Harmonic Analysis

LECTURE SCRIPT

Dr. J. Vybiral

May 1, 2014

Contents

1 Basic concepts

1.1 Approximation of identity

We denote by $\mathbb{R}^n$ the n-dimensional Euclidean space. Its open subsets $\Omega \subset \mathbb{R}^n$ are called *domains.*

Theorem 1.1.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain. The set of continuous functions with compact support contained in Ω is dense in $L_p(\Omega)$, $1 \le p < \infty$.*

Proof. We shall need two facts from measure theory.

i) Lebesgue measure λ in $\mathbb{R}^n$ is regular, i.e. $\lambda(A) = \inf\{\lambda(G) : G \supset A, G \,\text{open}\}$.

ii) The space of step functions, i.e. the linear span of the set $\{\chi_A : A \subset \Omega, A \text{ measurable}\}$, is dense in $L_p(\Omega)$ for every $1 \le p < \infty$.

We first consider bounded open sets $\Omega_j \subset \Omega$, $j \in \mathbb{N}$, such that $\Omega_j \subset \overline{\Omega_j} \subset \Omega_{j+1} \subset \Omega$ and $\bigcup_{j=1}^\infty \Omega_j = \Omega$.[1] Let us take $f \in L_p(\Omega)$. Then $f\chi_{\Omega_j} \to f$ in $L_p(\Omega)$ and we may restrict ourselves to $f \in L_p(\Omega)$ with compact support in Ω. Due to the second property of the Lebesgue measure, this function may be approximated by a step function $\sum_{k=1}^K \varrho_k \chi_{A_k}$ with $A_k \subset \operatorname{supp} f$. So, it is enough to approximate characteristic functions χ_B with $\overline{B}$ compact in Ω. Using the first property of the Lebesgue measure, we may restrict ourselves to bounded open sets $G \subset \overline{G} \subset \Omega$. Then the sequence of functions $x \to \max(0, 1 - k\,\mathrm{dist}(x, G))$ gives the desired approximation. $\square$

Lemma 1.1.2. *Let $f \in L_p(\mathbb{R}^n), 1 \le p < \infty$. Then $f(\cdot + h) \to f(\cdot)$ in $L_p(\mathbb{R}^n)$ if $h \to 0$.*

Proof. If f is continuous with compact support, then the result follows by uniform continuity of f and the Lebesgue dominated convergence theorem. If $f \in L_p(\mathbb{R}^n)$, we may find for every $t > 0$ a continuous function g with compact support such that $\|f - g\|_p < t$. Then

$$\begin{aligned}\|f(\cdot + h) - f(\cdot)\|_p &\le \|f(\cdot + h) - g(\cdot + h)\|_p + \|g(\cdot + h) - g(\cdot)\|_p + \|g(\cdot) - f(\cdot)\|_p \\ &\le 2t + \|g(\cdot + h) - g(\cdot)\|_p\end{aligned}$$

and the conclusion follows. $\square$

Theorem 1.1.3. *The family of functions $(K_\varepsilon)_{\varepsilon>0} \subset L_1(\mathbb{R}^n)$ is called the* approximation of identity*, if*

(K1) $\int_{\mathbb{R}^n} |K_\varepsilon(x)|dx \le C < \infty$ *for all* $\varepsilon > 0$,

(K2) $\int_{\mathbb{R}^n} K_\varepsilon(x)dx = 1$ *for all* $\varepsilon > 0$,

(K3) $\lim_{\varepsilon\to 0^+} \int_{|x|>\delta} |K_\varepsilon(x)|dx = 0$ *for all* $\delta > 0$.

Then

i) If $K \in L_1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} K(x)dx = 1$, then $K_\varepsilon(x) = \varepsilon^{-n}K(x/\varepsilon)$ is an approximation of identity.

[1] For example the sets $\Omega_j := \{x \in \Omega : |x| < j \text{ and } \mathrm{dist}(x, \partial\Omega) > 1/j\}$ will do.

ii) If $(K_\varepsilon)_{\varepsilon>0}$ is an approximation of the identity, then[2]

$$\lim_{\varepsilon\to 0^+} \|K_\varepsilon * f - f\|_p = 0$$

for every $1 \le p < \infty$ and $f \in L_p(\mathbb{R}^n)$.

Proof. (i) Let $K \in L_1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} K(x)dx = 1$. Then we get immediately

$$\varepsilon^{-n}\int_{\mathbb{R}^n} K(x/\varepsilon)dx = \int_{\mathbb{R}^n} K(x)dx = 1 \quad \text{and} \quad \varepsilon^{-n}\int_{\mathbb{R}^n} |K(x/\varepsilon)|dx = \int_{\mathbb{R}^n} |K(x)|dx = \|K\|_1 < \infty.$$

As for the third point, we have

$$\int_{|x|>\delta} |K_\varepsilon(x)|dx = \int_{|y|>\delta/\varepsilon} |K(y)|dy \to 0$$

as $\varepsilon \to 0^+$, due to the Lebesgue dominated convergence theorem.

(ii) We calculate for $p > 1$ and its conjugated index p' with $1/p + 1/p' = 1$ using Hölder's inequality[3] (if $p = 1$, the calculation becomes slightly simpler)

$$\begin{aligned}
\|K_\varepsilon * f - f\|_p^p &= \int_{\mathbb{R}^n} |(K_\varepsilon * f)(x) - f(x)|^p dx = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} K_\varepsilon(y) f(x-y)dy - f(x)\right|^p dx \\
&= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} K_\varepsilon(y)[f(x-y) - f(x)]dy\right|^p dx \\
&\le \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |K_\varepsilon(y)|^{1/p+1/p'} \cdot |f(x-y) - f(x)|dy\right)^p dx \\
&\le \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |K_\varepsilon(y)| \cdot |f(x-y) - f(x)|^p dy \cdot \left(\int_{\mathbb{R}^n} |K_\varepsilon(y)|dy\right)^{p/p'} dx \\
&\le C^{p/p'} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |K_\varepsilon(y)| \cdot |f(x-y) - f(x)|^p dydx \\
&= C^{p/p'} \int_{\mathbb{R}^n} |K_\varepsilon(y)| \cdot \|f(\cdot - y) - f(\cdot)\|_p^p dy \\
&\le C^{p/p'} \left\{ \int_{|y|\le\delta} |K_\varepsilon(y)| \cdot \|f(\cdot - y) - f(\cdot)\|_p^p dy + 2^p \|f\|_p^p \int_{|y|>\delta} |K_\varepsilon(y)|dy \right\}
\end{aligned}$$

for every $\delta > 0$. Using *(K3)* and previous Lemma, we obtain the conclusion of the theorem. □

Definition 1.1.4. Let $\Omega \subset \mathbb{R}^n$ be a domain. Then $C_c^\infty(\Omega)$ denotes the set of infinitely-differentiable functions compactly supported in Ω.

It is easy to show (but not completely trivial) that this class is actually non-empty. One example is the famous function

$$f(x) = \begin{cases} \exp\left(-\frac{1}{1-|x|^2}\right), & \text{if } |x| \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Theorem 1.1.5. *$C_c^\infty(\Omega)$ is dense in $L_p(\Omega)$ for every $1 \le p < \infty$ and every domain $\Omega \subset \mathbb{R}^n$.*

[2] $(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy = \int_{\mathbb{R}^n} f(y)g(x-y)dy$ is the convolution of f and g.

[3] $\int fg \le \int |fg| \le (\int |f|^p)^{1/p} (\int |g|^{p'})^{1/p'}$

Proof. Let $f \in L_p(\Omega)$. First, we approximate f by a continuous and compactly supported g (i.e. $\|f-g\|_p \le h$) and then consider the functions $\omega_\varepsilon * g$, where $\omega_\varepsilon(x) = \varepsilon^{-n}\omega(x/\varepsilon)$ and $\omega \in C_c^\infty(\mathbb{R}^n)$ has compact support and $\int \omega = 1$. It follows that $\omega_\varepsilon * g \in C_c^\infty(\mathbb{R}^n)$ (the support property is clear, the differentiability follows by taking differences and limits, use Lebesgue's dominated convergence theorem). Together with the formula $\|\omega_\varepsilon * g - g\|_p \to 0$, the conclusion follows. $\square$

1.2 Maximal operator

For $x \in \mathbb{R}^n$ and $r > 0$, we denote by $B(x,r)$ the ball in $\mathbb{R}^n$ with center at x and radius r, i.e. $B(x,r) = \{y \in \mathbb{R}^n : |x-y| < r\}$.

Let f be a locally integrable function on $\mathbb{R}^n$. Then we define the *Hardy-Littlewood maximal operator* of f by

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

Here stands $|B(x,r)|$ for the Lebesgue measure of $B(x,r)$.

Let us note, that maximal operator is not linear, but is sub-linear, i.e.

$$\begin{aligned} M(f+g)(x) &= \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)+g(y)|dy \\ &\le \sup_{r>0} \frac{1}{|B(x,r)|} \left(\int_{B(x,r)} |f(y)|dy + \int_{B(x,r)} |g(y)|dy \right) \\ &\le (Mf)(x) + (Mg)(x). \end{aligned}$$

Simple modifications include *cubic centered maximal operator*

$$M'f(x) = \sup_{Q} \frac{1}{|Q|} \int_{x+Q} |f(y)|dy,$$

cubic non-centered maximal operator

$$M''f(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_{Q} |f(y)|dy,$$

where the supremum runs over all cubes Q containing x, and the *dyadic maximal operator* M_d, when the supremum is taken over all *dyadic* cubes containing x. These are cubes of the type $2^k(q + [0,1)^n)$, where $k \in \mathbb{Z}$ and $q \in \mathbb{Z}^n$.

We shall study the mapping properties of the operator M in the frame of Lebesgue spaces $L_p(\mathbb{R}^n)$, $1 \le p \le \infty$. If $f \in L_\infty(\mathbb{R}^n)$, then

$$\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|dy \le \frac{\|f\|_\infty}{|B(x,r)|} \int_{B(x,r)} 1dy = \|f\|_\infty$$

holds for every $x \in \mathbb{R}^n$ and every $r > 0$ and $\|Mf\|_\infty \le \|f\|_\infty$ follows. To deal with other p's, we need some more notation first.

Let $f \in L_1(\mathbb{R}^n)$ and let $\alpha > 0$. Then

$$\begin{aligned} \alpha \cdot |\{x \in \mathbb{R}^n : |f(x)| > \alpha\}| &= \int_{\{x \in \mathbb{R}^n : |f(x)| > \alpha\}} \alpha dy \\ &\le \int_{\{x \in \mathbb{R}^n : |f(x)| > \alpha\}} |f(y)|dy \le \int_{\mathbb{R}^n} |f(y)|dy = \|f\|_1. \end{aligned}$$

The set of measurable functions f on $\mathbb{R}^n$ with

$$\|f\|_{1,w} := \sup_{\alpha>0} \alpha \cdot |\{x \in \mathbb{R}^n : |f(x)| > \alpha\}| < \infty$$

is denoted by $L_{1,w}(\mathbb{R}^n)$ and called *weak* L_1. We have just shown that $L_1(\mathbb{R}^n) \hookrightarrow L_{1,w}(\mathbb{R}^n)$. Let us mention that $\|\cdot\|_{1,w}$ is not a norm, but it still satisfies

$$\begin{aligned}
\|f+g\|_{1,w} &= \sup_{\alpha>0} \alpha \cdot |\{x \in \mathbb{R}^n : |f(x)+g(x)| > \alpha\}| \\
&\le \sup_{\alpha>0} \alpha \Big(|\{x \in \mathbb{R}^n : |f(x)| > \alpha/2\}| + |\{x \in \mathbb{R}^n : |g(x)| > \alpha/2\}| \Big) \\
&= 2 \sup_{\alpha>0} \alpha/2 \cdot \Big(|\{x \in \mathbb{R}^n : |f(x)| > \alpha/2\}| + |\{x \in \mathbb{R}^n : |g(x)| > \alpha/2\}| \Big) \\
&\le 2(\|f\|_{1,w} + \|g\|_{1,w}).
\end{aligned}$$

Finally, we observe, that the function $x \to \dfrac{1}{\|x\|^n} \in L_{1,w}(\mathbb{R}^n) \setminus L_1(\mathbb{R}^n)$.

The main aim of this section is to prove the following theorem.

Theorem 1.2.1. *Let f be a measurable function on $\mathbb{R}^n$. Then*

i) If $f \in L_p(\mathbb{R}^n)$, $1 \le p \le \infty$, then the function Mf is finite almost everywhere.

ii) If $f \in L_1(\mathbb{R}^n)$, then $Mf \in L_{1,w}(\mathbb{R}^n)$ and

$$\|Mf\|_{1,w} \le A\|f\|_1,$$

where A is a constant which depends only on the dimension (i.e. $A = 5^n$ will do).

iii) If $f \in L_p(\mathbb{R}^n)$ with $1 < p \le \infty$, then $Mf \in L_p(\mathbb{R}^n)$ and

$$\|Mf\|_p \le A_p\|f\|_p,$$

where A_p depends only on p and dimension n.

The proof is based on the following *covering lemma.*

Lemma 1.2.2. *Let E be a measurable subset of $\mathbb{R}^n$, which is covered by the union of a family of balls (B^j) with uniformly bounded diameter. Then from this family we can select a disjoint subsequence, $B_1, B_2, B_3, \dots$, such that*

$$\sum_k |B_k| \ge C|E|.$$

Here C is a positive constant that depends only on the dimension n; $C = 5^{-n}$ will do.

Proof. We describe first the choice of $B_1, B_2, \dots$. We choose B_1 so that it is essentially as large as possible, i.e.

$$\operatorname{diam}(B_1) \ge \frac{1}{2} \sup_j \operatorname{diam}(B^j).$$

The choice of B_1 is not unique, but that shall not hurt us.

If $B_1, B_2, \dots, B_k$ were already chosen, we take again B_{k+1} disjoint with $B_1, \dots, B_k$ and again nearly as large as possible, i.e.

$$\operatorname{diam}(B_{k+1}) > \frac{1}{2} \sup\{\operatorname{diam}(B^j) : B^j \text{ disjoint with } B_1, \dots, B_k\}.$$

In this way, we get a sequence $B_1, B_2, \dots, B_k, \dots$ of balls. It can be also finite, if there were no balls B^j disjoint with $B_1, B_2, \dots, B_k$.

If $\sum_k |B_k| = \infty$, then the conclusion of lemma is satisfied and we are done. If $\sum_k |B_k| < \infty$, we argue as follows.

We denote by B_k^* the ball with the same center as B_k and diameter five times as large. We claim that

$$\bigcup_k B_k^* \supset E,$$

which then immediately gives that $|E| \le \sum_k |B_k^*| = 5^n \sum_k |B_k|$.

We shall show that $\bigcup_k B_k^* \supset B^j$ for every j. This is clear if B^j is one of the balls in the preselected sequence. If it is not the case, we obtain $\mathrm{diam}(B_k) \to 0$ (as $\sum_k |B_k| < \infty$) and we choose the first k with $\mathrm{diam}(B_{k+1}) < \frac{1}{2}\mathrm{diam}(B^j)$. That means, that B^j must intersect one of the balls $B_1, \dots, B_k$, say B_{k_0}. Obvious geometric arguments (based on the inequality $\mathrm{diam}(B_{k_0}) \ge 1/2 \cdot \mathrm{diam}(B^j)$) then give that $B^j \subset B_{k_0}^*$. This finishes the proof. $\square$

Proof. (of Theorem 1.2.1). Let $\alpha > 0$ and let us consider $E_\alpha := \{x \in \mathbb{R}^n : Mf(x) > \alpha\}$. From the definition of M we obtain, that for every $x \in E_\alpha$, there is a ball B_x centered at x such that

$$\int_{B_x} |f(y)| dy > \alpha |B_x|.$$

Hence, we get $|B_x| < \|f\|_1/\alpha$ and $\cup_{x \in E_\alpha} B_x \supset E_\alpha$ and the balls $(B_x)_{x \in E_\alpha}$ satisfy the assumptions of Lemma 1.2.2. Using this covering lemma, we get a sequence of disjoint balls $(B_k)_k$, such that

$$\sum_k |B_k| \ge C|E_\alpha|.$$

We therefore obtain

$$\|f\|_1 \ge \int_{\cup_k B_k} |f(y)| dy = \sum_k \int_{B_k} |f| > \alpha \sum_k |B_k| \ge \alpha C |E_\alpha|,$$

which may be rewritten as $\sup_{\alpha > 0} \alpha \cdot |E_\alpha| = \|Mf\|_{1,w} \le \frac{1}{C}\|f\|_1$.

This proves the first assertion of the theorem for $p = 1$ and the second assertion.

We now consider $1 < p < \infty$. The proof follows from the information on the endpoints, i.e. from

$$\|Mf\|_{1,w} \le C\|f\|_1 \quad \text{and} \quad \|Mf\|_\infty \le \|f\|_\infty.$$

Let $\alpha > 0$ and put $f_1(x) := f(x)$ if $|f(x)| > \alpha/2$ and $f_1(x) := 0$ otherwise. Due to $|f(x)| \le |f_1(x)| + \alpha/2$, we have also $Mf(x) \le Mf_1(x) + \alpha/2$ and also

$$\{x \in \mathbb{R}^n : Mf(x) > \alpha\} \subset \{x \in \mathbb{R}^n : Mf_1(x) > \alpha/2\}.$$

Due to the second part of the theorem

$$\begin{aligned}|E_\alpha| &= |\{x \in \mathbb{R}^n : Mf(x) > \alpha\}| \le |\{x \in \mathbb{R}^n : Mf_1(x) > \alpha/2\}| \le \frac{2A}{\alpha}\|f_1\|_1 \\ &= \frac{2A}{\alpha} \int_{\{x \in \mathbb{R}^n : |f(x)| > \alpha/2\}} |f(y)| dy.\end{aligned}$$

We use the information of the size of the level sets of Mf to estimate the L_p-norm of Mf.

$$\begin{aligned}
\|Mf\|_p^p &= \int_{\mathbb{R}^n} Mf(x)^p dx = \int_0^\infty |\{x\in\mathbb{R}^n : Mf(x)^p > \alpha\}| d\alpha \\
&= \int_0^\infty |\{x\in\mathbb{R}^n : Mf(x) > \alpha^{1/p}\}| d\alpha \\
&= p\int_0^\infty \beta^{p-1} |\{x\in\mathbb{R}^n : Mf(x) > \beta\}| d\beta \\
&\le p\int_0^\infty \beta^{p-1} \left(\frac{2A}{\beta}\int_{\{x\in\mathbb{R}^n : |f(x)|>\beta/2\}} |f(y)|dy\right) d\beta \\
&= 2Ap\int_{\mathbb{R}^n} |f(y)| \int_0^{2|f(y)|} \beta^{p-2} d\beta dy = \frac{2Ap}{p-1}\int_{\mathbb{R}^n} |f(y)|\cdot|2f(y)|^{p-1}dy \\
&= \frac{2^pAp}{p-1}\int_{\mathbb{R}^n} |f(y)|^p dy,
\end{aligned}$$

where we used Fubini's theorem and the substitution $\beta := \alpha^{1/p}$ with $d\alpha = p\beta^{p-1}d\beta$. This gives the first and the third statement of the theorem with

$$A_p = 2\left(\frac{5^n p}{p-1}\right)^{1/p}, \quad 1 < p < \infty.$$

□

Corollary 1.2.3. (Lebesgue's differentiation theorem)
Let f be locally integrable on $\mathbb{R}^n$. Then

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

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

Proof. We may cover the set of "bad" points

$$\left\{x\in\mathbb{R}^n : \lim_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy \text{ does not exist or } \lim_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy \neq f(x)\right\}$$

by

$$\bigcup_{k=1}^\infty \left\{x\in\mathbb{R}^n : \limsup_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy - f(x) > \frac{1}{k}\right\}$$

united with

$$\bigcup_{k=1}^\infty \left\{x\in\mathbb{R}^n : \liminf_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy - f(x) < -\frac{1}{k}\right\}.$$

It is therefore enough to show, that each of these sets has measure zero. Let us fix $k \in \mathbb{N}$ and decompose $f = g + h$, where $g \in C(\mathbb{R}^n)$ and $\|h\|_1 \le t$, $t > 0$.

Obviously,

$$\lim_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} g(y)dy = g(x), \quad x \in \mathbb{R}^n,$$

which implies

$$\begin{aligned}&\left\{x\in\mathbb{R}^n:\limsup_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)}f(y)dy-f(x)>\frac{1}{k}\right\}\\&\quad=\left\{x\in\mathbb{R}^n:\limsup_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)}h(y)dy-h(x)>\frac{1}{k}\right\}\\&\quad\subset\left\{x\in\mathbb{R}^n:Mh(x)>\frac{1}{2k}\right\}\cup\left\{x\in\mathbb{R}^n:|h(x)|>\frac{1}{2k}\right\}.\end{aligned}$$

The measure of the first set is smaller than $2kA\|h\|_1$ and the measure of the second is smaller than $2k\|h\|_1$. As $\|h\|_1$ might be chosen arbitrary small, the measure of the original set is zero. The same argument works for lim inf instead of lim sup as well.

□

Theorem 1.2.4. *Let φ be a function which is non-negative, radial, decreasing (as function on $(0,\infty)$) and integrable. Then*

$$\sup_{t>0}|\varphi_t*f(x)|\le\|\varphi\|_1Mf(x),$$

where again $\varphi_t(x)=t^{-n}\varphi(x/t),x\in\mathbb{R}^n$.

Proof. Let in addition φ be a simple function. Then it can be written as

$$\varphi(x)=\sum_{j=1}^k a_j\chi_{B_{r_j}}(x),$$

with $a_j>0$ and $r_j>0$. Then

$$\varphi*f(x)=\sum_{j=1}^k a_j|B_{r_j}|\frac{1}{|B_{r_j}|}\chi_{B_{r_j}}*f(x)\le\|\varphi\|_1Mf(x),$$

since $\|\varphi\|_1=\sum_j a_j|B_{r_j}|$. As any normalized dilation of φ satisfies the same assumptions and has the same integral, it satisfies also the same inequality. Finally, any function satisfying the hypotheses of the Theorem can be approximated monotonically from below by a sequence of simple radial functions. This finishes the proof. □

Last theorem can be easily reformulated as a statement about boundedness of certain sublinear operator. Let φ be as in Theorem 1.2.4 and let Φ be the following operator

$$\Phi(f)(x):=\sup_{t>0}|(\varphi_t*f)(x)|.$$

Then $\Phi:L_1(\mathbb{R}^n)\to L_{1,w}(\mathbb{R}^n)$ and $\Phi:L_p(\mathbb{R}^n)\to L_p(\mathbb{R}^n)$ for $1<p\le\infty$. For example, if $f\in L_1(\mathbb{R}^n)$, we obtain

$$\|\Phi(f)\|_{1,w}\le\|\varphi\|_1\cdot\|Mf\|_{1,w}\le C\|\varphi\|_1\cdot\|f\|_1$$

(and similarly for $1<p\le\infty$).

2 Interpolation

We shall present two basic interpolation theorems, the Riesz-Thorin interpolation theorem and Marcinkiewicz interpolation theorem.

2.1 Marcinkiewicz interpolation theorems

The operator T mapping measurable functions to measurable functions is called sub-linear, if

$$|T(f_0+f_1)(x)| \le |Tf_0(x)| + |Tf_1(x)|,$$
$$|T(\lambda f)(x)| = |\lambda| \cdot |Tf(x)|, \quad \lambda \in \mathbb{C}.$$

Let (X,μ) and (Y,ν) be measure spaces and let T be a sub-linear operator mapping $L_p(X,\mu)$ into a space of measurable functions on (Y,ν). We say that T is strong type (p,q) if it is bounded from $L_p(X,\mu)$ into $L_q(Y,\nu)$. We say, that it is of weak type (p,q), $q<\infty$, if

$$\|Tf\|_{q,w} := \sup_{\lambda>0} \lambda \cdot \nu^{1/q}(\{y \in Y : |Tf(y)| > \lambda\}) \le C\|f\|_p, \quad f \in L_p(X,\mu),$$

i.e. if $T : L_p(X,\mu) \to L_{q,w}(Y,\nu)$.

Theorem 2.1.1. *(Marcinkiewicz interpolation theorem)*
Let (X,μ) and (Y,ν) be measure spaces, $1 \le p_0 < p_1 \le \infty$, and let T be a sublinear operator from $L_{p_0}(X,\mu)+L_{p_1}(X,\mu)$ to the measurable functions on Y that is weak (p_0,p_0) type and weak (p_1,p_1) type. Then T is strong (p,p) for $p_0<p<p_1$.

Proof. Let $\lambda>0$ be given and let $f \in L_p(X,\mu)$. Then we decompose f into $f=f_0+f_1$ with

$$f_0 = f\chi_{\{x:|f(x)|>\lambda\}},$$
$$f_1 = f\chi_{\{x:|f(x)|\le\lambda\}}.$$

The case $p_1=\infty$ appeared implicitly already in the proof of the boundedness of Hardy-Littlewood maximal operator, so we suppose that $p_1<\infty$. Then we have

$$\nu(\{y\in Y : |Tf(y)|>\lambda\}) \le \nu(\{y \in Y : |Tf_0(y)| > \lambda/2\}) + \nu(\{y\in Y : |Tf_1(y)|>\lambda/2\})$$

and

$$\nu(\{y \in Y : |Tf_i(y)| > \lambda/2\}) \le \left(\frac{2A_i}{\lambda}\|f_i\|_{p_i}\right)^{p_i}, \quad i=0,1.$$

We combine them to get

$$\begin{aligned}
\|Tf\|_p^p &= p\int_0^\infty \lambda^{p-1}\nu(\{x : |Tf(x)|>\lambda\})d\lambda \\
&\le p\int_0^\infty \lambda^{p-1-p_0}(2A_0)^{p_0}\int_{x:|f(x)|>\lambda}|f(x)|^{p_0}d\mu d\lambda \\
&\quad + p\int_0^\infty \lambda^{p-1-p_1}(2A_1)^{p_1}\int_{x:|f(x)|\le\lambda}|f(x)|^{p_1}d\mu d\lambda \\
&= p(2A_0)^{p_0}\int_X |f(x)|^{p_0}\int_0^{|f(x)|}\lambda^{p-1-p_0}d\lambda d\mu \\
&\quad + p(2A_1)^{p_1}\int_X |f(x)|^{p_1}\int_{|f(x)|}^\infty \lambda^{p-1-p_1}d\lambda d\mu \\
&= C\|f\|_p^p.
\end{aligned}$$

□

2.2 Riesz-Thorin interpolation theorem

Following theorem belongs also to the classical heart of interpolation theory.

Theorem 2.2.1. *(Riesz-Thorin Interpolation)*
Let (X,μ) and (Y,ν) be two maesure spaces. Let $1 \le p_0, p_1, q_0, q_1 \le \infty$, and for $0 < \theta < 1$ define p and q by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \quad \textit{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

If T is a linear(!) operator from $L_{p_0}(X) + L_{p_1}(X)$ to $L_{q_0}(Y) + L_{q_1}(Y)$, such that

$$\|Tf\|_{q_0} \le M_0 \|f\|_{p_0} \quad \textit{for } f \in L_{p_0}(X)$$

and

$$\|Tf\|_{q_1} \le M_1 \|f\|_{p_1} \quad \textit{for } f \in L_{p_1}(X),$$

then

$$\|Tf\|_q \le M_0^{1-\theta} M_1^{\theta} \|f\|_p \quad \textit{for } f \in L_p(X).$$

Remark 2.2.2. In this version, the theorem only holds for function spaces of complex-valued functions. In the real case, an additional factor 2 is necessary.

Proof. If $p_0 = p_1$, then the theorem collapses to Hölder's inequality. Hence, we consider only $p_0 \ne p_1$ and by symmetry, $p_0 < p_1$. This forces p to be finite. By multiplying μ, ν and/or T with some constants, we may assume that $M_0 = M_1 = 1$. By Hölder's inequality, we get

$$\left| \int_Y (Tf) g d\nu \right| \le \|Tf\|_{q_0} \cdot \|g\|_{q_0'} \le \|f\|_{p_0} \cdot \|g\|_{q_0'}, \quad f \in L_{p_0}(X), \quad g \in L_{q_0'}(Y) \tag{2.1}$$

and

$$\left| \int_Y (Tf) g d\nu \right| \le \|Tf\|_{q_1} \cdot \|g\|_{q_1'} \le \|f\|_{p_1} \cdot \|g\|_{q_1'}, \quad f \in L_{p_1}(X), \quad g \in L_{q_1'}(Y). \tag{2.2}$$

Then we claim that

$$\left| \int_Y (Tf) g d\nu \right| \le \|f\|_p \cdot \|g\|_{q'} \tag{2.3}$$

for all f, g simple functions with finite measure support. To see this, we first normalize $\|f\|_p = \|g\|_{q'} = 1$ and write $f = |f| \operatorname{sgn}(f)$ and $g = |g| \operatorname{sgn}(g)$. Then we define

$$F(s) := \int_Y (T[|f|^{(1-s)p/p_0 + sp/p_1} \operatorname{sgn}(f)])[|g|^{(1-s)q'/q_0' + sq'/q_1'} \operatorname{sgn}(g)] d\nu$$

with $q'/q_0' = q'/q_1' = 1$ if $q_0' = q_1' = q' = \infty$. We observe (by the linearity of T) that F is holomorphic function in s of at most exponential growth. We observe that

i) $F(\theta) = \int_Y (Tf) g d\nu$,

ii) $|F(0 + it)| \le 1$, by (2.1),

iii) $|F(1+it)| \le 1$, by (2.2).

The claim then follows by

Theorem 2.2.3. (Lindelöf's Three-Lines Theorem) *Let $s \mapsto F(s)$ be a holomorphic function on the strip $S := \{\sigma + it : 0 \le \sigma \le 1; t \in \mathbb{R}\}$, which obeys the bound*

$$|F(\sigma + it)| \le A \exp(\exp((\pi - \delta)t))$$

for all $\sigma + it \in S$ and some constants $A, \delta > 0$. Suppose also that $|F(0+it)| \le B_0$ and $|F(1+it)| \le B_1$ for all $t \in \mathbb{R}$. Then we have $|F(\theta + it)| \le B_0^{1-\theta} B_1^{\theta}$ for all $0 \le \theta \le 1$ and $t \in \mathbb{R}$.

To extend (2.3) to general functions f and g, one takes $f \in L_p(X)$ (keeping g simple with finite measure support) by decomposing f into a bounded function and a function of finite measure support, approximating the former in $L_p(X) \cap L_{p_1}(X)$ by simple functions of finite measure support, and approximating the latter in $L_p(X) \cap L_{p_0}(X)$ by simple functions of finite measure support, and taking limits using (2.1), (2.2) to justify the passage to the limit. One can then also allow arbitrary $g \in L_{q'}(Y)$ by using the monotone convergence theorem. The claim now follows from the duality between $L_{q_1}(Y)$ and $L_{q_1'}(Y)$. □

Proof. of Theorem 2.2.3:
We shall assume first that F is bounded on S. Define holomorpic functions

$$G(s) = F(s)(B_0^{1-s}B_1^s)^{-1} \quad \text{and} \quad G_n(s) = G(s)e^{(s^2-1)/n}$$

Since F is bounded on the closed unit strip and $B_0^{1-s}B_1^s$ is bounded from below, we conclude that G is bounded by some constant M on the closed strip. Also, G is bounded by one on its boundary. Since

$$|G_n(x+iy)| \le M e^{-y^2/n} e^{(x^2-1)/n} \le M e^{-y^2/n},$$

we deduce that $G_n(x+iy)$ converges to zero uniformly in $0 \le x \le 1$ as $|y| \to \infty$. Select $y(n) > 0$ such that for $|y| \ge y(n)$, $|G_n(x+iy)| \le 1$ uniformly in $x \in [0,1]$. By the maximum principle we obtain that $|G_n(s)| \le 1$ in the rectangle $[0,1] \times [-y(n), y(n)]$; hence $|G_n(s)| \le 1$ everywhere in the closed strip. Letting $n \to \infty$, we conclude that $|G(s)| \le 1$ in the closed strip.

The case of a general function F is then done in a similar way by considering

$$F(s)(B_0^{1-s}B_1^s)^{-1}G_\varepsilon(s)G_\varepsilon(1-s) \quad \text{and} \quad G_\varepsilon(s) = \exp(\varepsilon i \exp(i[\pi - \delta/2]s + \delta/4)).$$

□

Riesz-Thorin theorem has two straightforward applications. From properties of the Fourier transform $\mathcal{F} : L_1(\mathbb{R}^n) \to L_\infty(\mathbb{R}^n)$ and $\mathcal{F} : L_2(\mathbb{R}^n) \to L_2(\mathbb{R}^n)$, both with norm 1, we deduce the *Hausdorff-Young inequality* $\|\mathcal{F}f\|_{p'} \le \|f\|_p$ for every $1 \le p \le 2$.

The second application deals with convolution operator. Let $g \in L_1(\mathbb{R}^n)$ be a given function and let $\Phi_g(f) = f * g$. The (easy) estimates $\|f * g\|_1 \le \|f\|_1\|g\|_1$ and $\|f * g\|_\infty \le \|f\|_\infty \|g\|_1$ show, that $\Phi_g : L_1(\mathbb{R}^n) \to L_1(\mathbb{R}^n)$ and $\Phi_g : L_\infty(\mathbb{R}^n) \to L_\infty(\mathbb{R}^n)$. By Riesz-Thorin interpolation, we get also $\Phi_g : L_p(\mathbb{R}^n) \to L_p(\mathbb{R}^n)$ for any $1 < p < \infty$, i.e. $\|f * g\|_p \le \|f\|_p \|g\|_1$.

Next, we consider the convolution operator $\Psi_f(g) = f * g$. We have just proven that $\Psi_f : L_1(\mathbb{R}^n) \to L_p(\mathbb{R}^n)$ for any $f \in L_p(\mathbb{R}^n)$. The easy estimate (which follows at once by Hölder's inequality) $\|f * g\|_\infty \leq \|f\|_p \|g\|_{p'}$ for $1 \leq p \leq \infty$ then gives $\Psi_f : L_{p'}(\mathbb{R}^n) \to L_\infty(\mathbb{R}^n)$. By Riesz-Thorin Theorem, we obtain the *Young's inequality*, which states that

$$\|f * g\|_r \leq \|f\|_p \cdot \|g\|_q$$

with $1/r + 1 = 1/p + 1/q$. Indeed, let $1 < q < p'$, then we define $0 < \theta < 1$ by

$$\frac{1}{q} = \frac{1-\theta}{1} + \frac{\theta}{p'}, \quad \text{i.e. } \frac{1}{q} = 1 - \frac{\theta}{p}.$$

Then $\Psi_f : L_r(\mathbb{R}^n) \to L_q(\mathbb{R}^n)$, where

$$\frac{1}{r} = \frac{1-\theta}{p} + \frac{\theta}{\infty} = \frac{1}{p} - \frac{\theta}{p} = \frac{1}{p} + \frac{1}{q} - 1.$$

3 Weighted inequalities

3.1 Calderón-Zygmund decomposition

The aim of this is to present the decomposition method of Calderón and Zygmund. In its simple form, it was already used in the proof of boundedness of the maximal operator M.

Theorem 3.1.1. *Given a function f, which is integrable and non-negative, and given a positive number λ, there exists a sequence*[4] *(Q_j) of disjoint dyadic cubes such that*

i) $f(x) \le \lambda$ *for almost every* $x \notin \bigcup_j Q_j$*;*

ii) $\left|\bigcup_j Q_j\right| \le \frac{1}{\lambda}\|f\|_1$*;*

iii) $\lambda < \frac{1}{|Q_j|}\int_{Q_j} f \le 2^n\lambda$.

Proof. We denote by $\mathcal{Q}_k$ the collection of dyadic cubes with side length 2^{-k}, $k \in \mathbb{Z}$. Furthermore, we define

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

We define also

$$\Omega_k := \{x \in \mathbb{R}^n : E_k f(x) > \lambda \text{ and } E_j f(x) \le \lambda \text{ if } j < k\}$$

That is, $x \in \Omega_k$ if k is the smallest index with $E_k f(x) > \lambda$. Observe, that if $E_k f(x) > \lambda$ for at least one index $k \in \mathbb{Z}$, the integrability of f implies $E_k f(x) \to 0$ for $k \to -\infty$, and the smallest index k with $E_k f(x) > \lambda$ always exists. The sets Ω_k are clearly disjoint and each can be written as the union of cubes in $\mathcal{Q}_k$. Together, these cubes form the system (Q_j).

This gives the third statement of the theorem. The first follows by Lebesgue differentiation theorem: indeed, $E_k f(x) \le \lambda$ for all $k \in \mathbb{Z}$ implies $f(x) \le \lambda$ at almost every such point. The second follows just by

$$\left|\bigcup_j Q_j\right| = \sum_j |Q_j| \le \frac{1}{\lambda}\sum_j \int_{Q_j} f \le \frac{1}{\lambda}\|f\|_1.$$

□

3.2 First inequality

Theorem 3.2.1. *If w is a non-negative, measurable function and $1 < p < \infty$, then there exists a constant C_p such that*

$$\int_{\mathbb{R}^n} [Mf(x)]^p w(x)dx \le C_p \int_{\mathbb{R}^n} |f(x)|^p Mw(x)dx. \tag{3.1}$$

Furthermore,

$$\int_{\{x:Mf(x)>\lambda\}} w(x)dx \le \frac{C}{\lambda}\int_{\mathbb{R}^n} |f(x)|Mw(x)dx. \tag{3.2}$$

[4] possibly finite, or even empty

Proof. We shall show that

$$\|Mf\|_{L_\infty(w)} \le \|f\|_{L_\infty(Mw)} \tag{3.3}$$

and that the weak-type estimate (3.2) holds. The rest then follows by Marcinkiewicz interpolation theorem.

For (3.3), we argue as follows. If $Mw(x) = 0$ for any x, then $w(x) = 0$ a.e. and there is nothing to prove. Therefore, we may assume that $Mw(x) > 0$. Let $a > \|f\|_{L_\infty(Mw)}$. Then

$$\int_{\{x:|f(x)|>a\}} Mw(x)dx = 0,$$

and so $|\{x \in \mathbb{R}^n : |f(x)| > a\}| = 0$ and $|f(x)| \le a$ a.e. Therefore $Mf(x) \le a$ a.e. and $\|Mf\|_{L_\infty(w)} \le a$. This gives (3.3).

To show (3.2), we may assume that f is non-negative and integrable (If $f \in L_1(Mw)$, then $f\chi_{B(0,j)}$ is a monotone sequence of integrable functions converging to f.) Let (Q_j) be the Calderón-Zygmund decomposition of f at height $\lambda > 0$.

Let $x \notin \bigcup_j 2Q_j$, and let Q be any cube centered at x. Let $l(Q)$ denote the side length of Q. Take $k \in \mathbb{Z}$ with $2^{-(k+1)} \le l(Q) < 2^{-k}$. Then Q intersects $m \le 2^n$ dyadic cubes in $\mathcal{Q}_k$, which we denote $R_1, \dots, R_m$. As $x \notin \bigcup_j 2Q_j$, none of the cubes $R_1, \dots, R_m$ is contained in any of the Q_j's. By the construction of the Calderón-Zygmund decomposition, the average of f on each R_i is at most λ. Hence we obtain

$$\frac{1}{|Q|}\int_Q f = \frac{1}{|Q|}\sum_{i=1}^m \int_{Q\cap R_i} f \le \frac{1}{|Q|}\sum_{i=1}^m \frac{2^{-kn}}{|R_i|}\int_{R_i} f \le 2^n m\lambda \le 4^n\lambda.$$

Therefore,

$$\{x \in \mathbb{R}^n : M'f(x) > 4^n\lambda\} \subset \bigcup_j 2Q_j \tag{3.4}$$

and we obtain

$$\begin{aligned}
\int_{\{x:M'f(x)>4^n\lambda\}} w(x)dx &\le \sum_j \int_{2Q_j} w(x)dx \\
&= \sum_j 2^n|Q_j|\frac{1}{|2Q_j|}\int_{2Q_j} w(x)dx \\
&\le \frac{2^n}{\lambda}\sum_j \int_{Q_j} f(y)\Big(\frac{1}{|2Q_j|}\int_{2Q_j} w(x)dx\Big)dy \\
&\le \frac{2^nC}{\lambda}\int_{\mathbb{R}^n} f(y)M''w(y)dy.
\end{aligned}$$

Since $M''w \approx M'w \approx Mw$, we get (3.2). □

Let us observe that if $Mw(x) \le Cw(x)$, then the inequalities (3.1) and (3.2) simplify to boundedness of M on weighted spaces. This will be the starting point of the study of weighted spaces.

3.3 The Muckenhaupt A_p condition

If $1 \le p < \infty$ and w is a non-negative measurable function on $\mathbb{R}^n$, then we denote by

$$L_p(w) = \Big\{f : \|f\|_{L_p(w)} = \Big(\int_{\mathbb{R}^n} |f(x)|^p w(x)dx\Big)^{1/p} < \infty\Big\} \tag{3.5}$$

the *weighted Lebesgue spaces*. To avoid trivialities, it is good to assume that $w > 0$ a.e. and that w is locally integrable.

In this section, we shall denote by M the non-centered cubic maximal operator, earlier denoted by M''.

We are looking for necessary and sufficient conditions on w, such that

$$M : L_p(w) \to L_p(w), \tag{3.6}$$

or in the weak form

$$M : L_p(w) \to L_{p,w}(w). \tag{3.7}$$

Let us assume first, that (3.7) holds, i.e.

$$w(\{x \in \mathbb{R}^n : Mf(x) > \lambda\}) \le \frac{C}{\lambda^p} \int_{\mathbb{R}^n} |f(x)|^p w(x) dx \tag{3.8}$$

for every $\lambda > 0$.

Let $f \ge 0$ and let Q be a cube with $f(Q) = \int_Q f > 0$. Take $0 < \lambda < f(Q)/|Q|$. Then $Q \subset \{x \in \mathbb{R}^n : M(f\chi_Q)(x) > \lambda\}$ and (3.8) implies

$$w(Q) \le \frac{C}{\lambda^p} \int_Q |f(x)|^p w(x) dx. \tag{3.9}$$

As this holds for all the λ's as above, we get

$$w(Q)\Big(\frac{f(Q)}{|Q|}\Big)^p \le C \int_Q |f|^p w. \tag{3.10}$$

If $S \subset Q$ is a measurable subset, we take $f = \chi_S$ and obtain the necessary condition for (3.8),

$$w(Q)\Big(\frac{|S|}{|Q|}\Big)^p \le Cw(S) \quad \text{for all cubes } Q \quad \text{and for all } S \subset Q. \tag{3.11}$$

Let us interupt now with two short conclusions.
1. The weight w is either identically 0 or $w > 0$ a.e. - just consider S to be a bounded set of positive measure with $w = 0$ on S.
2. The weight w is either locally integrable, or identically infinite. If $w(Q) = \infty$, then the same is true for any larger cube, and therefore also for every bounded set S.

We want to further simplify (3.11). We shall distinguish two cases, $p = 1$ and $p > 1$.

Case 1: $p = 1$

In this case, (3.11) becomes

$$\frac{w(Q)}{|Q|} \le C\frac{w(S)}{|S|}.$$

Choosing $S_\varepsilon := \{x \in Q : w(x) \le \varepsilon + \mathrm{essinf}_{y\in Q} w(y)\}$, we get $w(Q)/|Q| \le C(\mathrm{essinf}_{y\in Q} w(y) + \varepsilon)$ for every $\varepsilon > 0$, i.e.

$$\frac{w(Q)}{|Q|} \le C\mathrm{essinf}_{y\in Q} w(y),$$

or equivalently

$$\frac{w(Q)}{|Q|} \le Cw(x), \quad \text{a.e. } x \in Q, \tag{3.12}$$

or equivalently

$$Mw(x) \le Cw(x), \quad \text{a.e. } x \in \mathbb{R}^n. \tag{3.13}$$

This is the so-called A_1 condition. Clearly, (3.13) implies (3.12). The converse follows by considering (the countably many) cubes with rational vertices.

Case 2: $1 < p < \infty$

We put $f = w^{1-p'}\chi_Q$ into (3.10) and obtain

$$w(Q)\Big(\frac{1}{|Q|}\int_Q w^{1-p'}\Big)^p \le C\int_Q w^{1-p'},$$

or equivalently

$$\Big(\frac{1}{|Q|}\int_Q w\Big)\Big(\frac{1}{|Q|}\int_Q w^{1-p'}\Big)^{p-1} \le C. \tag{3.14}$$

As we do not know a-priori that $w^{1-p'}$ is locally integrable, we take first $f = \min(w^{1-p'}, n)$ and pass to the limit $n \to \infty$. Only after that, we *conclude* from (3.14) that $w^{1-p'}$ is really locally integrable. Condition (3.14) was introduced by Muckenhaupt and is called A_p condition (i.e. we write $w \in A_p$ if w satisfies (3.14) for all cubes Q).

Theorem 3.3.1. *Let $1 \le p < \infty$. Then the weak (p,p) inequality*

$$w(\{x \in \mathbb{R}^n : Mf(x) > \lambda\}) \le \frac{C}{\lambda^p}\int_{\mathbb{R}^n} |f(x)|^p w(x)dx, \quad \lambda > 0 \tag{3.15}$$

holds if and only if $w \in A_p$.

Proof. The necessity follows from the discussion just given.

The case $p = 1$ follows directly from Theorem 3.2.1.

If $p > 1$, $w \in A_p$ and a function f is given, then we show (3.10). By Hölder's inequality

$$\begin{aligned}\Big(\frac{1}{|Q|}\int_Q |f|\Big)^p &= \Big(\frac{1}{|Q|}\int_Q |f|w^{1/p}w^{-1/p}\Big)^p\\ &\le \Big(\frac{1}{|Q|}\int_Q |f|^p w\Big)\Big(\frac{1}{|Q|}\int_Q w^{1-p'}\Big)^{p-1}\\ &\le C\Big(\frac{1}{|Q|}\int_Q |f|^p w\Big)\cdot\frac{|Q|}{w(Q)}.\end{aligned}$$

Therefore, also (3.10) is true and (3.11) also follows.

We may again assume that f is non-negative and integrable. Considering the Calderón-Zygmund decomposition of f at height $4^{-n}\lambda$, we get again

$$\{x \in \mathbb{R}^n : Mf(x) > \lambda\} \subset \bigcup_j 3Q_j,$$

as in (3.4), where the 3 is used due to the non-centered maximal operator. Then, putting this all together,

$$\begin{aligned}w(\{x \in \mathbb{R}^n : Mf(x) > \lambda\}) &\le \sum_j w(3Q_j) \le C3^{np}\sum_j w(Q_j)\\ &\le C3^{np}\sum_j\Big(\frac{|Q_j|}{f(Q_j)}\Big)^p\int_{Q_j}|f|^p w \le C3^{np}\Big(\frac{4^n}{\lambda}\Big)^p\int_{\mathbb{R}^n}|f|^p w.\end{aligned}$$

Here, the second inequality follows from (3.11), the third from (3.10) and the fourth from the properties of the Calderón-Zygmund decomposition $(Q_j)_j$. □

Proposition 3.3.2. *i) $A_p \subset A_q$ for $1 \le p < q$.*

ii) $w \in A_p$ if and only if $w^{1-p'} \in A_{p'}$.

iii) If $w_0, w_1 \in A_1$ then $w_0 w_1^{1-p} \in A_p$.

Proof. (i) If $p = 1$ then

$$\Big(\frac{1}{|Q|}\int_Q w^{1-q'}\Big)^{q-1} \le \mathrm{esssup}_{x\in Q} w(x)^{-1} = \Big(\mathrm{essinf}_{x\in Q} w(x)\Big)^{-1} \le C\Big(\frac{w(Q)}{|Q|}\Big)^{-1}.$$

If $p > 1$ then this follows immediately from Hölder's inequality.
(ii) The $A_{p'}$ condition for $w^{1-p'}$ is

$$\Big(\frac{1}{|Q|}\int_Q w^{1-p'}\Big)\Big(\frac{1}{|Q|}\int_Q w^{(1-p')(1-p)}\Big)^{p'-1} \le C,$$

and since $(p'-1)(p-1) = 1$, the left-hand side is the A_p condition raised to the power $p' - 1$.
(iii) We need to prove that

$$\Big(\frac{1}{|Q|}\int_Q w_0 w_1^{1-p}\Big)\Big(\frac{1}{|Q|}\int_Q w_0^{1-p'} w_1\Big)^{p'-1} \le C. \tag{3.16}$$

By the A_1 condition, for a.e. $x \in Q$ and $i = 0, 1$,

$$w_i(x)^{-1} \le \mathrm{esssup}_{x\in Q} w_i(x)^{-1} = \Big(\mathrm{essinf}_{x\in Q} w_i(x)\Big)^{-1} \le C\Big(\frac{w_i(Q)}{|Q|}\Big)^{-1}.$$

We substitute this into (3.16) for the negative powers and get the desired inequality. □

3.4 Strong-type inequalities

Theorem 3.4.1. *Let $1 < p < \infty$. Then M is bounded on $L_p(w)$ if and only if $w \in A_p$.*

Proof. If M is of strong (p, p) type, then it is also of weak (p, p) type, and therefore $w \in A_p$.

Let on the other hand $w \in A_p$. We shall show that there exists $q < p$ with $w \in A_q$. Then M is of weak (q, q) type and M is of strong (∞, ∞) type, as $L_\infty(w) = L_\infty$ ($w(E) = 0$ if and only if $|E| = 0$). The result then follows by Marcinkiewicz interpolation theorem. □

The existence of such a q is a consequence of

Theorem 3.4.2. *(**Reverse Hölder's Inequality**) Let $w \in A_p$, $1 \le p < \infty$. Then there exist constants C and $\varepsilon > 0$ depending only on p and the A_p constant of w, such that for any cube Q*

$$\Big(\frac{1}{|Q|}\int_Q w^{1+\varepsilon}\Big)^{1/(1+\varepsilon)} \le \frac{C}{|Q|}\int_Q w. \tag{3.17}$$

The name comes from the fact that the reverse of (3.17) is a consequence of Hölder's inequality.

Proof. The proof uses the following
Fact: Let $w \in A_p$, $1 \leq p < \infty$. Then for every $0 < \alpha < 1$, there exists $0 < \beta < 1$, such that for every cube Q and $S \subset Q$ with $|S| \leq \alpha |Q|$ also $w(S) \leq \beta w(Q)$ holds.
To prove the fact, just replace S by $Q \setminus S$ in (3.11) to get

$$w(Q)\Big(1 - \frac{|S|}{|Q|}\Big)^p \leq C(w(Q) - w(S)).$$

If $|S| \leq \alpha|Q|$, we get the statement with $\beta = 1 - C^{-1}(1-\alpha)^p$.

Fix a cube Q and form Calderón-Zygmund decompositions of w with respect to Q at heights $w(Q)/|Q| = \lambda_0 < \lambda_1 < \cdots < \lambda_k < \ldots$; we will fix the λ_k's below.

We get a family $\{Q_{k,j}\}$ of cubes with

$$w(x) \leq \lambda_k \text{ if } x \notin \Omega_k = \bigcup_j Q_{k,j},$$

$$\lambda_k < \frac{1}{|Q_{k,j}|}\int_{Q_{k,j}} w \leq 2^n \lambda_k.$$

By construction $\Omega_{k+1} \subset \Omega_k$ and $Q_{k,j_0} \cap \Omega_{k+1}$ is the union of cubes $Q_{k+1,i}$. Therefore

$$|Q_{k,j_0} \cap \Omega_{k+1}| = \sum_i |Q_{k+1,i}| \leq \frac{1}{\lambda_{k+1}} \sum_i \int_{Q_{k+1,i}} w \leq \frac{1}{\lambda_{k+1}} \int_{Q_{k,j_0}} w \leq \frac{2^n \lambda_k}{\lambda_{k+1}} |Q_{k,j_0}|.$$

Fix $0 < \alpha < 1$ and put $\lambda_k := (2^n \alpha^{-1})^k w(Q)/|Q|$, i.e. $2^n \lambda_k / \lambda_{k+1} = \alpha$. Then $|Q_{k,j_0} \cap \Omega_{k+1}| \leq \alpha |Q_{k,j_0}|$ and (by the Fact) $w(Q_{k,j_0} \cap \Omega_{k+1}) \leq \beta w(Q_{k,j_0})$.

We sum on the level k and get $w(\Omega_{k+1}) \leq \beta w(\Omega_k)$, i.e. $w(\Omega_k) \leq \beta^k w(\Omega_0)$ and by the same argument $|\Omega_k| \leq \alpha^k |\Omega_0|$.
Therefore,

$$\begin{aligned}
\frac{1}{|Q|}\int_Q w^{1+\varepsilon} &= \frac{1}{|Q|}\int_{Q\setminus\Omega_0} w^{1+\varepsilon} + \frac{1}{|Q|}\sum_{k=0}^{\infty}\int_{\Omega_k\setminus\Omega_{k+1}} w^{1+\varepsilon} \\
&\leq \lambda_0^{\varepsilon}\frac{w(Q)}{|Q|} + \frac{1}{|Q|}\sum_{k=0}^{\infty}\lambda_{k+1}^{\varepsilon} w(\Omega_k) \\
&\leq \lambda_0^{\varepsilon}\frac{w(Q)}{|Q|} + \frac{1}{|Q|}\sum_{k=0}^{\infty}(2^n\alpha^{-1})^{(k+1)\varepsilon}\lambda_0^{\varepsilon}\beta^k w(\Omega_0).
\end{aligned}$$

Choose $\varepsilon > 0$ with $(2^n\alpha^{-1})^{\varepsilon}\beta < 1$; then the series converges and the last term is bounded by $C\lambda_0^{\varepsilon} w(Q)/|Q|$. Since $\lambda_0 = w(Q)/|Q|$, the proof is finished. ◻

Proposition 3.4.3. *i)* $A_p = \bigcup_{q<p} A_q$, $1 < p < \infty$.

ii) If $w \in A_p, 1 \leq p < \infty$, then there exists $\varepsilon > 0$ such that $w^{1+\varepsilon} \in A_p$.

iii) If $w \in A_p, 1 \leq p < \infty$, then there exists $\delta > 0$ such that given a cube Q and $S \subset Q$,

$$\frac{w(S)}{w(Q)} \leq C\Big(\frac{|S|}{|Q|}\Big)^{\delta}.$$

The last condition is called A_∞. With this notation, *(i)* holds also for $p = \infty$.

Proof. (i) Let $w \in A_p$. Then by Proposition 3.3.2, we get that $w^{1-p'} \in A_{p'}$ and by Reverse Hölder's inequality

$$\Big(\frac{1}{|Q|}\int_Q w^{(1-p')(1+\varepsilon)}\Big)^{1/(1+\varepsilon)} \le \frac{C}{|Q|}\int_Q w^{1-p'}$$

for all cubes Q and some $\varepsilon > 0$. Fix q with $q'-1 = (p'-1)(1+\varepsilon)$. Then $q < p$ and we get

$$\begin{aligned}\Big(\frac{1}{|Q|}\int_Q w\Big)\Big(\frac{1}{|Q|}\int_Q w^{1-q'}\Big)^{q-1} &\le \Big(\frac{1}{|Q|}\int_Q w\Big)\Big(\frac{C}{|Q|}\int_Q w^{1-p'}\Big)^{(1+\varepsilon)(q-1)}\\ &= \Big(\frac{1}{|Q|}\int_Q w\Big)\Big(\frac{C}{|Q|}\int_Q w^{1-p'}\Big)^{p-1} \le C'\end{aligned}$$

by $w \in A_p$ and $(1+\varepsilon)(q-1) = p-1$. This implies that $w \in A_q$.
(ii) If $p > 1$, then we choose $\varepsilon > 0$ small enough, so that both w and $w^{1-p'}$ satisfy the Reverse Hölder's inequality with ε. We get from the A_p condition

$$\Big(\frac{1}{|Q|}\int_Q w^{1+\varepsilon}\Big)\Big(\frac{1}{|Q|}\int_Q w^{(1+\varepsilon)(1-p')}\Big)^{p-1} \le C\Big(\frac{1}{|Q|}\int_Q w\Big)^{1+\varepsilon}\Big(\frac{C'}{|Q|}\int_Q w^{1-p'}\Big)^{(1+\varepsilon)(p-1)} \le C''.$$

If $p = 1$, we get

$$\frac{1}{|Q|}\int_Q w^{1+\varepsilon} \le C\Big(\frac{1}{|Q|}\int_Q w\Big)^{1+\varepsilon} \le C' w(x)^{1+\varepsilon}, \quad \text{for a.e. } x \in Q.$$

(iii) Fix $S \subset Q$ and suppose that w satisfies the Reverse Hölder's Inequality with $\varepsilon > 0$. Then (by Hölder's inequality!)

$$w(S) = \int_Q \chi_S w \le \Big(\int_Q w^{1+\varepsilon}\Big)^{1/(1+\varepsilon)}|S|^{\varepsilon/(1+\varepsilon)} \le C w(Q)\Big(\frac{|S|}{|Q|}\Big)^{\varepsilon/(1+\varepsilon)}.$$

Hence, we may choose $\delta = \varepsilon/(1+\varepsilon) = 1/(1+\varepsilon)'$. $\square$

4 Hilbert transform

One of the main objects of harmonic analysis are singular integrals, from which the most important is the Hilbert transform.

4.1 Hilbert Transform

The *Hilbert transform* of a measurable function f on $\mathbb{R}$ is defined as

$$Hf(x) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(x-y)}{y} dy.$$

As the integral does not converge absolutely, it has to be interpreted in an appropriate limiting sense, which uses its *cancellation property*, i.e.

$$Hf(x) = \lim_{\varepsilon \to 0^+} \frac{1}{\pi} \int_{y:|y|>\varepsilon} \frac{f(x-y)}{y} dy.$$

With this definition, $Hf(x)$ makes sense for all smooth functions, especially for $f \in \mathscr{S}(\mathbb{R})$. [5]

If we define the distribution

$$\text{p.v.}\frac{1}{x}(\varphi) := \lim_{\varepsilon \to 0^+} \int_{x:|x|>\varepsilon} \frac{\varphi(x)}{x} dx, \quad \varphi \in \mathscr{S}(\mathbb{R}),$$

then $Hf := \frac{1}{\pi}\text{p.v.}\frac{1}{x} * f$. This formula suggests that we look for the Fourier transform of Hf. Therefore, we regularize $1/x$. This can be done using complex distributions $\frac{1}{\pi(x \pm i\varepsilon)}$ and letting $\varepsilon \to 0$ or (and that is what we shall do) by defining

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

Obviously, $\lim_{t\to 0} Q_t(x) = \frac{1}{\pi x}$ holds pointwise and, as we shall show below, also in $\mathscr{S}'(\mathbb{R})$. As

$$\begin{aligned}
\mathcal{F}^{-1}(\text{sgn}(x)e^{-a|x|})(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \text{sgn}(x) e^{-a|x|} \cdot e^{ix\xi} dx \\
&= \frac{1}{\sqrt{2\pi}} \left\{ -\int_{-\infty}^{0} e^{x(a+i\xi)} dx + \int_{0}^{\infty} e^{x(-a+i\xi)} dx \right\} \\
&= \frac{1}{\sqrt{2\pi}} \left\{ \frac{-1}{a+i\xi} + \frac{1}{a-i\xi} \right\} = \sqrt{\frac{2}{\pi}} \cdot \frac{i\xi}{a^2+\xi^2},
\end{aligned}$$

we obtain

$$\widehat{Q}_t(\xi) = \frac{-i}{\sqrt{2\pi}} \text{sgn}(\xi) e^{-t|\xi|},$$

we have also $\lim_{t\to 0} \widehat{Q}_t(\xi) = -i\,\text{sgn}(\xi)/\sqrt{2\pi}$. As this convergence is uniform on compact sets, it holds also in $\mathscr{S}'(\mathbb{R})$. Finally, due to the continuity of Fourier transform, we obtain

$$\left(\frac{1}{\pi}\text{p.v.}\frac{1}{x}\right)^{\wedge} = [\lim_{t\to 0} Q_t]^{\wedge} = \lim_{t\to 0} \widehat{Q}_t = \frac{-i\,\text{sgn}(\cdot)}{\sqrt{2\pi}}.$$

[5] Observe, that the restriction to $n = 1$ is both natural and essential for the cancellation property. Furthermore, from now on, we shall denote the Fourier transform of a function f also by the more usual $\widehat{f}$.

Theorem 4.1.1. *In $\mathscr{S}'(\mathbb{R})$,*

$$\lim_{t\to 0} Q_t(x) = \frac{1}{\pi}\text{p.v.}\frac{1}{x}.$$

Proof. For each $\varepsilon > 0$, the functions $\psi_\varepsilon(x) = x^{-1}\chi_{\{y:|y|>\varepsilon\}}(x)$ are bounded and define tempered distributions with $\lim_{\varepsilon\to 0^+}\psi_\varepsilon = \text{p.v.}\frac{1}{x}$. Therefore, it is enough to show that

$$\lim_{t\to 0}\,(\pi Q_t - \psi_t) = 0$$

in $\mathscr{S}'(\mathbb{R})$. This follows by

$$\begin{aligned}
(\pi Q_t - \psi_t)(\varphi) &= \int_{\mathbb{R}} \frac{x\varphi(x)}{t^2+x^2}dx - \int_{x:|x|>t}\frac{\varphi(x)}{x}dx \\
&= \int_{x:|x|<t}\frac{x\varphi(x)}{t^2+x^2}dx + \int_{x:|x|>t}\left(\frac{x}{t^2+x^2}-\frac{1}{x}\right)\varphi(x)dx \\
&= \int_{x:|x|<1}\frac{x\varphi(xt)}{1+x^2}dx - \int_{x:|x|>1}\frac{\varphi(tx)}{x(1+x^2)}dx
\end{aligned}$$

for $\varphi \in \mathscr{S}(\mathbb{R})$. As $t \to 0$, we apply Lebesgue dominated convergence theorem and use the symmetry of the integrands to conclude, that the limit is zero. □

$$\begin{tikzcd}
Q_t \arrow[r, "\mathscr{S}'(\mathbb{R})", "t\to 0"'] \arrow[d, "\mathcal{F}"'] & \frac{1}{\pi}\text{p.v.}\frac{1}{x} \arrow[d, "\mathcal{F}"'] \\
\frac{-i}{\sqrt{2\pi}}\operatorname{sgn}(\xi)e^{-t|\xi|} \arrow[r, "\mathscr{S}'(\mathbb{R})", "t\to 0"'] & \frac{-i\operatorname{sgn}(\cdot)}{\sqrt{2\pi}},
\end{tikzcd}$$

Summarizing, one defines the Hilbert transform Hf for $f \in \mathscr{S}(\mathbb{R}^n)$ by any of these formulas:

$$\begin{aligned}
Hf &= \frac{1}{\pi}\text{p.v.}\frac{1}{x} * f, \\
Hf &= \lim_{t\to 0} Q_t * f, \\
(Hf)^\wedge(\xi) &= -i\operatorname{sgn}(\xi)\hat{f}(\xi).
\end{aligned}$$

Using the third expression, we can extend the definition of H to $L_2(\mathbb{R})$ and it holds

$$\begin{aligned}
&\|Hf\|_2 = \|(Hf)^\wedge\|_2 = \|-i\operatorname{sgn}(\cdot)\hat{f}\|_2 = \|\hat{f}\|_2 = \|f\|_2, \\
&H(Hf) = (-i\operatorname{sgn}(\cdot)(Hf)^\wedge)^\vee = ((-i\operatorname{sgn}(\cdot))^2\hat{f})^\vee = (-\hat{f})^\vee = -f, \\
&\langle Hf, Hg\rangle = \langle f, g\rangle, \quad \text{by polarization}, \\
&\int Hf\cdot g = -\int f\cdot Hg,
\end{aligned}$$

where the last identity follows from

$$\langle Hf, g\rangle = \langle (Hf)^\wedge, \hat{g}\rangle = \langle -i\operatorname{sgn}(\cdot)\hat{f}, \hat{g}\rangle = \langle \hat{f}, \operatorname{sgn}(\cdot)\hat{g}\rangle = \langle \hat{f}, -(Hg)^\wedge\rangle = -\langle f, Hg\rangle$$

and the simple fact that $\overline{Hg} = H\overline{g}$.

Theorem 4.1.2. *For $f \in \mathscr{S}(\mathbb{R})$, the following is true.*

(i) (Kolmogorov) H is weak type $(1,1)$,

$$\|Hf\|_{1,w} \le C\|f\|_1, \quad i.e. \quad |\{x \in \mathbb{R} : |Hf(x)| > \lambda\}| \le \frac{C}{\lambda}\|f\|_1, \quad \lambda > 0, f \in L_1(\mathbb{R}^n).$$

(ii) (M. Riesz) H is of strong type (p,p) for every $1 < p < \infty$, i.e.

$$\|Hf\|_p \le C_p\|f\|_p.$$

Proof. Step 1.: We show the weak type $(1,1)$ by exploiting Theorem 3.1.1. Let f be non-negative and let $\lambda > 0$, then Theorem 3.1.1 gives a sequence of disjoint intervals (I_j), such that

$$f(x) \le \lambda \text{ for a.e. } x \notin \Omega = \bigcup_j I_j,$$
$$|\Omega| \le \frac{1}{\lambda}\|f\|_1,$$
$$\lambda < \frac{1}{|I_j|}\int_{I_j} f \le 2\lambda.$$

Given this decomposition of $\mathbb{R}$, we decompose f into "good" and "bad" part defined by

$$g(x) = \begin{cases} f(x), & x \notin \Omega, \\ \frac{1}{|I_j|}\int_{I_j} f, & x \in I_j, \end{cases} \qquad b(x) = \sum_j b_j(x) = \sum_j \left(f(x) - \frac{1}{|I_j|}\int_{I_j} f \right)\chi_{I_j}(x).$$

Then $g(x) \le 2\lambda$ almost everywhere, and b_j is supported on I_j and has zero integral. Since $Hf = Hg + Hb$, we have

$$|\{x \in \mathbb{R} : |Hf(x)| > \lambda\}| \le |\{x \in \mathbb{R} : |Hg(x)| > \lambda/2\}| + |\{x \in \mathbb{R} : |Hb(x)| > \lambda/2\}|.$$

We estimate the first term using the L_2-boundedness of H by

$$|\{x \in \mathbb{R} : |Hg(x)| > \lambda/2\}| \le \frac{4}{\lambda^2}\int_{\mathbb{R}} |Hg(x)|^2 dx = \frac{4}{\lambda^2}\int_{\mathbb{R}} g(x)^2 dx \le \frac{8}{\lambda}\int_{\mathbb{R}} g(x)dx = \frac{8}{\lambda}\int_{\mathbb{R}} f(x)dx.$$

Let $2I_j$ be the interval with the same center as I_j and twice the length. Let $\Omega^* = \bigcup_j 2I_j$. Then $|\Omega^*| \le 2|\Omega|$ and

$$\begin{aligned} |\{x \in \mathbb{R} : |Hb(x)| > \lambda/2\}| &\le |\Omega^*| + |\{x \notin \Omega^* : |Hb(x)| > \lambda/2\}| \\ &\le \frac{2}{\lambda}\|f\|_1 + \frac{2}{\lambda}\int_{\mathbb{R}\setminus\Omega^*} |Hb(x)|dx. \end{aligned}$$

As $|Hb(x)| \le \sum_j |Hb_j(x)|$, it is enough to show that

$$\sum_j \int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx \le C\|f\|_1.$$

Denote the center of I_j by c_j and use that b_j has zero integral to get

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

The last inequality follows from $|y-c_j| < |I_j|/2$ and $|x-y| > |x-c_j|/2$. The inner integral equals 2, so

$$\sum_j \int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx \le 2 \sum_j \int_{I_j} |b_j(y)|dy \le 4\|f\|_1.$$

Our proof of the weak $(1,1)$ inequality is for non-negative f, but this is sufficient since an arbitrary real function can be decomposed into its positive and negative parts, and a complex function into its real and imaginary parts.

Step 2.: Since H is weak type $(1,1)$ and strong type $(2,2)$, it is also strong type (p,p) for $1 < p < 2$. If $p > 2$, we apply duality, i.e.

$$\begin{aligned}\|Hf\|_p &= \sup\left\{ \left| \int_{\mathbb{R}} Hf \cdot g \right| : \|g\|_{p'} \le 1 \right\} \\ &= \sup\left\{ \left| \int_{\mathbb{R}} f \cdot Hg \right| : \|g\|_{p'} \le 1 \right\} \\ &\le \|f\|_p \cdot \sup\left\{ \|Hg\|_{p'} : \|g\|_{p'} \le 1 \right\} \le C_{p'}\|f\|_p.\end{aligned}$$

□

The strong (p,p) inequality is false if $p=1$ or $p=\infty$; this can easily be seen if we let $f = \chi_{[0,1]}$. Then

$$Hf(x) = \frac{1}{\pi} \log \left| \frac{x}{x-1} \right|,$$

and Hf is neither integrable nor bounded.

4.2 Connection to complex analysis

Let D be the (open) unit disc of the complex plane, i.e. $D = \{z \in \mathbb{C} : |z| < 1\}$ and let $\partial D = \{z \in \mathbb{C} : |z| = 1\}$ be its boundary. If $f \in L_2(\partial D)$ is a real-valued function, then the Cauchy integral

$$f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(\xi)}{\xi - z} d\xi, \quad z \in D,$$

defines a holomorphic function on D. There are two main questions connected with this construction:

i) In which sense are the values of f on ∂D also the boundary values of f defined on the whole D? Or, equivalently, in which sense do the functions $f_r(\varphi) := \mathrm{Re}(f(re^{i\varphi})), \varphi \in [0, 2\pi]$ converge to f if $r \to 1^-$.

ii) How do the imaginary parts of the functions $\varphi \to f(re^{i\varphi})$ look like, do they also converge to some other function $f^\dagger$ on ∂D, and how does $f^\dagger$ depend on f?

The study of the first question leads to Poisson formula and approximations of identity. The study of the second question is closely connected to Hilbert transform. Indeed, let u, v be two real-valued functions of two real variables be defined by

$$(u+iv)(r,\varphi) = f(re^{i\varphi}) = \frac{1}{2\pi i}\int_{\partial D}\frac{f(\xi)}{\xi - re^{i\varphi}}d\xi.$$

Then

$$\begin{aligned}
v(r,\varphi) &= \mathrm{Im}\Big(\frac{1}{2\pi i}\int_0^{2\pi}\frac{f(e^{it})}{e^{it}-re^{i\varphi}}ie^{it}dt\Big) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\mathrm{Im}\Big(\frac{e^{it}}{e^{it}-re^{i\varphi}}\Big)dt\\
&= \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\mathrm{Im}\Big(\frac{1}{1-re^{i(\varphi-t)}}\Big)dt\\
&= \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\mathrm{Im}\Big(\frac{1}{1-r\cos(\varphi-t)-ir\sin(\varphi-t)}\Big)dt\\
&= \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\frac{r\sin(\varphi-t)}{1+r^2-2r\cos(\varphi-t)}dt.
\end{aligned}$$

If we now let $r \to 1^+$, we obtain the function

$$\begin{aligned}
\varphi \to \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\frac{\sin(\varphi-t)}{2(1-\cos(\varphi-t))}dt &= \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\frac{2\sin((\varphi-t)/2)\cos((\varphi-t)/2)}{4\sin^2((\varphi-t)/2)}dt\\
= \frac{1}{4\pi}\int_0^{2\pi} f(e^{it})\cot\big(\frac{\varphi-t}{2}\big)dt.
\end{aligned}$$

Considering this function as a periodic variable on $[0,2\pi]$, we obtain that

$$f^\dagger(\varphi) = \frac{1}{4\pi}\int_0^{2\pi} f(t)\cot\big(\frac{\varphi-t}{2}\big)dt.$$

Finally, using Taylor's series, one obtains that $\cot(\alpha)$ behaves like $1/\alpha$ for α close to zero. The answer to the question (ii) posed above is therefore, that $f^\dagger$ is (up to higher order terms) the Hilbert transform of f.

4.3 Connection to Fourier series

We use the complex notation of Fourier series, i.e.

$$f \approx t \to \sum_{j\in\mathbb{Z}}\hat{f}(j)e^{ijt}, \quad \hat{f}(j) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ijt}dt.$$

Let $\mathrm{sgn}(x)$ be defined as 1 for $x > 0$, as 0 or $x = 0$ and as -1 for $x < 0$. Then the Hilbert transform on $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ is defined as

$$Hf = \sum_{j\in\mathbb{Z}} h_j\hat{f}(j)e^{ijt}, \quad h_j = -i\,\mathrm{sgn}(j).$$

Naturally, we define the operator of the Nth partial sum as

$$S_N(f) = \sum_{|j|\le N}\hat{f}(j)e^{ijt} = \sum_{j\in\mathbb{Z}}\chi_{[-N,N]}(j)\hat{f}(j)e^{ijt}$$

and the projections

$$P_N(f) = \hat{f}(N)e^{iNt}.$$

We combine the formulas

$$\chi_{[-N,N]}(j) = \frac{1}{2}\big[\mathrm{sgn}(j+N) - \mathrm{sgn}(j-N)\big] + \frac{1}{2}\big[\chi_{\{N\}}(j) + \chi_{\{-N\}}(j)\big]$$

with

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \mathrm{sgn}(j+N)\hat{f}(j)e^{ijt} &= \sum_{j\in\mathbb{Z}} \mathrm{sgn}(j)\hat{f}(j-N)e^{i(j-N)t} = ie^{-iNt}\sum_{j\in\mathbb{Z}}(-i)\,\mathrm{sgn}(j)\hat{f}(j-N)e^{ijt}\\ &= ie^{-iNt}\sum_{j\in\mathbb{Z}} h_j(f\cdot e^{iN\cdot})\hat{\ }(j)e^{ijt} = ie^{-iNt}H[e^{iN\cdot}\cdot f](t)\end{aligned}$$

to obtain a reformulation of $S_N f$ into

$$S_N f = ie^{-iN\cdot}H[e^{iN\cdot}f] - ie^{iN\cdot}H[e^{-iN\cdot}f] + \frac{1}{2}[P_N f + P_{-N}f].$$

If now $f \in L_p(\mathbb{T})$, then the boundedness of H on $L_p(\mathbb{T})$ implies that also $S_N f \in L_p(\mathbb{T})$ for $1 < p < \infty$. Actually it follows that $\|S_N f\|_p \le c_p\|f\|_p$ with c_p independent on N and that $S_N f \to f$ for all trigonometric polynomials, which are actually dense in $L_p(\mathbb{T})$. It follows that $\|S_N f - f\|_p \to 0$ for all $f \in L_p(\mathbb{T})$. Indeed, let $f \in L_p(\mathbb{T})$ and let g be a trigonometric polynomial with $\|f-g\|_p \le \varepsilon$. Let N be the degree of g. Then it follows for every $M \ge N$ that

$$\|S_M f - f\|_p = \|S_M f - S_M g + g - f\|_p \le \|S_M(f-g)\|_p + \|f-g\|_p \le (c_p+1)\varepsilon.$$

4.4 Connection to maximal operator

Let us state (without proof, cf. Fourier Analysis of Javier Duoandikoetxea, page 56) the following result, called Cotlar's inequality.

Lemma 4.1. *If $f \in \mathscr{S}(\mathbb{R}^n)$ then $H^*f(x) \le M(Hf)(x) + CMf(x)$, where $H^*f(x) = \sup_{\varepsilon>0}|H_\varepsilon f(x)|$ and $H_\varepsilon f(x) = \frac{1}{\pi}\int_{|y|>\varepsilon}\frac{f(x-y)}{y}dy$.*

5 H^1 and BMO

5.1 H^1 - atomic

Let $\varphi \in \mathscr{S}(\mathbb{R})$. Then $H\varphi \in L_1(\mathbb{R})$ if, and only if, $\int_{\mathbb{R}} \varphi = 0$. We show only one direction, the other resembles the proof of Theorem 5.1.3. Let $\varphi \in L_1(\mathbb{R})$ and $H\varphi \in L_1(\mathbb{R})$. Then, by basic properties of the Fourier transform, $\hat{\varphi}$ and $(H\varphi)^\wedge = -i\,\mathrm{sgn}(\cdot)\hat{\varphi}$ are both continuous. Obviously, this is only possible if $\hat{\varphi}(0) = 0$.

We shall define a subspace of $L_1(\mathbb{R})$, which will be mapped by H into $L_1(\mathbb{R})$. As this space plays the same role also for other singular integrals on $\mathbb{R}^n$, we define the space for general $n \in \mathbb{N}$.

Definition 5.1.1. An *atom* is a complexed-valued function defined on $\mathbb{R}^n$, which is supported on a cube Q and satisfies

$$\int_Q a(x)dx = 0 \quad \text{and} \quad \|a\|_\infty \le \frac{1}{|Q|}.$$

Observe, that this implies that $\|a\|_1 \le |Q| \cdot \|a\|_\infty \le 1$.

Definition 5.1.2. The atomic space $H^1_{\mathrm{at}}(\mathbb{R}^n)$ is defined by

$$H^1_{\mathrm{at}}(\mathbb{R}^n) = \left\{ \sum_j \lambda_j a_j : a_j \text{ atoms}, \lambda_j \in \mathbb{C}, \sum_j |\lambda_j| < \infty \right\}$$

and normed by

$$\|f|H^1_{\mathrm{at}}(\mathbb{R}^n)\| = \inf \left\{ \sum_j |\lambda_j| : f = \sum_j \lambda_j a_j \right\}.$$

We state (without the quite easy proof), that this expression is really a norm and that $H^1_{\mathrm{at}}(\mathbb{R}^n)$ is really a Banach space. Furthermore, the *atomic decomposition* converges in $L_1(\mathbb{R}^n)$ and $H^1_{\mathrm{at}}(\mathbb{R}^n)$ is a subspace of $L_1(\mathbb{R}^n)$. Both these statements follow easily from

$$\|\sum_j \lambda_j a_j\|_1 \le \sum_j |\lambda_j| \cdot \|a_j\|_1 \le \sum_j |\lambda_j|.$$

Theorem 5.1.3. *Let $n = 1$.*
(i) There exists a constant $C > 0$, such that for every atom a,

$$\|Ha\|_1 \le C.$$

(ii) $H : H^1_{\mathrm{at}}(\mathbb{R}) \to L_1(\mathbb{R})$.

Proof. (i) Since $a \in L_2(\mathbb{R})$, Ha is well defined and we get for Q^* co-centric with Q and length 2 times larger

$$\int_{Q^*} |Ha(x)|dx \le |Q^*|^{1/2} \left(\int_{Q^*} |Ha(x)|^2 dx \right)^{1/2} \le C|Q|^{1/2} \left(\int_Q |a(x)|^2 dx \right)^{1/2} \le C.$$

Using that a has zero average, we get for c_Q the center of Q and of Q^*

$$\begin{aligned}
\int_{\mathbb{R}\setminus Q^*} |Ha(x)|dx &= \frac{1}{\pi} \int_{\mathbb{R}\setminus Q^*} \left| \int_Q \frac{a(y)}{x-y} dy \right| dx \\
&\le \int_{\mathbb{R}\setminus Q^*} \left| \int_Q \left[\frac{1}{x-y} - \frac{1}{x-c_Q} \right] a(y)dy \right| dx \\
&\le \int_Q \int_{\mathbb{R}\setminus Q^*} \left| \frac{1}{x-y} - \frac{1}{x-c_Q} \right| dx \cdot |a(y)|dy \le C.
\end{aligned}$$

(ii) Follows directly from (i) and the definition of $H^1_{\mathrm{at}}(\mathbb{R})$. □

5.2 *BMO*

Hilbert transform acts rather badly on $L_\infty(\mathbb{R})$. Not only is H unbounded on $L_\infty(\mathbb{R})$, it can not be easily defined on a dense subset of $L_\infty(\mathbb{R})$. The definition

$$Hf(y) = \frac{1}{\pi}\int_{\mathbb{R}} \frac{f(x)}{y-x}dx$$

runs for $f \in L_\infty(\mathbb{R})$ into troubles for x near y and x near infinity. If we look onto differences, the situation changes to

$$Hf(y) - Hf(y') = \frac{1}{\pi}\int_{\mathbb{R}} f(x)\left(\frac{1}{y-x} - \frac{1}{y'-x}\right)dx.$$

This improves the situation for x near infinity, as $1/(x-y) - 1/(x-y') = O(1/x^2)$ in this case.

We say that f and g are equivalent modulo a constant if $f(x) = g(x) + C$ for some (complex) constant C and almost every $x \in \mathbb{R}$. Given $f \in L_\infty(\mathbb{R})$ and $y \in \mathbb{R}$, then we take an open interval $B \subset \mathbb{R}$ with center at zero containing y. Then $f\chi_B \in L_2(\mathbb{R})$ and we define $Hf(y)$ to be

$$Hf(y) := H(f\chi_B)(y) + \frac{1}{\pi}\int_{\mathbb{R}\setminus B} f(x)\left(\frac{1}{y-x} + \frac{1}{x}\right)dx. \tag{5.1}$$

The first term is defined by the L_2 definition of H, the integral in the second term converges absolutely. The definition depends on B, but choosing another interval $B' \supset B$ with the center at the origin leads to a difference

$$\begin{aligned}
&H(f\chi_B)(y) + \frac{1}{\pi}\int_{\mathbb{R}\setminus B} f(x)\left(\frac{1}{y-x} + \frac{1}{x}\right)dx - H(f\chi_{B'})(y) - \frac{1}{\pi}\int_{\mathbb{R}\setminus B'} f(x)\left(\frac{1}{y-x} + \frac{1}{x}\right)dx\\
&= H(f\chi_B - f\chi_{B'})(y) + \frac{1}{\pi}\int_{B'\setminus B} f(x)\left(\frac{1}{y-x} + \frac{1}{x}\right)dx\\
&= -H(f\chi_{B'\setminus B})(y) + \frac{1}{\pi}\int_{B'\setminus B} f(x)\left(\frac{1}{y-x} + \frac{1}{x}\right)dx = \frac{1}{\pi}\int_{B'\setminus B}\frac{f(x)}{x}dx,
\end{aligned}$$

which does not depend on y. This defines Hf modulo constant for $f \in L_\infty(\mathbb{R})$. Of course, such a definition does not allow to measure Hf in the usual norms, as L_p. Instead of that, we need a space of functions defined modulo constants.

Definition 5.2.1. (Bounded mean oscillation). Let $f : \mathbb{R}^n \to \mathbb{C}$ be a function defined modulo a constant. The *BMO* (or Bounded Mean Oscillation) norm of f is defined

$$\|f|BMO(\mathbb{R}^n)\| := \sup_B \frac{1}{|B|}\int_B \left| f - \frac{1}{|B|}\int_B f\right|$$

where B ranges over all balls. Note that

$$\int_B \left| f - \frac{1}{|B|}\int_B f\right| = \int_B \left| f(y) - \frac{1}{|B|}\int_B f(x)dx\right| dy = \frac{1}{|B|}\int_B \left|\int_B (f(y)-f(x))dx\right| dy$$

if one shifts f by a constant, the BMO norm is unchanged, so this norm is well-defined for functions defined modulo constants. We denote by $BMO(\mathbb{R}^n)$ the space of all functions with finite BMO norm.

Example 5.2.2. Let $f(x) = \operatorname{sgn}(x)$. Let $|y| < a$. We take $B = (-a, a)$ and apply the definition of Hf as presented above. This gives (for $y > 0$)

$$\begin{aligned}
\pi Hf(y) &= \text{p.v.} \int_{-a}^{a} \frac{\operatorname{sgn}(x)}{y-x} dx + \int_{(-\infty,-a)\cup(a,\infty)} \operatorname{sgn}(x) \left(\frac{1}{y-x} + \frac{1}{x} \right) dx \\
&= \lim_{\varepsilon \to 0^+} \left(\int_{(-a,a)\setminus(y-\varepsilon,y+\varepsilon)} \frac{\operatorname{sgn}(x)}{y-x} dx \right) \underbrace{- \int_{-\infty}^{-a} \left(\frac{1}{x} - \frac{1}{x-y} \right) dx}_{-\ln \frac{a}{a+y}} + \underbrace{\int_{a}^{\infty} \left(\frac{1}{x} - \frac{1}{x-y} \right) dx}_{=\ln \frac{x}{x-y} \Big|_{x=a}^{x=\infty} = -\ln \frac{a}{a-y}} \\
&= \lim_{\varepsilon \to 0^+} \left(\int_{-a}^{0} \frac{1}{x-y} dx - \int_{0}^{y-\varepsilon} \frac{1}{x-y} dx - \int_{y+\varepsilon}^{a} \frac{1}{x-y} \right) - \ln \frac{a}{a+y} - \ln \frac{a}{a-y} \\
&= \lim_{\varepsilon \to 0^+} \left(\ln \frac{y}{a+y} + \ln \frac{y}{\varepsilon} - \ln \frac{a-y}{\varepsilon} \right) - \ln \frac{a}{a+y} - \ln \frac{a}{a-y} \\
&= 2 \ln y - 2 \ln a.
\end{aligned}$$

If $y < 0$, similar calculation applies as well. Hence, $\pi Hf(y) = 2 \ln |y| - 2 \ln a$. Hence, ignoring the constant,

$$H(\operatorname{sgn} x) = \frac{2}{\pi} \ln |x|.$$

Let us observe, that

$$\frac{1}{|B|} \int_B \left| f - \frac{1}{|B|} \int_B f \right| \approx \inf_{c \in \mathbb{R}} \frac{1}{|B|} \int_B |f - c|$$

holds for every ball B with universal constants. Indeed, the left-hand side is obviously larger than the right hand side. On the other hand, we get as well

$$\begin{aligned}
\frac{1}{|B|} \int_B \left| f - \frac{1}{|B|} \int_B f \right| &\le \frac{1}{|B|} \int_B |f - c| + \frac{1}{|B|} \int_B \left| c - \frac{1}{|B|} \int_B f \right| \\
&= \frac{1}{|B|} \int_B |f - c| + \frac{1}{|B|} \int_B \left| \frac{1}{|B|} \int_B (c - f) \right| \\
&\le \frac{1}{|B|} \int_B |f - c| + \frac{1}{|B|} \int_B \left(\frac{1}{|B|} \int_B |c - f| \right) \\
&= \frac{2}{|B|} \int_B |f - c| .
\end{aligned}$$

5.3 Connection to Hilbert transform

Theorem 5.3.1. *(H maps $L_\infty(\mathbb{R})$ into $BMO(\mathbb{R})$)*
Let $f \in L_\infty(\mathbb{R})$. Then

$$\|Hf|BMO(\mathbb{R})\| \lesssim \|f\|_\infty.$$

Proof. Due to the observation above, it is enough to show that for every ball B, there is a constant $c = c(B)$ such that

$$\frac{1}{|B|} \int_B |Hf - c| \lesssim \|f\|_\infty.$$

Similarly to (5.1), we define $Hf(x)$ for $x \in B$ to be

$$Hf(x) = H(f\chi_{2B})(x) + H(f\chi_{\mathbb{R}\setminus 2B})(x) = H(f\chi_{2B})(x) + \frac{1}{\pi} \int_{y: y \notin 2B} f(y) \left(\frac{1}{x-y} - \frac{1}{\gamma - y} \right) dy,$$

where $2B$ is a ball with the same center as B, but twice the radius and γ is the center of B. With this definition of Hf, we can actually take $c = 0$.

First we get

$$\begin{aligned}\frac{1}{|B|}\int_B |H(f\chi_{2B})| &\leq \frac{1}{|B|}\left(\int_B |H(f\chi_{2B})|^2\right)^{1/2} \cdot \left(\int_B 1\right)^{1/2}\\ &\leq \frac{1}{\sqrt{|B|}}\left(\int_{2B}|f|^2\right)^{1/2} \leq \frac{\|f\|_\infty}{\sqrt{|B|}}\sqrt{|2B|} \lesssim \|f\|_\infty.\end{aligned}$$

This deals with the “local” part of Hf. For the “global” part, observe that for $x \in B$ we have

$$\pi H(f\chi_{\mathbb{R}\setminus 2B})(x) = \int_{y:y\notin 2B} f(y)\left(\frac{1}{x-y} - \frac{1}{\gamma - y}\right)dy$$

and

$$\begin{aligned}&\frac{1}{|B|}\int_B\left|\int_{\mathbb{R}\setminus 2B} f(y)\left(\frac{1}{x-y} - \frac{1}{\gamma-y}\right)dy\right|dx\\ &\leq \frac{1}{|B|}\int_B\int_{\mathbb{R}\setminus 2B}\left|f(y)\left(\frac{1}{x-y} - \frac{1}{\gamma-y}\right)\right|dydx\\ &\leq \frac{\|f\|_\infty}{|B|}\int_B\int_{\mathbb{R}\setminus 2B}\left|\frac{1}{x-y} - \frac{1}{\gamma-y}\right|dydx.\end{aligned}$$

By shifting, we may assume, that $\gamma = 0, B = (-a, a)$ and $2B = (-2a, 2a)$. We estimate

$$\frac{1}{2a}\int_0^a\int_{2a}^\infty \frac{x}{|y(x-y)|}dydx \leq \frac{1}{2a}\int_0^a xdx\int_{2a}^\infty \frac{1}{|y(y-a)|}dy \leq \frac{a}{4}\int_{2a}^\infty \frac{1}{(y-a)^2}dy \leq c$$

(and similarly for the remaining parts). Altogether, this gives

$$\frac{1}{|B|}\int_B |H(f\chi_{\mathbb{R}\setminus 2B})| \lesssim \|f\|_\infty.$$

Adding the two facts, we obtain the claim. ☐

5.4 Interpolation with BMO and good-λ-inequality

When proving that the Hardy-Littlewood maximal operator M is bounded on $L_p(\mathbb{R}^n), 1 < p \leq \infty$, we used interpolation. As M is not bounded on $L_1(\mathbb{R}^n)$, we needed to replace this by weak boundedness of M. This was still sufficient for the interpolation argument (as described by the Marcinkiewicz theorem). Hilbert transform is not bounded on $L_\infty(\mathbb{R}^n)$, but in the same spirit, we can hope that for the interpolation to work, this could be replaced by weaker information at the endpoint. This is indeed the case.

Theorem 5.4.1. *Let T be a linear operator which is bounded on $L_{p_0}(\mathbb{R}^n)$ for some $1 < p_0 < \infty$, and bounded from $L_\infty(\mathbb{R}^n)$ into $BMO(\mathbb{R}^n)$. Then T is bounded on $L_p(\mathbb{R}^n)$ for all $p_0 < p < \infty$.*

The proof of this statement is based on the so-called *good-λ inequality.* We use the notation M_d for the dyadic maximal operator and

$$M^\# f(x) = \sup_{Q\ni x}\frac{1}{|Q|}\int_Q |f - f_Q|,$$

where $f_Q = \frac{1}{|Q|}\int_Q f$ denotes the average of f over Q.

Lemma 5.4.2. *If $f \in L_{p_0}(\mathbb{R}^n)$ for some $p_0, 1 \le p_0 < \infty$, then*

$$|\{x \in \mathbb{R}^n : M_d f(x) > 2\lambda, M^{\#} f(x) \le \gamma\lambda\}| \le 2^n \gamma |\{x \in \mathbb{R}^n : M_d f(x) > \lambda\}|.$$

for all $\gamma > 0$ and $\lambda > 0$.

Proof. We may assume that f is non-negative and that $\gamma > 0$ and $\lambda > 0$ are fixed. We form the Calderón-Zygmund decomposition of f at height λ. Then the set $\{x \in \mathbb{R}^n : M_d f(x) > \lambda\}$ can be written as the union of disjoint, maximal dyadic cubes. Let Q be one of these cubes. Then it is enough to prove that

$$|\{x \in Q : M_d f(x) > 2\lambda, M^{\#} f(x) \le \gamma\lambda\}| \le 2^n \gamma |Q|. \tag{5.2}$$

Let $\tilde{Q}$ be the dyadic cube, which contains Q whose sides are twice that long. Since Q was maximal, $f_{\tilde{Q}} \le \lambda$. If $x \in Q$ and $M_d f(x) > 2\lambda$, then also $M_d(f\chi_Q)(x) > 2\lambda$. Hence, for such x's,

$$M_d((f - f_{\tilde{Q}})\chi_Q)(x) \ge M_d(f\chi_Q)(x) - f_{\tilde{Q}} > \lambda.$$

By the weak $(1,1)$ inequality for M_d (which actually holds with norm 1, although we did not prove that), we get

$$\begin{aligned} |\{x \in Q : M_d((f - f_{\tilde{Q}})\chi_Q)(x) > \lambda\}| &\le \frac{1}{\lambda} \int_Q |f(x) - f_{\tilde{Q}}| dx \\ &\le \frac{2^n |Q|}{\lambda} \cdot \frac{1}{|\tilde{Q}|} \int_{\tilde{Q}} |f(x) - f_{\tilde{Q}}| dx \le \frac{2^n |Q|}{\lambda} \inf_{x \in Q} M^{\#} f(x). \end{aligned}$$

If the left-hand side of (5.2) is zero, there is nothing to prove. Otherwise there is an $x \in Q$ with $M^{\#} f(x) \le \gamma\lambda$ and (5.2) follows. □

Of course, $M_d f(x) \le C M^{\#} f(x)$ is not true in general. But it holds at least in L_p-norm.

Lemma 5.4.3. *If $1 < p_0 \le p < \infty$ and $f \in L_{p_0}(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n} M_d f(x)^p \le C \int_{\mathbb{R}^n} M^{\#} f(x)^p dx. \tag{5.3}$$

Proof. For $N > 0$, we get

$$\begin{aligned} I_N &= \int_0^N p\lambda^{p-1} |\{x \in \mathbb{R}^n : M_d f(x) > \lambda\}| d\lambda = 2^p \int_0^{N/2} p\lambda^{p-1} |\{x \in \mathbb{R}^n : M_d f(x) > 2\lambda\}| d\lambda \\ &\le 2^p \int_0^{N/2} p\lambda^{p-1} \Big(|\{x \in \mathbb{R}^n : M_d f(x) > 2\lambda, M^{\#} f(x) \le \gamma\lambda\}| + |\{x \in \mathbb{R}^n : M^{\#} f(x) > \gamma\lambda\}| \Big) d\lambda \\ &\le 2^{p+n} \gamma I_N + \frac{2^p}{\gamma^p} \int_0^{\gamma N/2} p\lambda^{p-1} |\{x \in \mathbb{R}^n : M^{\#} f(x) > \lambda\}| d\lambda. \end{aligned}$$

Now we choose $\gamma > 0$ such that $2^{p+n}\gamma = 1/2$, and get

$$I_N \le \frac{2^{p+1}}{\gamma^p} \int_0^{\gamma N/2} p\lambda^{p-1} |\{x \in \mathbb{R}^n : M^{\#} f(x) > \lambda\}| d\lambda.$$

This step is only justified, if $I_N < \infty$. This follows from the fact, that $f \in L_{p_0}(\mathbb{R}^n)$ implies $M_d f \in L_{p_0}(\mathbb{R}^n)$ and

$$I_N \le \frac{p}{p_0} N^{p-p_0} \int_0^N p_0 \lambda^{p_0-1} |\{x \in \mathbb{R}^n : M_d f(x) > \lambda\}| d\lambda < \infty.$$

If the right hand side of (5.3) is infinite, there is nothing to prove. If it is finite, we let $N \to \infty$ and (5.3) follows. □

Proof. of Theorem 5.4.1: The composition $M^{\#} \circ T$ is a sublinear operator. It is bounded on $L_{p_0}(\mathbb{R}^n)$, and on $L_\infty(\mathbb{R}^n)$, as

$$\|M^{\#}(Tf)\|_\infty = \|Tf|BMO(\mathbb{R}^n)\| \le C\|f\|_\infty.$$

By Marcinkiewicz theorem, it is bounded on $L_p(\mathbb{R}^n)$, $p_0 < p < \infty$.

Now let $f \in L_p(\mathbb{R}^n)$ with compact support. Then $f \in L_{p_0}(\mathbb{R}^n)$ and so $Tf \in L_{p_0}(\mathbb{R}^n)$. We apply Lemma 5.4.3 to Tf and due to $|Tf(x)| \le M_d(Tf)(x)$ a.e., we get

$$\int_{\mathbb{R}^n} |Tf(x)|^p dx \le \int_{\mathbb{R}^n} M_d(Tf)(x)^p dx \le C\int_{\mathbb{R}^n} [M^{\#}(Tf)(x)]^p dx \le C\int_{\mathbb{R}^n} |f(x)|^p dx.$$

□

5.5 John-Nirenberg inequality

On the example of $\log(1/|x|)$ we have seen that $BMO(\mathbb{R}^n)$ contains also unbounded functions. On the interval $(-a,a)$, its average is $1-\log a$ and for $\lambda > 1$, the set of x's with

$$|\log(1/|x|) - (1-\log a)| > \lambda$$

has measure $2ae^{-\lambda-1}$, i.e. is exponentially small. This is the case for all unbounded functions from $BMO(\mathbb{R}^n)$.

Theorem 5.5.1. *(John-Nirenberg inequality)*
Let $f \in BMO(\mathbb{R}^n)$. Then there exist constants C_1 and C_2, depending only on the dimension, such that given any cube Q in $\mathbb{R}^n$ and any $\lambda > 0$,

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \le C_1 \exp(-C_2\lambda/\|f|BMO(\mathbb{R}^n)\|)|Q|.$$

Proof. As the inequality is homogeneous, we may assume that $\|f|BMO(\mathbb{R}^n)\| = 1$, hence

$$\frac{1}{|Q|}\int_Q |f(x) - f_Q| dx \le 1.$$

We form the Calderón-Zygmund decomposition of $|f - f_Q|$ with respect to Q at height 2. This gives us a family of cubes $\{Q_{1,j}\}$ such that

$$2 < \frac{1}{|Q_{1,j}|}\int_{Q_{1,j}} |f(x) - f_Q| dx \le 2^{n+1}$$

and $|f(x) - f_Q| \le 2$ if $x \notin \bigcup_j Q_{1,j}$. In particular,

$$\sum_j |Q_{1,j}| \le \frac{1}{2}\int_Q |f(x) - f_Q| dx \le \frac{1}{2}|Q|,$$

and

$$|f_{Q_{1,j}} - f_Q| = \left|\frac{1}{|Q_{1,j}|}\int_{Q_{1,j}} (f(x) - f_Q) dx\right| \le 2^{n+1}.$$

On each cube $Q_{1,j}$ we form the Calderón-Zygmund decomposition of $|f - f_{Q_{1,j}}|$ at height 2. We obtain a family of cubes $\{Q_{1,j,k}\}$ with

$$|f_{Q_{1,j,k}} - f_{Q_{1,j}}| \le 2^{n+1},$$
$$|f(x) - f_{Q_{1,j}}| \le 2 \text{ if } x \in Q_{1,j} \setminus \bigcup_k Q_{1,j,k},$$
$$\sum_k |Q_{1,j,k}| \le \frac{1}{2}|Q_{1,j}|.$$

Collect all the cubes $\{Q_{1,j,k}\}$ into one sequence $\{Q_{2,j}\}$. Then we have $\sum_j |Q_{2,j}| \leq 1/4|Q|$ and if $x \notin \bigcup_j Q_{2,j}$

$$|f(x) - f_Q| \leq |f(x) - f_{Q_{1,j}}| + |f_{Q_{1,j}} - f_Q| \leq 2 + 2^{n+1} \leq 2 \cdot 2^{n+1}.$$

Repeat this process indefinitely. We get for each N a family of cubes $\{Q_{N,j}\}$ with

$$|f(x) - f_Q| \leq N \cdot 2^{n+1}, \quad \text{if } x \notin \bigcup_j Q_{N,j}$$

and $\sum_j |Q_{N,j}| \leq 2^{-N}|Q|$. Fix $\lambda > 2^{n+1}$ and let N be such that $N2^{n+1} \leq \lambda < (N+1)2^{n+1}$. Then

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \leq \sum_j |Q_{N,j}| \leq \frac{1}{2^N}|Q| = e^{-N\log 2}|Q| \leq e^{-C_2\lambda}|Q|,$$

where $C_2 = (\log 2)/2^{n+2}$.
If $\lambda < 2^{n+1}$, then $C_2\lambda < \log(\sqrt{2})$ and

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| \leq |Q| \leq e^{\log(\sqrt{2}) - C_2\lambda}|Q| = \sqrt{2}e^{-C_2\lambda}|Q|,$$

hence $C_1 = \sqrt{2}$. $\square$

6 Singular integrals

Hilbert transform is the most important example of the so-called *singular integrals*. Its n-dimensional analogue are the Riesz transforms.

6.1 Riesz transforms

Riesz transforms are defined as

$$R_j f(x) = c_n \,\text{p.v.} \int_{\mathbb{R}^n} \frac{y_j}{\|y\|^{n+1}} f(x-y)dy, \quad 1 \le j \le n, \tag{6.1}$$

where

$$c_n = \Gamma\Big(\frac{n+1}{2}\Big)\pi^{-\frac{n+1}{2}}.$$

The constants c_n are chosen to have

$$(R_j f)\hat{}(\xi) = -i\frac{\xi_j}{\|\xi\|}\hat{f}(\xi), \tag{6.2}$$

which quickly implies also

$$\sum_{j=1}^{n} R_j^2 = -I.$$

To prove (6.2), we use the technique of homogeneous functions.

A function f is homogeneous of degree $a \in \mathbb{R}$ if for any $x \in \mathbb{R}^n$ and any $\lambda > 0$

$$f(\lambda x) = \lambda^a f(x).$$

To extend this notion also to distributions, let first f be a homogeneous function of degree a and we calculate

$$\int_{\mathbb{R}^n} f(x)\varphi(x)dx = \lambda^{-a}\int_{\mathbb{R}^n} f(\lambda x)\varphi(x)dx = \lambda^{-a}\int_{\mathbb{R}^n} f(x)\varphi(x/\lambda)\frac{dx}{\lambda^n}.$$

We therefore say that $T \in \mathscr{S}'(\mathbb{R}^n)$ is homogeneous of degree $a \in \mathbb{R}$, if for every $\varphi \in \mathscr{S}(\mathbb{R}^n)$

$$T(\varphi_\lambda) = \lambda^a T(\varphi),$$

where $\varphi_\lambda(\cdot) = \lambda^{-n}\varphi(\cdot/\lambda)$.

Proposition 6.1. *If T is a homogeneous distribution of degree a, then its Fourier transform is homogeneous of degree $-n-a$.*

Proof. Indeed, we have

$$\hat{T}(\varphi_\lambda) = T((\varphi_\lambda)\hat{}) = T(\hat{\varphi}(\lambda\cdot)) = \lambda^{-n}T((\hat{\varphi})_{\lambda^{-1}}) = \lambda^{-n-a}T(\hat{\varphi}) = \lambda^{-n-a}\hat{T}(\varphi).$$

□

This proposition allows us to calculate easily the Fourier transform of $f(x) = \|x\|^{-a}$ for $0 < a < n$. Since f is rotationally invariant and homogeneous of degree $-a$, $\hat{f}$ is also rotationally invariant and homogeneous of degree $a-n$, i.e.

$$\hat{f}(\xi) = c_{a,n}\|\xi\|^{a-n}.$$

To find the constants $c_{a,n}$, we calculate for $n/2 < a < n$

$$\int_{\mathbb{R}^n} e^{-\|x\|^2/2}\|x\|^{-a}dx = \int_{\mathbb{R}^n}(e^{-\|x\|^2/2})^\wedge(\xi)(\|x\|^{-a})^\wedge(\xi)d\xi = c_{a,n}\int_{\mathbb{R}^n} e^{-\|\xi\|^2/2}\|\xi\|^{a-n}d\xi.$$

Using polar coordinates, we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-\|x\|^2/2}\|x\|^b dx &= \omega_{n-1}\int_0^\infty e^{-r^2/2}r^{b+n-1}dr = 2^{\frac{b+n}{2}-1}\omega_{n-1}\int_0^\infty e^{-s}s^{\frac{b+n}{2}-1}ds\\ &= 2^{\frac{b+n}{2}-1}\omega_{n-1}\Gamma\Big(\frac{b+n}{2}\Big),\end{aligned}$$

where ω_{n-1} is the area of the n-dimensional unit sphere. Hence

$$c_{a,n} = \frac{2^{\frac{n-a}{2}-1}\Gamma((n-a)/2)}{2^{\frac{a}{2}-1}\Gamma(a/2)} = 2^{\frac{n}{2}-a}\cdot\frac{\Gamma((n-a)/2)}{\Gamma(a/2)}.$$

By the inversion formula, the same holds also for $0 < a < n/2$ and by taking the limit also for $a = n/2$.

Finally, the Fourier transform of the Riesz transform follows from the distributional formula

$$\frac{\partial}{\partial x_j}\|x\|^{-n+1} = (1-n)\,\text{p.v.}\frac{x_j}{\|x\|^{n+1}},$$

which gives

$$\begin{aligned}\Big(\text{p.v.}\frac{x_j}{\|x\|^{n+1}}\Big)^\wedge(\xi) &= \frac{1}{1-n}\Big(\frac{\partial}{\partial x_j}\|x\|^{-n+1}\Big)^\wedge(\xi) = \frac{i\xi_j}{1-n}(\|x\|^{-n+1})^\wedge(\xi)\\ &= \frac{i\xi_j}{1-n}\frac{2^{1-\frac{n}{2}}\Gamma(1/2)}{\Gamma((n-1)/2)}\|\xi\|^{-1} = \frac{-i2^{-n/2}\sqrt{\pi}}{\Gamma\Big(\frac{n+1}{2}\Big)}\cdot\frac{\xi_j}{\|\xi\|}.\end{aligned}$$

Formula (6.2) then follows by this, and by taking the $(2\pi)^{n/2}$ factor into account, which appears in the formula for Fourier transform of convolutions.

Theorem 6.1.1. *The Riesz transforms are bounded on $L_p(\mathbb{R}^n)$ for $1 < p < \infty$.*

Proof. The proof writes the Riesz transforms as a linear (integral) combination of directional one-dimensional Hilbert transforms and then uses the boundedness of H on $L_p(\mathbb{R})$.

$$\begin{aligned}R_jf(x) &= c_n\lim_{\varepsilon\to 0}\int_{\|y\|>\varepsilon}\frac{y_j}{\|y\|^{n+1}}f(x-y)dy = c_n\lim_{\varepsilon\to 0}\int_\varepsilon^\infty\int_{r\mathbb{S}^{n-1}}\frac{y_j}{\|y\|^{n+1}}f(x-y)dydr\\ &= c_n\lim_{\varepsilon\to 0}\int_\varepsilon^\infty\int_{\mathbb{S}^{n-1}}z_jf(x-rz)dz\frac{dr}{r} = \frac{c_n}{2}\lim_{\varepsilon\to 0}\int_{\mathbb{S}^{n-1}}z_j\int_{|r|>\varepsilon}f(x-rz)\frac{dr}{r}dz\\ &= \frac{c_n\pi}{2}\int_{\mathbb{S}^{n-1}}z_jH_zf(x)dz,\end{aligned}$$

where

$$H_zf(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|r|>\varepsilon}f(x-rz)\frac{dr}{r} = H(f(\overline{x}+\cdot z))(x_1),$$

if $x = x_1z+\overline{x}$, $x_1\in\mathbb{R}$ and $\overline{x}\perp z$. We observe first, that H_z is bounded on $L_p(\mathbb{R}^n)$ for every $z\in\mathbb{R}^n$. Indeed,

$$\begin{aligned}\int_{\mathbb{R}^n}|H_zf(x)|^pdx &= \int_{z^\perp}\int_{\mathbb{R}}|H_zf(x_1z+\overline{x})|^pdx_1d\overline{x} = \int_{z^\perp}\int_{\mathbb{R}}|H(f(\overline{x}+\cdot z))(x_1)|^pdx_1d\overline{x}\\ &\le C_p^p\int_{z^\perp}\int_{\mathbb{R}}|f(\overline{x}+x_1z)|^pdx_1d\overline{x} = C_p^p\|f\|_p^p.\end{aligned}$$

Combining these two steps, we obtain

$$\|R_j f\|_p \le \frac{c_n\pi}{2}\int_{\mathbb{S}^{n-1}} |z_j|\cdot\|H_z f\|_p dz \le \frac{c_n\pi C_p\|f\|_p}{2}\int_{\mathbb{S}^{n-1}}|z_j|dz.$$

□

6.2 Riesz potentials

If $\varphi \in \mathscr{S}(\mathbb{R}^n)$, then $(-\Delta\varphi)^\wedge(\xi) = \|\xi\|^2\hat{\varphi}(\xi)$. The fractional Laplace operator may then be defined as

$$((-\Delta)^{a/2}\varphi)^\wedge(\xi) = \|\xi\|^a\hat{\varphi}(\xi).$$

We define

$$I_a(f) = (-\Delta)^{-a/2}(f), \quad 0 < a < n,$$

i.e.

$$I_a f(x) = \frac{1}{\gamma_a}\int_{\mathbb{R}^n}\frac{f(y)}{\|x-y\|^{n-a}}dy,$$

where

$$\gamma_a = \pi^{n/2-a}\frac{\Gamma(a/2)}{\Gamma((n-a)/2)}.$$

Let $\Lambda_\delta(f)(x) = f(\delta x)$, then

$$\Lambda_{\delta^{-1}} I_a \Lambda_\delta = \delta^{-a} I_a, \quad \|\Lambda_\delta(f)\|_p = \delta^{-n/p}\|f\|_p. \quad \|\Lambda_{\delta^{-1}} I_a(f)\|_q = \delta^{n/q}\|I_a(f)\|_q.$$

If $\|I_a f\|_q \le C\|f\|_p$ is true, then we get by

$$\begin{aligned}\|I_a(\Lambda_\delta f)\|_q &= \|\Lambda_\delta(\delta^{-a} I_a f)\|_q = \delta^{-a}\|\Lambda_\delta I_a f\|_q \\ &= \delta^{-a}\|I_a f\|_q\delta^{-n/q} \le C\delta^{-a-n/q}\|f\|_p = \delta^{-a-n/q+n/p}\|\Lambda_\delta f\|_p\end{aligned}$$

that

$$\frac{1}{q} = \frac{1}{p} - \frac{a}{n}. \tag{6.3}$$

Theorem 6.2.1. *Let $0 < a < n$, $1 \le p < n/a$ and define q by (6.3). Then $\|I_a f\|_q \le C\|f\|_p$ for $p > 1$ and $\|I_a f\|_{q,w} \le C\|f\|_p$ for $p = 1$.*

Proof. We give a proof, which is based on the following inequality due to Hedberg (1972):

$$I_a f(x) \le C_a\|f\|_p^{ap/n}\cdot Mf(x)^{1-ap/n}. \tag{6.4}$$

This inequality implies for $p > 1$

$$\|I_a f\|_q \le C_a\|f\|_p^{ap/n}\|(Mf)^{1-ap/n}\|_q = C_a\|f\|_p^{ap/n}\|Mf\|_p^{1-ap/n} \lesssim \|f\|_p,$$

and the weak bound follows for $p = 1$ in the same manner. To show (6.4), we argue as follows.

Fix $x \in \mathbb{R}^n$ and split (for $A > 0$ to be chosen later on)

$$\gamma_a I_a f(x) = \int_{y:\|x-y\|\le A}\frac{f(y)}{\|x-y\|^{n-a}}dy + \int_{y:\|x-y\|>A}\frac{f(y)}{\|x-y\|^{n-a}}dy.$$

We apply Theorem 1.2.4 to the first integral

$$\int_{y:\|x-y\|\le A}\frac{f(y)}{\|x-y\|^{n-a}}dy \le C\int_{y:\|x-y\|\le A}\frac{dy}{\|x-y\|^{n-a}}\cdot Mf(x) \tag{6.5}$$
$$\le C'\int_0^A r^{-(n-a)}r^{n-1}dr\cdot Mf(x) = C''A^a\cdot Mf(x).$$

The second part may be estimated by Hölder's inequality as follows

$$\int_{y:\|x-y\|>A}\frac{f(y)}{\|x-y\|^{n-a}}dy \le C\|f\|_p\cdot\Big(\int_{y:\|x-y\|>A}\|x-y\|^{(a-n)p'}dy\Big)^{1/p'} \tag{6.6}$$
$$\le C'\Big(\int_A^\infty r^{n-1}r^{(a-n)p'}dr\Big)^{1/p'}$$
$$\le C''\|f\|_p\cdot A^{((a-n)p'+n)/p'} = C''\|f\|_p\cdot A^{a-n+n/p'} = C''\|f\|_pA^{a-n/p}$$

Choosing A in such a way, that (6.5) and (6.6) coincide, we get $A=(\|f\|_p/Mf(x))^{p/n}$ and (6.4) follows. □

Definition 6.2.2. Let $k\in\mathbb{N}_0$ and let $1\le p\le\infty$. Then the Sobolev space $W_p^k(\mathbb{R}^n)$ is the set of all functions from $L_p(\mathbb{R}^n)$, such that all its (distributional) derivatives up to order k belong to $L_p(\mathbb{R}^n)$.

Theorem 6.2.3. (Sobolev's embedding theorem) *Let $k\in\mathbb{N}$, $1\le p\le\infty$ and $1/q=1/p-k/n$.*

(i) If $q<\infty$, i.e. $p<n/k$, then $W_p^k(\mathbb{R}^n)\hookrightarrow L_q(\mathbb{R}^n)$.

(ii) If $q=\infty$. i.e. $p=n/k$, then the restriction of any $f\in W_p^k(\mathbb{R}^n)$ to a compact subset of $\mathbb{R}^n$ belongs to $L_r(\mathbb{R}^n)$ for any $r<\infty$.

(iii) If $p>n/k$, then every $f\in W_p^k(\mathbb{R}^n)$ can be modified on a set of measure zero so that the resulting function is continuous.

Proof. For a direct proof, we refer to the book of Stein, Chapter 5. Let us sketch the proof using Riesz transforms and potentials.

$$\Big(R_j\Big(\frac{\partial}{\partial x_j}f\Big)\Big)^\wedge(\xi)=\frac{\xi_j^2}{\|\xi\|}\hat f(\xi)$$

and

$$f=I_1\Big(\sum_{j=1}^n R_j\Big(\frac{\partial}{\partial x_j}f\Big)\Big). \tag{6.7}$$

We deal with $k=1$, higher order follow by iteration. Let $f\in W_p^1(\mathbb{R}^n)$.

Let $1<p<n$. Then all $\partial f/\partial x_j$ belong to $L_p(\mathbb{R}^n)$ and their Riesz transforms $R_j(\partial f/\partial x_j)$ belong also to $L_p(\mathbb{R}^n)$. Finally, I_1 maps $L_p(\mathbb{R}^n)$ into $L_q(\mathbb{R}^n)$. This gives the proof of (i) for $1<p<\infty$. We leave out the proof for $p=1$.

For the proof of (ii), we consider only ηf, where η is a smooth function with compact support. We apply (6.7) to ηf, which (together with its first order derivatives) belongs not only to $L_p(\mathbb{R}^n)$ but also to all $L_s(\mathbb{R}^n)$, $1<s<p=n$. Riesz transforms R_j then map these to $L_s(\mathbb{R}^n)$ again, and the Riesz potential I_1 then maps the outcome into $L_r(\mathbb{R}^n)$ for $1/r-1/s-1/n$, i.e. into every $L_r(\mathbb{R}^n)$, $r<\infty$. The proof of (iii) is again left out. □

6.3 Calderón-Zygmund operators

These are convolutions with kernel K, which might have a singularity at origin.

Theorem 6.3.1. (Calderón-Zygmund) *Let $K \in \mathscr{S}'(\mathbb{R}^n)$ be a tempered distribution, which is associated to a locally integrable function on $\mathbb{R}^n \setminus \{0\}$ and satisfies*

(CZ1) $\quad |\hat{K}(\xi)| \leq A, \quad \xi \in \mathbb{R}^n,$

(CZ2) $\displaystyle\int_{\|x\|_2 \geq 2\|y\|_2} |K(x-y) - K(x)| dx \leq B, \quad y \in \mathbb{R}^n.$

Then, for $1 < p < \infty$,

$$\|K * f\|_p \leq C_p \|f\|_p \quad \textit{and} \quad \|K * f\|_{1,w} \leq C\|f\|_1.$$

The proof copies very much the proof of the same statement for the Hilbert transform, and we leave out the details.

The condition (CZ2) is sometimes called *Hörmander condition.* By the help of mean value theorem, it is satisfied for example if

$$\|\nabla K(x)\|_2 \leq \frac{C}{\|x\|_2^{n+1}}, \quad x \neq 0.$$

An important and non-trivial generalisation of the theory of singular integrals is given by considering the vector-valued analogues. By this, we mean the following.

- $H, \tilde{H}$ are (complex) Hilbert spaces.
- For $0 < p < \infty$, $L_p(\mathbb{R}^n \to H)$ is the set of measurable functions $f : \mathbb{R}^n \to H$, such that $\int_{\mathbb{R}^n} \|f(x)\|_H^p dx < \infty$.
- Let $K : \mathbb{R}^n \to \mathscr{L}(H, \tilde{H})$. Then $Tf(x) = \displaystyle\int_{\mathbb{R}^n} K(y) f(x-y) dy$ takes values in $\tilde{H}$.
- Under same (just appropriately interpreted) conditions as above, T is bounded from $L_p(\mathbb{R}^n \to H)$ into $L_p(\mathbb{R}^n \to \tilde{H})$. Especially, the gradient condition above is still valid in this case.

7 Special role of $p = 2$

7.1 Khintchine inequality

We denote by

$$r_n(t) := \text{sign } \sin(2^n \pi t), \quad t \in [0,1], n \in \mathbb{N}_0.$$

the *Rademacher functions.*

The system $(r_n)_{n=0}^\infty$ forms an orthonormal system in $L_2(0,1)$, but it is not a basis (consider i.e. the function $f(t) = 1 - 2t$).

Theorem 7.1.1. *Let $p \in [1,\infty)$. Then there are positive constants A_p and B_p such that*

$$A_p \left(\sum_{n=1}^m |a_n|^2 \right)^{1/2} \leq \left(\int_0^1 \left| \sum_{n=1}^m a_n r_n(t) \right|^p dt \right)^{1/p} \leq B_p \left(\sum_{n=1}^m |a_n|^2 \right)^{1/2}$$

holds for every $m \in \mathbb{N}$ and every sequence of real numbers $a_1, \ldots, a_m$.

Proof. By A_p and B_p we denote the best possible constants (which are actually known, but we shall derive slightly weaker estimates). Furthermore, orthogonality of Rademacher functions gives immediately $A_2 = B_2 = 1$. Finally, due to monotonicity of the L_p-norms, we have $A_r \leq A_p$ and $B_r \leq B_p$ for $r \leq p$.

So, it is enough to show that $A_1 > 0$ and $B_{2k} < \infty$ for all $k \in \mathbb{N}$.

We start with B_{2k}. Let us observe that

$$\begin{aligned} E := \int_0^1 \left| \sum_{n=1}^m a_n r_n(t) \right|^{2k} dt &= \int_0^1 \left(\sum_{n=1}^m a_n r_n(t) \right)^{2k} dt \\ &= \sum_{|\alpha|=2k} \frac{(2k)!}{\alpha_1! \ldots \alpha_m!} a_1^{\alpha_1} \ldots a_m^{\alpha_m} \int_0^1 r_1^{\alpha_1}(t) \ldots r_m^{\alpha_m}(t) dt \\ &= \sum_{|\alpha|=k} \frac{(2k)!}{(2\alpha_1)! \ldots (2\alpha_m)!} a_1^{2\alpha_1} \ldots a_m^{2\alpha_m} \int_0^1 r_1^{2\alpha_1}(t) \ldots r_m^{2\alpha_m}(t) dt \\ &= \sum_{|\alpha|=k} \frac{(2k)!}{(2\alpha_1)! \ldots (2\alpha_m)!} a_1^{2\alpha_1} \ldots a_m^{2\alpha_m}, \end{aligned}$$

where we have used the multinomial theorem (a generalisation of the binomial theorem to a bigger number of summands) and the fact that

$$\int_0^1 r_1^{\alpha_1}(t) \ldots r_m^{\alpha_m}(t) dt$$

is equal to zero if some of the α_i's is odd and equal to 1 if all of them are even.

Let $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}_0^m$ be integers with $|\alpha| = k$, then

$$2^k \alpha_1! \ldots \alpha_m! = (2^{\alpha_1} \alpha_1!) \ldots (2^{\alpha_m} \alpha_m!) \leq (2\alpha_1)! \ldots (2\alpha_m)!.$$

This implies

$$\begin{aligned} E &\leq \frac{(2k)!}{2^k k!} \sum_{|\alpha|=k} \frac{k!}{\alpha_1! \ldots \alpha_m!} a_1^{2\alpha_1} \ldots a_m^{2\alpha_m} \\ &= \frac{(2k)!}{2^k k!} \left(\sum_{n=1}^m |a_n|^2 \right)^k = \frac{(2k)!}{2^k k!} \|a\|_2^{2k} \end{aligned}$$

and

$$E^{1/(2k)} \le \left(\frac{(2k)!}{2^k k!}\right)^{1/(2k)} \|a\|_2.$$

Hence, the statement holds with [6]

$$B_{2k} := \left(\frac{(2k)!}{2^k k!}\right)^{1/(2k)}.$$

Finally, we have to show the existence of $A_1 > 0$. We proceed by a nice duality trick using the (already proven) first part of this theorem.

Let $f(t) := \sum_{n=1}^m a_n r_n(t)$. By Hölder's inequality for $p = 3/2$ and $p' = 3$, we have

$$\begin{aligned}\int_0^1 |f(t)|^2 dt &= \int_0^1 |f(t)|^{2/3} \cdot |f(t)|^{4/3} dt \le \left(\int_0^1 |f(t)| dt\right)^{2/3} \cdot \left(\int_0^1 |f(t)|^4\right)^{1/3}\\ &\le \left(\int_0^1 |f(t)| dt\right)^{2/3} B_4^{4/3} \cdot \|a\|_2^{4/3} = \left(\int_0^1 |f(t)| dt\right)^{2/3} B_4^{4/3} \cdot \|f\|_2^{4/3}.\end{aligned}$$

Therefore,

$$\left(\int_0^1 |f(t)| dt\right)^{2/3} \ge B_4^{-4/3} \left(\int_0^1 |f(t)|^2 dt\right)^{1/3},$$

that is

$$\int_0^1 |f(t)| dt \ge B_4^{-2} \left(\int_0^1 |f(t)|^2 dt\right)^{1/2} = B_4^{-2} \|a\|_2.$$

Hence, $A_1 \ge B_4^{-2}$. □

Remark 7.1.2. Stochastic reformulation of Khintchine's inequalities sounds as follows. Let $\varepsilon_i, i = 1, \dots, m$ be independent variables with $\mathbb{P}(\varepsilon_i = 1) = 1/2$ and $\mathbb{P}(\varepsilon_i = -1) = 1/2$. Let $1 \le p < \infty$. Then there are constants A_p, B_p such that for every $a_1, \dots, a_m \in \mathbb{R}$ [7]

$$A_p \|a\|_2 \le \left(\mathbb{E}\left|\sum_{i=1}^m a_i \varepsilon_i\right|^p\right)^{1/p} \le B_p \|a\|_2.$$

Choosing p large enough, this estimate gives very quickly the so-called *tail bound estimates on sum of independent Rademacher variables*, i.e. the assymptotic estimates of

$$\mathbb{P}\left(\left|\sum_{i=1}^m a_i \varepsilon_i\right| > t\right)$$

for $t \to \infty$.

We use this reformulation of Khintchine's inequalities to give another proof of Theorem 7.1.1.

Proof. (of the upper estimate in Theorem 7.1.1).
We normalize to $\|a\|_2 = 1$. Then

$$\mathbb{E} \exp\left(\sum_{i=1}^m a_i \varepsilon_i\right) = \mathbb{E} \prod_{i=1}^m \exp(a_i \varepsilon_i) = \prod_{i=1}^m \mathbb{E} \exp(a_i \varepsilon_i) = \prod_{i=1}^m \cosh(a_i).$$

[6] By Stirling's formula, one can show quite easily that B_{2k} grows as $\sqrt{2k}$ for $k \to \infty$.
[7] Also $a_1, \dots, a_m \in \mathbb{C}$ can be considered with slightly modified proof.

Using Taylor's expansion, one obtains $\cosh(a_j) \le \exp(a_j^2/2)$. Hence,

$$\mathbb{E}\exp\left(\sum_{i=1}^m a_i\varepsilon_i\right) \le \prod_{i=1}^m \exp(a_i^2/2) \lesssim 1,$$

and by Markov's inequality

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^m a_i\varepsilon_i > \lambda\right) &= \mathbb{P}\left(\exp\left(\sum_{i=1}^m a_i\varepsilon_i\right) > \exp(\lambda)\right) = \mathbb{P}\left(\exp\left(\sum_{i=1}^m a_i\varepsilon_i - \lambda\right) > 1\right)\\ &\le \mathbb{E}\exp\left(\sum_{i=1}^m a_i\varepsilon_i - \lambda\right) \lesssim e^{-\lambda}.\end{aligned}$$

By symmetry of ε_i's, we get also $\mathbb{P}\left(\left|\sum_{i=1}^m a_i\varepsilon_i\right| > \lambda\right) \lesssim e^{-\lambda}$. The rest then follows by distributional representation of the L_p-norm. □

Khintchine's inequalities have an interesting application in operator theory. Let $1 \le p < \infty$ and let $T : L_p(\mathbb{R}^n) \to L_p(\mathbb{R}^n)$ be a bounded linear operator. Then

$$\left\|\left(\sum_{j=0}^N |Tf_j|^2\right)^{1/2}\right\|_p \le c_p\left\|\left(\sum_{j=0}^N |f_j|^2\right)^{1/2}\right\|_p, \tag{7.1}$$

where the constant c_p depends only on p and $\|T\|$.

The proof follows by considering Rademacher functions $r_1, \dots, r_N$ and

$$\begin{aligned}\left\|\left(\sum_{j=0}^N |Tf_j|^2\right)^{1/2}\right\|_p^p &= \int_{\mathbb{R}^n}\left(\sum_{j=1}^N |(Tf_j)(x)|^2\right)^{p/2} dx \le c\int_{\mathbb{R}^n}\left(\int_0^1 |\sum_{j=1}^N Tf_j(x)r_j(t)|^p dt\right)^{p/p} dx\\ &= c\int_0^1\int_{\mathbb{R}^n} |\sum_{j=1}^N Tf_j(x)r_j(t)|^p dxdt = c\int_0^1\int_{\mathbb{R}^n} |T\Big(\sum_{j=1}^N f_jr_j(t)\Big)(x)|^p dxdt\\ &\le c\|T\|^p\int_0^1\int_{\mathbb{R}^n} |\sum_{j=1}^N f_j(x)r_j(t)|^p dxdt \le c'\int_{\mathbb{R}^n}\left(\sum_{j=1}^N |f_j(x)|^2\right)^{p/2} dx\\ &= c_p^p\left\|\left(\sum_{j=0}^N |f_j|^2\right)^{1/2}\right\|_p^p.\end{aligned}$$

By letting $N \to \infty$, the same result holds also for infinite sums.

7.2 Littlewood-Paley Theory

Let $\{I_j\}$ be a sequence of intervals on the real line, finite or infinite, and let $\{S_j\}$ be the sequence of operators defined by $(S_jf)^\wedge(\xi) = \chi_{I_j}(\xi)\hat{f}(\xi)$. Later on, we shall concentrate on the *dyadic decomposition* of $\mathbb{R}$ (strictly speaking of $\mathbb{R} \setminus \{0\}$) given by

$$I_j = (-2^{j+1}, -2^j] \cup [2^j, 2^{j+1}), \quad j \in \mathbb{Z}. \tag{7.2}$$

Furthermore, we denote $S_j^* = S_{j-1} + S_j + S_{j+1}$. Let us observe that this implies $S_j^*S_j = S_jS_j^* = S_j$.

Finally, we adopt this concept also to smooth dyadic decompositions. Let $\psi \in \mathscr{S}(\mathbb{R})$ be non-negative, have support in $1/2 \le \|\xi\|_2 \le 4$ and be equal to 1 on $1 \le \|\xi\|_2 \le 2$. Then we define

$$\psi_j(\xi) = \psi(2^{-j}\xi) \quad \text{and} \quad (\tilde{S}_j f)^\wedge(\xi) = \psi_j(\xi)\hat{f}(\xi), \quad \xi \in \mathbb{R}.$$

Theorem 7.2.1. (Littlewood-Paley Theory) *Let $1 < p < \infty$.*

i) Then there exist two constants $C_p > c_p > 0$, such that

$$c_p\|f\|_p \le \left\| \left(\sum_j |S_j f|^2 \right)^{1/2} \right\|_p \le C_p\|f\|_p.$$

The same holds for S_j^*.

ii) There exists a constant $C_p > 0$ such that

$$\left\| \left(\sum_j |\tilde{S}_j f|^2 \right)^{1/2} \right\|_p \le C_p\|f\|_p.$$

iii) Finally, if $\sum_j |\psi(2^{-j}\xi)|^2 = 1$, then there is also a constant $c_p > 0$, such that

$$c_p\|f\|_p \le \left\| \left(\sum_j |\tilde{S}_j f|^2 \right)^{1/2} \right\|_p.$$

Proof. Step 1.
We know, that $(S_j f)^\wedge = \chi_{I_j}\hat{f}$, where I_j was defined by (7.2). We define

$$I_j^- := (-2^{j+1}, -2^j], \qquad I_j^+ := [2^j, 2^{j+1}), \quad j \in \mathbb{Z}$$

and split $S_j = S_j^- + S_j^+$, where $(S_j^- f)^\wedge = \chi_{I_j^-}\hat{f}$ and $(S_j^+ f)^\wedge = \chi_{I_j^+}\hat{f}$.

We observe, that

$$\chi_{I_j^+}(x) = \frac{1}{2}\left(\operatorname{sgn}(x - 2^j) - \operatorname{sgn}(x - 2^{j+1})\right) \quad \text{for (almost) all } x \in \mathbb{R},$$

and

$$S_j^+ f = (\chi_{I_j^+}\hat{f})^\vee = \frac{1}{2}\left((\operatorname{sgn}(\cdot - 2^j)\hat{f})^\vee - (\operatorname{sgn}(\cdot - 2^{j+1})\hat{f})^\vee\right).$$

Finally, we write

$$\operatorname{sgn}(\xi - 2^j)\hat{f}(\xi) = \tau_{2^j}[\operatorname{sgn}(\xi)\hat{f}(\xi + 2^j)] = \tau_{2^j}[\operatorname{sgn}(\xi) \cdot \tau_{-2^j}\hat{f}(\xi)],$$

leading to

$$\begin{aligned}
(\operatorname{sgn}(\xi - 2^j)\hat{f}(\xi))^\vee &= M_{2^j}(\operatorname{sgn} \cdot \tau_{-2^j}\hat{f})^\vee = (2\pi)^{-1/2} \cdot M_{2^j}(\operatorname{sgn}(\cdot)^\vee * (\tau_{-2^j}\hat{f})^\vee) \\
&= (2\pi)^{-1/2} M_{2^j}\left(\frac{\sqrt{2\pi}}{-i}\left(\frac{1}{\pi} p.v.\frac{1}{x}\right) * M_{-2^j} f\right) \\
&= iM_{2^j}\left(\left(\frac{1}{\pi} p.v.\frac{1}{x}\right) * M_{-2^j} f\right) \\
&= iM_{2^j} H M_{-2^j} f.
\end{aligned}$$

Using the boundedness of H on $L_p(\mathbb{R})$ for $1 < p < \infty$, we immediately obtain that $\|S_j f\|_p \le c\|f\|_p$, and the same is true also for S_j^+ and S_j^-.

Step 2.

We combine Step 1. with (7.1) to obtain

$$\begin{aligned}\Big\|\Big(\sum_{j\in\mathbb{Z}}|S_j^+ f_j|^2\Big)^{1/2}\Big\|_p &\le \frac{1}{2}\left\{\Big\|\Big(\sum_{j\in\mathbb{Z}}|(\mathrm{sgn}(\xi-2^j)\hat{f}_j(\xi))^\vee|^2\Big)^{1/2}\Big\|_p + \Big\|\Big(\sum_{j\in\mathbb{Z}}|(\mathrm{sgn}(\xi-2^{j+1})\hat{f}_j(\xi))^\vee|^2\Big)^{1/2}\Big\|_p\right\}\\ &\le \frac{1}{2}\left\{\Big\|\Big(\sum_{j\in\mathbb{Z}}|M_{2^j}HM_{-2^j}f_j|^2\Big)^{1/2}\Big\|_p + \dots\right\}\\ &= \frac{1}{2}\left\{\Big\|\Big(\sum_{j\in\mathbb{Z}}|HM_{-2^j}f_j|^2\Big)^{1/2}\Big\|_p + \dots\right\}\\ &\le c_p\left\{\Big\|\Big(\sum_{j\in\mathbb{Z}}|M_{-2^j}f_j|^2\Big)^{1/2}\Big\|_p + \dots\right\} = c_p'\Big\|\Big(\sum_{j\in\mathbb{Z}}|f_j|^2\Big)^{1/2}\Big\|_p.\end{aligned}$$

The same holds of course for S_j^- and, therefore, also for S_j.

Step 3.

This, together with the identity $S_j = S_j\tilde{S}_j$ implies

$$\Big\|\Big(\sum_{j\in\mathbb{Z}}|S_j f|^2\Big)^{1/2}\Big\|_p = \Big\|\Big(\sum_{j\in\mathbb{Z}}|S_j\tilde{S}_j f|^2\Big)^{1/2}\Big\|_p \le \Big\|\Big(\sum_{j\in\mathbb{Z}}|\tilde{S}_j f|^2\Big)^{1/2}\Big\|_p.$$

Step 4.

This shows, that the second inequality in part (i) of the theorem follows from (ii). Therefore, we prove (ii) now.

Let $\hat{\Psi} = \psi$ and $\Psi_j(x) = 2^j\Psi(2^j x)$. Then $\hat{\Psi}_j = \psi_j$ and $\tilde{S}_j f = \Psi_j * f$. It is enough to prove that the vector-valued mapping

$$f \to (\tilde{S}_j f)_j$$

is bounded from L_p to $L_p(\ell_2)$. If $p = 2$, this follows by Plancherel theorem:

$$\left\|\left(\sum_j |\tilde{S}_j f|^2\right)^{1/2}\right\|_2^2 = \int_{\mathbb{R}}\sum_j |\psi_j(\xi)|^2\cdot|\hat{f}(\xi)|^2 d\xi \le 3\|f\|_2^2.$$

The proof for $p \neq 2$ is an application of the vector-valued Calderón-Zygmund theory, and especially the Hörmander's condition for vector-valued singular integrals.

Let us first give the notation. Let $H = \mathbb{C}$ and $\tilde{H} = \ell_2(\mathbb{Z})$. This means that H is just one-dimensional Hilbert space and $\tilde{H}$ is the ℓ_2 space of sequences indexed by integers. Then $L_p(\mathbb{R}\to H) = L_p(\mathbb{R}\to\mathbb{C})$ is just the usual $L_p(\mathbb{R})$ space of complex-valued functions. And $L_p(\mathbb{R}\to\tilde{H})$ is a space of functions $g:\mathbb{R}\to\ell_2(\mathbb{Z})$. Each $g(x)$ is then a sequence $(\dots, g_{-1}(x), g_0(x), g_1(x), \dots)$ and the $L_p(\mathbb{R}\to\tilde{H})$ norm of g is given by

$$\begin{aligned}\|g|L_p(\mathbb{R}\to\tilde{H})\| = \|(g_j)_{j\in\mathbb{Z}}|L_p(\mathbb{R}\to\tilde{H})\| &= \Big(\int_{\mathbb{R}}\|(g_j(x))_{j\in\mathbb{Z}}\|_{\tilde{H}}^p dx\Big)^{1/p}\\ &= \Big(\int_{\mathbb{R}}\Big(\sum_{j\in\mathbb{Z}}|g_j(x)|^2\Big)^{p/2}dx\Big)^{1/p} = \Big\|\Big(\sum_{j\in\mathbb{Z}}|g_j(x)|^2\Big)^{1/2}\Big\|_p.\end{aligned}$$

Now, we define a vector-valued Calderón-Zygmund operator T based on a kernel K. We need $K(x) \in \mathcal{L}(H, \tilde{H}) = \mathcal{L}(\mathbb{C}, \ell_2(\mathbb{Z}))$ for every $x \in \mathbb{R}$. Therefore, we first characterize the elements of $\mathcal{L}(\mathbb{C}, \ell_2(\mathbb{Z}))$. As the source space is one-dimensional, it is easy to see that each element $Z \in \mathcal{L}(\mathbb{C}, \ell_2(\mathbb{Z}))$ is given by

$$Z(\lambda) = \lambda \cdot z$$

for some $z \in \ell_2(\mathbb{Z})$. We therefore define

$$K(x) : \lambda \to \lambda \cdot (\dots, \Psi_{-1}(x), \Psi_0(x), \Psi_1(x), \dots).$$

Without giving the formal detail on vector-valued integration, we get

$$\begin{aligned} Tf(x) &= \int_{\mathbb{R}} K(y) f(x-y) dy = \int_{\mathbb{R}} f(x-y) \cdot (\dots, \Psi_{-1}(y), \Psi_0(y), \Psi_1(y), \dots) dy \\ &= \Big(\dots, \int_{\mathbb{R}} f(x-y)\Psi_{-1}(y)dy, \int_{\mathbb{R}} f(x-y)\Psi_0(y)dy, \int_{\mathbb{R}} f(x-y)\Psi_1(y)dy, \dots\Big) \\ &= \Big(\dots, \Psi_{-1} * f(x), \Psi_0 * f(x), \Psi_1 * f(x), \dots\Big) \\ &= \Big(\dots, \tilde{S}_{-1} f, \tilde{S}_0 f, \tilde{S}_1 f, \dots\Big) = \text{"}K * f(x)\text{"}. \end{aligned}$$

When we show, that Calderón-Zygmund theorem can be applied, this will give

$$\|Tf | L_p(\mathbb{C} \to \ell_2(Z))\| \le C \|f\|_p.$$

Due to what we said before, this is exactly the inequality in (ii).

So, we are left with verifying the two conditions $(CZ1)$ and $(CZ2)$. The first one is equivalent to showing that

$$(\dots, \hat{\Psi}_{-1}(\xi), \hat{\Psi}_0(\xi), \hat{\Psi}_1(\xi), \dots) = (\dots, \psi_{-1}(\xi), \psi_0(\xi), \psi_1(\xi), \dots)$$

is uniformly bounded in $\ell_2(\mathbb{Z})$ for every $\xi \in \mathbb{R}$. But that follows easily from the support properties of the functions ψ_j.

To verify the second, we need to show that

$$\|\Psi_j'(x)\|_{\ell_2} \le C|x|^{-2}, \quad x \in \mathbb{R}.$$

Using that $\Psi \in \mathscr{S}(\mathbb{R})$, we obtain

$$\begin{aligned} \Big(\sum_{j\in\mathbb{Z}} |\Psi_j'(x)|^2\Big)^{1/2} &\le \sum_{j\in\mathbb{Z}} |\Psi_j'(x)| = \sum_j 2^{2j} |\Psi'(2^j x)| \le C \sum_{j\in\mathbb{Z}} 2^{2j} \min(1, (2^j|x|)^{-3}) \\ &= C \sum_{j:2^j|x|\le 1} \dots + C \sum_{j:2^j|x|>1} \dots \\ &= C \sum_{j:2^j|x|\le 1} 2^{2j} + C \sum_{j:2^j|x|>1} 2^{2j} (2^j|x|)^{-3} \le \frac{C'}{|x|^2}. \end{aligned}$$

Step 5.

Finally, we prove the first inequality in part (i) and part (iii) of the theorem. Surprisingly enough, they follow quite quickly from previous steps and duality.

The identity

$$\Big\| \Big(\sum_{j \in \mathbb{Z}} |S_j f|^2 \Big)^{1/2} \Big\|_2 = \|f\|_2 \tag{7.3}$$

follows by Plancherel's theorem and, by polarization, also

$$\int_{\mathbb{R}} \sum_j S_j f \cdot \overline{S_j g} = \int_{\mathbb{R}} f \overline{g}$$

follows. Using this, and the first part of the theorem for p' with $1/p + 1/p' = 1$ allows the following estimate.

$$\begin{aligned}
\|f\|_p &= \sup \left\{ \left| \int_{\mathbb{R}} f \overline{g} \right| : \|g\|_{p'} \le 1 \right\} \\
&= \sup \left\{ \left| \int_{\mathbb{R}} \sum_j S_j f \cdot \overline{S_j g} \right| : \|g\|_{p'} \le 1 \right\} \\
&\le \sup \left\{ \Big\| \Big(\sum_{j \in \mathbb{Z}} |S_j f|^2 \Big)^{1/2} \Big\|_p \cdot \Big\| \Big(\sum_{j \in \mathbb{Z}} |S_j g|^2 \Big)^{1/2} \Big\|_{p'} : \|g\|_{p'} \le 1 \right\} \\
&\le c_p \Big\| \Big(\sum_{j \in \mathbb{Z}} |S_j f|^2 \Big)^{1/2} \Big\|_p .
\end{aligned}$$

Part (iii) of the theorem follows in exactly the same way - the assumption of the theorem gives exactly the identity (7.3). □

References

[1] W. Rudin, *Functional Analysis*, McGraw-Hill, 1991.

[2] D. Werner, *Funktionalanalysis*, 7. Aufl., Springer, 2011.

[3] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J., 1970.

[4] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.

[5] H. Triebel, *Higher analysis*, University Books for Mathematics, Johann Ambrosius Barth Verlag GmbH, Leipzig, 1992.

[6] L. Grafakos, *Classical Fourier analysis*, Second edition, Graduate Texts in Mathematics, 249. Springer, New York, 2008.

[7] L. Grafakos, *Modern Fourier analysis*, Second edition, Graduate Texts in Mathematics, 250. Springer, New York, 2009.

[8] J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics 29, American Mathematical Society, 2000.