j\}$, because otherwise we would have $x \in \cup_j 3Q_j$. Hence the average on f on each R_i is at most λ and, consequently

$$\fint_Q f = |Q|^{-1} \sum_i^m \int_{Q \cap R_i} f \leq \frac{2^{kn}}{|Q|} \sum_i^m \fint_{R_i} f \leq 2^n m\lambda \leq 4^n \lambda.$$

It follows that

$$\{x \in \mathbf{R}^n : \ Mf(x) > 4^n\lambda\} \subset \bigcup_j 3Q_j$$

and hence

$$|\{x \in \mathbf{R}^n : \ Mf(x) > 4^n\lambda\}| \leq 3^n \sum_j |Q_j|$$

and the result follows. □

Theorem 6.51. *Let $f \in S(\mathbf{R}^n)$. Then there are constants $c_p = c_p(n, p, A_1, A_2)$ and $c_1 = c_1(n, A_1, A_2)$ such that $T^{(*)}$ is strong (p,p) with the constant c_p and weak $(1,1)$ with the constant c_1 when restricted to $S(\mathbf{R}^n)$.*

Proof. Strong (p,p) is an immediate consequence of Lemma 6.48 with $\nu = 1$, because both T and M are strong (p,p).

To prove weak (1,1), we argue as follows. Starting from (6.49), we have

$$\begin{aligned} &\left|\left\{x \in \mathbf{R}^n \ : \ T^{(*)} f(x) > \lambda\right\}\right| \\ &\quad \leq |\{x \in \mathbf{R}^n \ : \ Mf(x) > \lambda/(2c)\}| \\ &\qquad + \left|\left\{x \in \mathbf{R}^n \ : \ M(|Tf|^\nu)(x)^{1/\nu} > \lambda/(2c)\right\}\right| . \end{aligned}$$

For the first term on the right we obviously have

$$|\{x \in \mathbf{R}^n \ : \ Mf(x) > \lambda/(2c)\}| \leq \frac{c}{\lambda} \|f\|_1.$$

For the second we first use the lemma above to obtain

$$\begin{aligned}&\left|\left\{x \in \mathbf{R}^n \ : \ M(|Tf|^\nu)(x)^{1/\nu} > (\lambda/(2c))^\nu\right\}\right|\\ &\qquad \le 3^n \left|\left\{x \in \mathbf{R}^n \ : \ M_d(|Tf|^\nu)(x) > 4^{-n}\,(\lambda/(2c))^\nu\right\}\right| .\end{aligned}$$

Forming the Calderón-Zygmund decomposition of $|Tf|^\nu$ at the level $4^{-n}\,(\lambda/(2c))^\nu$ we obtain disjoint dyadic cubes $\{Q_j\}$ such that

$$E := \left\{x \in \mathbf{R}^n \ : \ M_d(|Tf|^\nu)(x) > 4^{-n}\,(\lambda/(2c))^\nu\right\} = \bigcup_j Q_j$$

Since $f \in S(\mathbf{R}^n)$, E has finite measure. Moreover, since

$$|Q_j| \le \frac{4^n(2c)^\nu}{\lambda^\nu} \int_{Q_j} |Tf|^\nu,$$

we have that

$$|E| \le \frac{4^n(2c)^\nu}{\lambda^\nu} \int_E |Tf|^\nu.$$

Applying then Lemma 6.47 (indeed, T is weak (1,1)), we obtain

$$\int_E |Tf|^\nu \le \frac{c_1^\nu}{1-\nu} |E|^{1-\nu} \|f\|_1^\nu,$$

immediately giving

$$|E| \le \frac{c_1}{(1-\nu)^{1/\nu}} \frac{4^{n/\nu} 2c}{\lambda} \|f\|_1.$$

This concludes the proof for example taking $\nu = 1/2$. □

Previous results allow us to conclude with the L^p theory of singular integrals of convolution type. The proof is completely analogous to the case of the Hilbert transform and we omit it.

Theorem 6.52. *Let $f \in L^p(\mathbf{R}^n)$. Then the limit*

$$Tf(x) = \lim_{\varepsilon \downarrow 0} T^{(\varepsilon)} f(x)$$

exists for almost every $x \in \mathbf{R}^n$. Moreover, there are constants $c_p = c_p(n, p, A_1, A_2)$ and $c_1 = c_1(n, A_1, A_2)$ such that T is strong (p,p) with the constant c_p and weak $(1,1)$ with the constant c_1.

6.4. **Singular integrals of nonconvolution type: Calderón-Zygmund operators.** In this final section of the course we very briefly and informally comment more general class of singular integral operators, namely Calderón-Zygmund operators. We will consider functions

$$K : (\mathbf{R}^n \times \mathbf{R}^n) \setminus \Delta \mapsto \mathbf{C}, \qquad \Delta := \{(x,x) \ : \ x \in \mathbf{R}^n\}$$

that satisfy for some constant $A > 0$ the size bound

$$|K(x,y)| \le \frac{A}{|x-y|^n}, \qquad x, y \in \mathbf{R}^n,\ x \ne y,$$

and for some $\delta > 0$ the regularity conditions

$$|K(x,y)-K(z,y)| \le A\frac{|x-z|^\delta}{(|x-y|+|z-y|)^{n+\delta}}, \qquad x,z,y \in \mathbf{R}^n,\ x,z \neq y,$$

whenever $|x-z| \le 2^{-1}\max\{|x-y|,|z-y|\}$, and

$$|K(x,z)-K(x,y)| \le A\frac{|y-z|^\delta}{(|x-y|+|x-z|)^{n+\delta}}, \qquad x,z,y \in \mathbf{R}^n,\ x \neq y,z,$$

whenever $|y-z| \le 2^{-1}\max\{|x-y|,|x-z|\}$.

Function satisfying the above three conditions are called *standard kernels* and the set of them denoted by $SK(\delta, A)$. These conditions imply so-called Hörmander conditions

$$\int_{|x-y|>2|y-z|} |K(x,y)-K(x,z)|\,dx \le c \qquad \forall y,z \in \mathbf{R}^n \tag{6.53}$$

and

$$\int_{|x-y|>2|x-z|} |K(x,y)-K(z,y)|\,dy \le c \qquad \forall x,z \in \mathbf{R}^n. \tag{6.54}$$

An operator T is called a (generalized) Calderón-Zygmund operator if T is strong $(2,2)$ and there exists a standard kernel $K \in SK(\delta, A)$ associated to T such that for $f \in L^2(\mathbf{R}^n)$ with a compact support,

$$Tf(x) = \int_{\mathbf{R}^n} K(x,y)f(y)\,dy, \qquad x \notin \mathrm{spt}(f).$$

Here $\mathrm{spt}(f)$ stands for the support of f. The class of such operators is denoted by $CZO(\delta, A)$. Following ideas from previous sections it is possible to show that T defined this way is weak $(1,1)$ and strong (p,p). Actually, for this one needs only conditions (6.53) and (6.54).

To define the principal values, one very natural candidate would be to define them as limits of the approximating operator

$$T^{(\varepsilon)}f(x) = \int_{|x-y|>\varepsilon} K(x,y)f(y)\,dy.$$

However, it turns out that the limit as $\varepsilon \to 0$ needs not to exist or it may exist but be different from $Tf(x)$. The non-existence of limits occurs for example with the kernel $K(x,y) = |x-y|^{-n-it}$, $t > 0$, and the second phenomenon happens when considering the identity operator $T = I$. This is an operator in $CZO(\delta, A)$ with the standard kernel $K(x,y) = 0$. Indeed, clearly I is strong $(2,2)$ and $If(x) = 0$ whenever $x \notin \mathrm{spt}(f)$. More in general, any pointwise multiplication $Tf(x) = a(x)f(x)$, $a \in L^\infty(\mathbf{R}^n)$, has the same peculiar property.

The existence of limits can be described via the following proposition.

Proposition 6.55. *Let $T \in CZO(\delta, A)$ be associated with the kernel $K \in SK(\delta, A)$, and let Let $f \in C_0^\infty(\mathbf{R}^n)$. Then the limit*

$$\lim_{\varepsilon \downarrow 0} T^{(\varepsilon)}f(x)$$

exists if and only if the limit

$$\lim_{\varepsilon \downarrow 0} \int_{\varepsilon < |x-y| < 1} K(x,y)\,dy$$

exists for almost every $x \in \mathbf{R}^n$*.*

Proof. Suppose first that the limit

$$(Lf)(x) := \lim_{\varepsilon \downarrow 0} T^{(\varepsilon)} f(x)$$

exists for all $f \in C_0^\infty(\mathbf{R}^n)$. Fix $\phi \in C_0^\infty(\mathbf{R}^n)$ such that $\phi = 1$ in $B(x,1)$. Then

$$(L\phi)(x) = \lim_{\varepsilon \downarrow 0} \int_{\varepsilon < |x-y| < 1} K(x,y)\,dy + \int_{|x-y| \geq 1} K(x,y)\phi(y)\,dy.$$

The second integral on the right exists since $|K(x,y)| \leq A$ on $\{|x-y| \geq 1\}$. Therefore, also the limit

$$\lim_{\varepsilon \downarrow 0} \int_{\varepsilon < |x-y| < 1} K(x,y)\,dy = (L\phi)(x) - \int_{|x-y| \geq 1} K(x,y)\phi(y)\,dy$$

exists.

Conversely, assume that

$$L(x) := \lim_{\varepsilon \downarrow 0} \int_{\varepsilon < |x-y| < 1} K(x,y)\,dy$$

exists. Then

$$\begin{aligned}\lim_{\varepsilon \downarrow 0} T^{(\varepsilon)} f(x) =& Lf(x) + \lim_{\varepsilon \downarrow 0} \int_{\varepsilon < |x-y| < 1} K(x,y)(f(y) - f(x))\,dy \\ &+ \int_{|x-y| \geq 1} K(x,y)(y)\,dy\end{aligned}$$

The limit on the right exists by the dominated convergence, since

$$|K(x,y)(f(y) - f(x))| \chi_{\varepsilon < |x-y| < 1} \leq \frac{A}{|x-y|^{n-1}} \|Df\|_\infty,$$

which is integrable over $B(x,1)$ with respect to y. Thus

$$\begin{aligned}&\lim_{\varepsilon \downarrow 0} T^{(\varepsilon)} f(x) = L(x) f(x) \\ &\quad + \int_{|x-y| < 1} K(x,y)(f(y) - f(x))\,dy + \int_{|x-y| \geq 1} K(x,y)(y)\,dy.\end{aligned}$$

□

As noted before, the first limit above does not necessarily coincide with $Tf(x)$. Nonetheless, from the above proof, we have the following.

Proposition 6.56. *If two operators* $T_1, T_2 \in CZO(\delta, A)$ *are associated with the same kernel* $K \in SK(\delta, A)$*, then their difference is a pointwise multiplication operator.*

An operator is called a Calderón-Zygmund singular integral operator if it satisfies

$$Tf(x) = \lim_{\varepsilon \downarrow 0} T^{(\varepsilon)} f(x) \tag{6.57}$$

for almost every $x \in \mathbf{R}^n$.[6] As before, defining

$$T^{(*)} f(x) := \sup_{\varepsilon > 0} |T^{(\varepsilon)} f(x)|,$$

it can be shown that $T^{(*)}$ is strong (p,p) and weak $(1,1)$, readily implying that

$$\left\{ f \in L^p(\mathbf{R}^n) \ : \ \lim_{\varepsilon \downarrow 0} T^{(\varepsilon)} f \text{ exists for a.e. } x \in \mathbf{R}^n \right\}$$

is closed in $L^p(\mathbf{R}^n)$. Thus, if we can verify (6.57) for a dense subset of $L^p(\mathbf{R}^n)$, say for instance $C_0^\infty(\mathbf{R}^n)$, then the limit exists for all $f \in L^p(\mathbf{R}^n)$, $1 \leq p < \infty$.

We yet briefly comment the assumption that $T \in CZO(\delta, A)$ satisfies strong $(2,2)$. In practice, given $K \in SK(\delta, A)$, verifying strong $(2,2)$ is utterly non-trivial. As we have seen in the case of singular integrals of convolution type, this is loosely speaking the starting point in the analysis we have pursued. For a long time strong (2,2) was one of the major open problems - and in full generality still is - in the field of Harmonic Analysis. This question was eventually answered by G. David and J-L. Journé in 1984 with a theorem nowadays commonly cited as $T1$ Theorem (in most concise form it says that if T has the weak boundedness property and both $T1$ and T^*1 belong to function space of bounded mean oscillations, BMO, then T is strong $(2,2)$). The proof, however, goes well beyond the scope of this course. Another important result from Harmonic Analysis, proved in 70s by Fefferman-Stein, considers images of L^∞-functions under the singular integral operators. This topic we have also intentionally excluded from our approach. It turns out that $T \in CZO(\delta, A)$ maps $L^\infty(\mathbf{R}^n)$ to BMO.

Finally, since the function space BMO appears in these two important contexts, we give the definition of this function space. Let $f \in L^1_{loc}(\mathbf{R}^n)$ and $Q \subset \mathbf{R}^n$ be a cube. Define the sharp maximal function as

$$M^\# f(x) := \sup_{Q \ni x} ⨍_Q |f - (f)_Q| \, dx, \qquad (f)_Q = ⨍_Q f \, dx.$$

We say that $f \in L^1_{loc}(\mathbf{R}^n)$ belongs to BMO if $M^\# f \in L^\infty(\mathbf{R}^n)$. Define the corresponding seminorm as

$$\|f\|_{BMO} := \|M^\# f\|_\infty.$$

[6] Actually, in the literature, the same notion is used in the context that there is merely a subsequence $\{\varepsilon_j\}_j$ converging to zero with the above property.