HARMONIC ANALYSIS

Contents

Date: 3.11.2010.

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*

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

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*

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

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*

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

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

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

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

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

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

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

because

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

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

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

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)

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

(ii)

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

(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*

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

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

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

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:

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

□

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 $Q_1, \ldots, Q_{i-1}$. 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 $Q_1, Q_2, \ldots$, because otherwise

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

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

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

It follows that

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

for every $i = 1, 2, \ldots$, 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, $i = 1, 2, \ldots$. 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

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

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 $f \in L^p(\mathbf{R}^n)$, $1 < p \le \infty$, then $Mf \in L^p(\mathbf{R}^n)$ and there exists $C = C(n,p)$ (meaning C depends on n,p) such that*

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

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 $1 \le p \le q \le \infty$. 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

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

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

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

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

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

□

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

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

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

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

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

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

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

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

$$|Tf| \leq |Tf_1| + |Tf_2|$$

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

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

Then by Lemma 2.9, it follows that

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

Further by Fubini's theorem

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

and

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

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

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

It follows that

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

Then by using Lemma 2.9 again, we see that

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

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

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

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*

$$\begin{aligned}
\int_{B(x,r)} \frac{\left|\nabla u(y)\right|}{\left|x-y\right|^{n-1}}\,\mathrm{d}y &= \sum_{i=0}^\infty \int_{B(x,2^{-i}r)\setminus B(x,2^{-i-1}r)} \frac{\left|\nabla u(y)\right|}{\left|x-y\right|^{n-1}}\,\mathrm{d}y \\
&\le \sum_{i=0}^\infty \int_{B(x,2^{-i}r)\setminus B(x,2^{-i-1}r)} \frac{\left|\nabla u(y)\right|}{(2^{-i-1}r)^{n-1}}\,\mathrm{d}y \\
&\le \sum_{i=0}^\infty \frac{2^{-i}r}{2^{-i}r} \int_{B(x,2^{-i}r)} 2^{n-1}\frac{\left|\nabla u(y)\right|}{(2^{-i}r)^{n-1}}\,\mathrm{d}y \\
&\le C\sum_{i=0}^\infty 2^{n-1}2^{-i}r \fint_{B(x,2^{-i}r)} \left|\nabla u(y)\right|\,\mathrm{d}y \\
&\le C2^{n-1}rM\left|\nabla u\right|(x)\sum_{i=0}^\infty 2^{-i}
\end{aligned}$$

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

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

Then we calculate

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

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 $f \in L^p(\mathbf{R}^n)$, $1 \leq p \leq \infty$ and $g \in L^1(\mathbf{R}^n)$, then $(f * g)(x)$ exists for almost all $x \in \mathbf{R}^n$ and*

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

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

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

we have

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

Case $p = \infty$:

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

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*

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

(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*

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

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

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

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

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

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 $|a+b|^p \le (|a| + |b|)^p = (\frac{1}{2}2\,|a| + \frac{1}{2}2\,|b|)^p \le \frac{1}{2}(2\,|a|)^p + \frac{1}{2}(2\,|b|)^p = 2^{p-1}\,|a|^p + 2^{p-1}\,|b|^p$. By using these tools, and by adding and subtracting g, we can estimate

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

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

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

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*

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

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

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

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

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

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

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

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

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 $i=1,2,\ldots$. 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, $i=1,2,\ldots$. 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*

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

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

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

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

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

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),

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

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*

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

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. □

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

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

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).

"$\Rightarrow$" 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

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

Case $p = 1$**.**

"$\Leftarrow$" was already proven before (4.7).

"$\Rightarrow$" 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

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

□

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
30.9.2010 analysis and in the theory of PDEs.

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

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

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

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

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

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

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

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)

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

and

(iii)

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

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*

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

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

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

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

$$R_m \cap Q \ne \emptyset,\ m = 1, \ldots, l.$$

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

$$\fint_{R_m} |f(y)| \,\mathrm{d}y \le \lambda$$

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

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

Moreover,

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

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

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

5.10.2010

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

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

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

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

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

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

Similarly

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

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

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

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

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

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) $||\hat{f}||_\infty \leq ||f||_1$,
(vi) $\widehat{f(\varepsilon x)} = \frac{1}{\varepsilon}\hat{f}(\frac{\xi}{\varepsilon}) = \hat{f}_\varepsilon(\xi), \varepsilon > 0$,
(vii) $f\widehat{(x+h)} = \hat{f}(\xi)e^{2\pi i h \xi}$,
(viii) $f\widehat{(x)e^{2\pi i h x}} = \hat{f}(\xi - h)$,

Proof. (i) Integral is linear.
(ii)

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

(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)

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

(v)

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

(vi)

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

(vii)

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

(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*

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

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

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

□

5.2. **On L^1.** As stated above for $f \in L^1(\mathbf{R})$, the Fourier transform $\hat{f}(\xi) = \int_{\mathbf{R}} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x$ 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 $\int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi i x\xi}\,\mathrm{d}\xi$ 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*

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

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

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

To this end, recall that $\widehat{\phi(-\varepsilon x)} = \hat{\phi}_\varepsilon(-\xi)$ and $\widehat{f(x)e^{2\pi i hx}} = \hat{f}(\xi - h)$ 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

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

When dealing with convolutions, we showed in Theorem 3.7 that

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

If $\hat{f} \in L^1(\mathbf{R})$, then the inversion formula $f(x) = \int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi i x\xi}\,\mathrm{d}\xi$ 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

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

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 $⨍_B |f| \,\mathrm{d}x \le \left(⨍_B |f|^2 \,\mathrm{d}x \right)^{1/2}$.

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

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

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

$$||Tf_1||_\infty \le C_1 ||f_1||_1$$

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

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

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 $R \to \infty$.

References