Harmonic Analysis Lecture Notes

Guo, Jingwei

June 25, 2025

Contents

Glossary

$\lvert A\rvert$	the Lebesgue measure of the set $A \subset \mathbb{R}^n$
$d_f(\alpha)$	the measure of the level set $\{f > \alpha\}$
$f = O(g)$	$\lvert f\rvert \leq C\lvert g\rvert$ for some unimportant constant C
$f = O_n(g)$	$\lvert f\rvert \leq C\lvert g\rvert$ for some constant C depending on n
$f \precsim g$	the same meaning with $f = O(g)$
$f \precsim_n g$	the same meaning with $f = O_n(g)$
$f \approx g$	$f \precsim g$ and $g \precsim f$
$f \approx_n g$	$f \precsim_n g$ and $g \precsim_n f$
p'	the conjugate exponent of $p \geq 1$ when there is no special claim
$B(x,r)$	the open ball in $\mathbb{R}^n$ with center x and radius r
B_r	$B(0,r)$ for short
ν_n	the volumn of B_1
ω_n	the area of $\mathbb{S}^{n-1}$
C_c	the space of continuous functions with compact support
C_0	the space of continuous functions vanishing at infinity
C^∞	the space of smooth functions
C_c^∞	the space of smooth functions with compact support
$\mathscr{S}$	the Schwartz space
$\mathscr{S}'$	the space of tempered distributions
τ^y	the translation operator, $\tau^y f(x) = f(x-y)$
δ^a	the dilation operator, $\delta^a f(x) = f(ax)$
$\sim$	the reflection operator when there is no special claim, $\tilde{f}(x) = f(-x)$
$-$	the conjugate operator when there is no special claim, $\bar{f}(x) = \overline{f(x)}$
$\lvert\alpha\rvert$	$\alpha_1 + ... + \alpha_n$ for muti-index $\alpha = (\alpha_1, ..., \alpha_n)$
x^α	$x_1^{\alpha_1}...x_n^{\alpha_n}$ for muti-index $\alpha = (\alpha_1, ..., \alpha_n)$
D^α	the differential operator of order $\lvert\alpha\rvert$, $D^\alpha f = \frac{\partial^{\lvert\alpha\rvert}}{\partial x_1^{\alpha_1}...\partial x_n^{\alpha_n}} f$
D_j	the same meaning with $\frac{\partial}{\partial x_j}$
f_ε	the dilation of f, $f_\varepsilon(x) = \varepsilon^{-n} f(\varepsilon^{-1}x)$
$\langle u, \varphi\rangle$	$u \in \mathscr{S}'$ act on $\varphi \in \mathscr{S}$
$\hat{f}$	the Fourier transform of f

$\check{f}$	the Fourier inverse transform transform of f
$\Gamma(z)$	the value of gamma function at point z
$DG(p)$	the derivation of $G : \mathbb{R}^n \to \mathbb{R}^n$ at p,
$J_G(p)$	the Jacobian of $G : \mathbb{R}^n \to \mathbb{R}^n$ at p
$H_f(p)$	the Hessian of $f : \mathbb{R}^n \to \mathbb{R}$ at p
$2B$	$B(x, 2r)$ if $B = B(x, r)$
$2Q$	the cube with the same center and twice the sidelength of Q

1 Preliminaries

1.1 Interpolation

We will not go into details of contents of L^p space, here we just mention the following technique of function decomposition: assume $0 < p_0 < p < p_1 \leq \infty$, then we can decompose $f \in L^p$ as follows:

$$f = f^\gamma + f_\gamma, \quad \text{where } f^\gamma = f\chi_{\{|f|>\gamma\}} \in L^{p_0}, f_\gamma = f\chi_{\{|f|\leq\gamma\}} \in L^{p_1}.$$

More specifically, we have:

$$\|f^\gamma\|_{p_0}^{p_0} = p_0 \int_\gamma^\infty \alpha^{p_0-1} d_f(\alpha) d\alpha + \gamma^{p_0} d_f(\gamma) < \infty. \tag{1.1.1}$$

$$\|f_\gamma\|_{p_1}^{p_1} = p_1 \int_0^\gamma \alpha^{p_1-1} d_f(\alpha) d\alpha - \gamma^{p_1} d_f(\gamma) < \infty. \tag{1.1.2}$$

To see why the right end of (1.1.1),(1.1.2) is $< \infty$ (that is, why the integral is convergent), just use the following Markov inequality:

$$d_f(\alpha) \leq \alpha^{-p} \|f\|_p^p \tag{1.1.3}$$

However, if we just ask for the (1.1.3)-type inequality (i.e. $d_f(\alpha) = O(\alpha^{-p})$), we do not need to require f to be a L^p function. It inspires the conception of "weaker" L^p functions.

Definition 1.1.1. Let $0 < p < \infty$, f is a measurable function. The weak L^p norm of f and weak L^p space $L^{p,\infty}$ are defined as follows:

$$\|f\|_{p,\infty} := \sup_{\alpha>0} \alpha d_f(\alpha)^{1/p}, \quad L^{p,\infty} = \{f : \|f\|_{p,\infty} < \infty\}.$$

The equivalence in $L^{p,\infty}$ is understood as "equal almost everywhere". Besides, $L^{\infty,\infty} := L^\infty$.

Remark. (a) By definition, we get (1.1.3)-type inequality automatically as follows:

$$d_f(\alpha) \leq \alpha^{-p} \|f\|_{p,\infty}^p$$

In fact, we have $\|f\|_{p,\infty} \leq \|f\|_p$ by (1.1.3), hence $L^p \subset L^{p,\infty}$.

(b) We will not make further discussion on $L^{p,\infty}$(such as its topology), its definition is enough for most cases.

□

Before stating interpolation theorems, let us recall some names for operators enjoying common natures.

Definition 1.1.2. (i) Say $T : L^p \to L^q$ is bounded (not necessarily linear) if and only if

$$\|Tf\|_q \lesssim_{p,q} \|f\|_p, \quad \forall f \in L^p. \tag{1.1.4}$$

T is also called a strong type (p, q) operator, the minimal implicit constant in $\lesssim$ in (1.1.4) is denoted as $\|T\|_{L^p \to L^q}$, or $\|T\|_{p\to q}$ for short.

(ii) Say $T : L^p \to L^{q,\infty}$ is bounded (not necessarily linear) if and only if

$$\|Tf\|_{q,\infty} \lesssim_{p,q} \|f\|_p, \quad \forall f \in L^p. \tag{1.1.5}$$

T is also called a weak type (p, q) operator, the minimal implicit constant in $\lesssim$ in (1.1.5) is denoted as $\|T\|_{L^p\to L^{q,\infty}}$, or $\|T\|_{p\to q,\infty}$ for short.

(iii) Say T (defined on some proper function space) is sub-linear, if and only if

$$|T(f + g)| \le |Tf| + |Tg|, \quad |T(\lambda f)| = |\lambda| \cdot |Tf|, \lambda \in \mathbb{C}.$$

Theorem 1.1.3 (The Marcinkiewicz interpolation theorem)**.** Let $0 < p_0 < p_1 \le \infty$, T is a sub-linear operator defined on $L^{p_0} + L^{p_1}$ (mapping to measurable functions). If T is both weak type (p_0, p_0) and weak type (p_1, p_1), then for any $p_0 < p < p_1$, T is strong type (p, p). More specifically, if

$$\|T\|_{p_0\to p_0,\infty} \le A_0, \quad \|T\|_{p_1\to p_1,\infty} \le A_1.$$

Then for any $p_0 < p < p_1$ with $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ for some $0 < \theta < 1$, we have

$$\|T\|_{p\to p} \lesssim_{p,p_0,p_1} A_0^{1-\theta} A_1^{\theta}.$$

Proof. **Case I:** $p_1 < \infty$. Fix $\alpha > 0$ and $f \in L^p$, decompose f as $f = f_{\delta\alpha} + f^{\delta\alpha}$ for some $\delta > 0$ to be chosen later(here, $f_{\delta\alpha} = f\chi_{\{f\le\delta\alpha\}}, f^{\delta\alpha} = f\chi_{\{f>\delta\alpha\}}$). By the sub-linearity of T, we have

$$d_{Tf}(\alpha) \le d_{T(f^{\delta\alpha})}\left(\frac{\alpha}{2}\right) + d_{T(f_{\delta\alpha})}\left(\frac{\alpha}{2}\right) \le \frac{A_0^{p_0} \left\|f^{\delta\alpha}\right\|_{p_0}^{p_0}}{(\alpha/2)^{p_0}} + \frac{A_1^{p_1} \|f_{\delta\alpha}\|_{p_1}^{p_1}}{(\alpha/2)^{p_1}}.$$

Multiply α^{p-1} on both sides and then integrate with respect to α, we have (using Fubini on the right end) :

$$\|Tf\|_p^p \le p\left(\frac{(2A_0)^{p_0}}{p - p_0}\frac{1}{\delta^{p-p_0}} + \frac{(2A_1)^{p_1}}{p_1 - p}\delta^{p_1-p}\right)\|f\|_p^p.$$

Chose δ to satisfy $\frac{(2A_0)^{p_0}}{p-p_0}\frac{1}{\delta^{p-p_0}}=\frac{(2A_1)^{p_1}}{p_1-p}\delta^{p_1-p}$, we are done.

Case II: $p_1=\infty$**.** Chose $\delta\left(=\frac{1}{2A_1}\right)$ to satisfy

$$\|Tf_{\delta\alpha}\|_\infty\leq A_1\|f_{\delta\alpha}\|_\infty\leq A_1\delta\alpha=\frac{\alpha}{2}.$$

Then

$$d_{Tf}(\alpha)\leq d_{T(f^{\delta\alpha})}\left(\frac{\alpha}{2}\right)\leq\frac{A_0^{p_0}\left\|f^{\delta\alpha}\right\|_{p_0}^{p_0}}{(\alpha/2)^{p_0}}.$$

Hence (by the same method used in Case I):

$$\|Tf\|_p^p\leq p\left(\frac{(2A_0)^{p_0}}{p-p_0}\frac{1}{\delta^{p-p_0}}\right)\|f\|_p^p=\frac{p\,(2A_0)^{p_0}\,(2A_1)^{p-p_0}}{p-p_0}\|f\|_p^p.$$

□

Theorem 1.1.4 (The Riesz-Thorin interpolation theorem)**.** Let $1\leq p_0,q_0,p_1,q_1\leq\infty$, T is a linear operator defined on $S(X)$, the space of simple functions with finite-measure support. If

$$\|T\|_{q_0}\leq A_0\|f\|_{p_0},\quad\|T\|_{q_1}\leq A_1\|f\|_{p_1},\quad\forall f\in S(X).$$

Then for any $0<\theta<1$ and $f\in S(X)$, we have

$$\|Tf\|_q\leq A_0^{1-\theta}A_1^{\theta}\|f\|_p,\qquad\text{where }\left(\frac{1}{p},\frac{1}{q}\right)=(1-\theta)\left(\frac{1}{p_0},\frac{1}{q_0}\right)+\theta\left(\frac{1}{p_1},\frac{1}{q_1}\right).\quad(1.1.6)$$

When $p<\infty$, T can be continuously extended to the whole L^p in an unique way such that (1.1.6) holds for any $f\in L^p$.

Remark. When all index considered is restricted on $[1,\infty)$, the Riesz-Thorin interpolation theorem can be roughly stated as follows: If linear operator T is both strong type (p_0,q_0) and strong type (p_1,q_1) , then for any (p,q) satisfying $\left(\frac{1}{p},\frac{1}{q}\right)=(1-\theta)\left(\frac{1}{p_0},\frac{1}{q_0}\right)+\theta\left(\frac{1}{p_1},\frac{1}{q_1}\right)$ for some $\theta\in(0,1)$, T is strong type (p,q). □

Proof. The proof of this theorem depends on Hadamard three-line Theorem of complex analysis, we omit the details here. □

1.2 Convolution, approximate identities and maximal operators.

We already know two things of convolution on $L^p, 1\leq p\leq\infty$: first, if $f\in L^1, g\in L^p$, then $f*g\in L^p$; second, if $f\in L^p, g\in L^{p'}$ (p' denotes the conjugate index of p), then $f*g$ is a uniformly continuous function.

Especially, convolution can be considered as multiplication on L^1 with, however, no multiplicative unit (that is, there is no $g \in L^1$ such that for any $f \in L^1$, $f * g = f$). This inspires the means of compensation called "approximate identity": using a sequence of functions to mimic the effect of multiplicative unit in the limit sense.

Definition 1.2.1. Call $\{k_\varepsilon\}$ an approximate identity as $\varepsilon \to 0$, if and only if

(i) $\{\|k_\varepsilon\|_1\}$ is bounded, and $\int k_\varepsilon \equiv 1$;

(ii) for any neighborhood V of 0, $\int_{V^c} |k_\varepsilon| \to 0$ as $\varepsilon \to 0$.

Remark. Let $\varphi \in L^1$ and $\int \varphi = 1$, then $\{\varphi_\varepsilon\}$ given by $\varphi_\varepsilon(x) := \varepsilon^{-n}\varphi\left(\varepsilon^{-1}x\right)$ is an approximate identity as $\varepsilon \to 0$. φ_ε is also called the L^1- dilation of φ (since φ_ε and φ share the same integral and L^1-norm). □

Theorem 1.2.2. Let $\{k_\varepsilon\}$ be an approximate identity as $\varepsilon \to 0$.

(i) if $f \in L^p, 1 \le p < \infty$, then $k_\varepsilon * f$ converge to f in L^p;

(ii) if $f \in L^\infty$ and f is uniformly continuous on $A \subset \mathbb{R}^n$, then $k_\varepsilon * f$ converge uniformly to f on A . Especially, if f is bounded and continuous at point x_0, then $(k_\varepsilon * f)(x_0)$ converge to $f(x_0)$.

Proof. (i) By Jensen's inequality,

$$\begin{aligned}
&\int |\left(k_\varepsilon * f\right)(x) - f(x)|^p\, dx \\
\le &\int dx \left|\int |k_\varepsilon(y)||f(x-y) - f(x)|dy\right|^p \\
\le &\|k_\varepsilon\|_1^{p-1} \int dx \int |k_\varepsilon(y)|\,|f(x-y) - f(x)|^p dy \\
= &\|k_\varepsilon\|_1^{p-1} \int \|\tau^y f - f\|_p^p\, |k_\varepsilon(y)|\, dy \\
= &\|k_\varepsilon\|_1^{p-1} \left(\int_V \|\tau^y f - f\|_p^p\, |k_\varepsilon(y)|\, dy + \int_{V^c} \|\tau^y f - f\|_p^p\, |k_\varepsilon(y)|\, dy\right).
\end{aligned}$$

Here V is some neighborhood of 0, $\tau^y f$ denotes the translation of f .

It is easy to see that: the first item can be small enough if V is taken small (by the L^p- continuous of translate), the second item can be small enough if ε is taken small (by the definition of approximate identity).

(ii) when $x \in A$,

$$|\left(k_\varepsilon * f\right)(x) - f(x)| \le \int_V |k_\varepsilon(y)|\,|f(x-y) - f(x)|dy + \int_{V^c} |k_\varepsilon(y)|\,|f(x-y) - f(x)|dy.$$

the first item can be small enough if V is taken small (by the uniform continuous property of f), the second item can be small enough if ε is taken small (by the definition of approximate identity). $\square$

Remark. If $\{k_\varepsilon\}$ enjoy the following properties which are slightly different from that of approximate identity:

(i) $\{\|k_\varepsilon\|_1\}$ is bounded, and $\int k_\varepsilon \equiv a$;

(ii) for any neighborhood V of 0, $\int_{V^c} |k_\varepsilon| \to 0$ as $\varepsilon \to 0$.

Then almost the same argument as before shows that (only the case $a = 0$ needs to be checked additionally):

(i) if $f \in L^p$, $1 \le p < \infty$, then $k_\varepsilon * f$ converge to af in L^p ;

(ii) if $f \in L^\infty$ and f is uniformly continuous on A , then $k_\varepsilon * f$ converge uniformly to af on A . $\square$

By Theorem 1.2.2, if $\{k_\varepsilon\}$ is an approximate identity, then the convolution of $f \in L^p$ with k_ε converge to f in L^p, hence there exists a subsequence converges to f pointwise almost everywhere. Later we shall show that, when $\{k_\varepsilon\}$ enjoy some good properties, the whole sequence $f * k_\varepsilon$ (rather than a subsequence) converge to f pointwise almost everywhere.

Next, let us recall contents of Hardy-Littlewood maximal function/operator.

The Hardy-Littlewood maximal function Mf of $f \in L^1_{loc}$ is defined pointwise as follows:

$$(Mf)(x) := \sup_{r>0} \frac{\int_{B(x,r)} |f|}{|B(x,r)|}. \tag{1.2.1}$$

We already know that:

(i) M is sub-linear and weak type (1,1) . Besides, M is obviously strong type (∞, ∞), and hence strong type (p,p) by The Marcinkiewicz interpolation theorem.

(ii) if $f \in L^1_{loc}$ is not identically zero, then $Mf \notin L^1$, since

$$(Mf)(x) \ge \frac{\int_{B(x,|x|+R)} |f|}{\nu_n(|x|+R)^n} \ge \frac{\int_{B_R} |f|}{\nu_n(|x|+R)^n}, \quad \forall R. \tag{1.2.2}$$

(1.2.2) also tells us if $Mf = 0$ at some point, then f is identically zero.

(iii) if x is a Lebesgue point of f, then $\frac{\int_{B(x,\varepsilon)} f}{|B(x,\varepsilon)|} \to f(x)$. Since almost every point of $\mathbb{R}^n$ are Lebesgue points of $f \in L^1_{loc}$, so

$$\frac{\int_{B(x,\varepsilon)} f}{|B(x,\varepsilon)|} \to f(x) \text{ a.e.} \tag{1.2.3}$$

An important observation is, (1.2.1) can be converted into form of convolution as follows:

$$(Mf)(x) = \sup_{\varepsilon>0} (f * k_\varepsilon)(x), \quad k = \nu_n^{-1}\chi_{B_1}.$$

Here k_ε is the L^1-dilation of k .

By Theorem 1.2.2, we have known that there exists some subsequence of $f * k_\varepsilon$ converging to f pointwise almost everywhere. However, (1.2.3) shows that the whole $f * k_\varepsilon$ converge to f pointwise almost everywhere rather than a subsequence. Thus we are believed that when approximate identity enjoys some good properties, it give us better results concerning convergence pointwise when conducting convolution.

Next, we shall exhibit the standard argument concerning convergence pointwise in harmonic analysis.

Theorem 1.2.3. Let $0 < p, q < \infty$, $\{T_\varepsilon\}$ is a family of linear operators defined on L^p, $T_*f := \sup_{\varepsilon>0} |T_\varepsilon f|$ is well-defined almost everywhere for any $f \in L^p$, D is dense in L^p.

If T_* is weak type (p, q) with norm at most B , and $T_\varepsilon f$ converge to an almost-everywhere-finite function for any $f \in D$. Then $T_\varepsilon f$ converge to an almost-everywhere-finite function for any $f \in L^p$.

Thus we obtain a linear operator $T := \lim_{\varepsilon\to 0} T_\varepsilon$ on L^p, T is also weak type (p, q) since $|Tf| \le T_*f$.

Proof. For simplicity, We may assume that T_*f is well-defined everywhere for every $f \in L^p$. Define the oscillation of f at point y as:

$$O_f(y) := \limsup_{\varepsilon\to 0} \limsup_{\theta\to 0} |(T_\varepsilon f)(y) - (T_\theta f)(y)| .$$

It is easy to see that:

(i) $O_f(y) = 0$ if and only if $\{(T_\varepsilon f)(y)\}$ is Cauchy sequence;

(ii) $O_{f+g}(y) \le O_f(y) + O_g(y)$; (iii) $O_f(y) \le 2(T_*f)(y)$.

we have known that when $f \in D$, $O_f = 0$ $a.e.$ To complete the proof, we just need to show $|\{O_f > \delta\}| = 0$ for any $f \in L^p$ and $\delta > 0$ .Given $\eta > 0$, we first choose $g \in D$ such that $\|f - g\|_p < \eta$, then by the properties of O_f listed before,

$$|\{O_f > \delta\}| \le |\{O_{f-g} > \delta\}| \le \left|\left\{T_*(f-g) > \frac{\delta}{2}\right\}\right| \le \left(\frac{2B\|f-g\|_p}{\delta}\right)^q \le \left(\frac{2B\eta}{\delta}\right)^q.$$

Thus $|\{O_f > \delta\}| = 0$ since η is arbitrary given. $\square$

Let $\varphi \in L^1, \{\varphi_t\}$ is L^1- dilation of φ. We have known from the remark of Theorem 1.2.2 that: $f * \varphi_t$ converge uniformly to $af\ \left(a = \int \varphi\right)$ when $f \in C_c$. If we hope to establish the pointwise convergence property for $f \in L^p$, we need to study the boundness of the maximal operator of $\{f * \varphi_t\}$ via Theorem 1.2.3.

As shown below, when φ satisfies some majority condition, the maximal operator of $\{f * \varphi_t\}$ can be controlled by Hardy-Littlewood maximal operator pointwise, hence is weak type (p, p) .

Theorem 1.2.4. (i) Let $k : [0, \infty) \to [0, \infty)$ be a non-increasing function with possibly finite discontinuous points; $K(x) = k(|x|)$ is the radical function given by K. If $K \in L^1$, then for any $f \in L^1_{loc}$ and $x \in \mathbb{R}^n$, we have

$$\sup_{\varepsilon>0} \left(|f| * K_\varepsilon\right)(x) \leq \|K\|_1 \cdot (Mf)(x). \tag{1.2.4}$$

Here K_ε is the L^1- dilation of K.

(ii) if $K \in L^1$ is controlled by some radical function K_0 induced as in (i) (that is, there exists a non-increasing function $k_0 : [0, \infty) \to [0, \infty)$ with possibly finite discontinuous points such that $K(x) \leq K_0(x) := k_0(|x|)$), then we say K_0 is a integrable non-increasing radical majorant of K .

If there exists a integrable non-increasing radical majorant Φ for $\varphi \in L^1$, then for any $f \in L^1_{loc}$ and $x \in \mathbb{R}^n$, we have

$$\sup_{t>0} |f * \varphi_t|(x) \leq \sup_{t>0} \left(|f| * \Phi_t\right)(x) \leq \|\Phi\|_1 \cdot (Mf)(x).$$

(iii) Let $1 \leq p < \infty$ and there exists a integrable non-increasing radical majorant for $\varphi \in L^1$. then for any $f \in L^p$, $f * \varphi_t$ converge to af pointwise almost everywhere, here $a = \int \varphi$.

Proof. It suffices to prove (i). Without loss of generality, we may assume $f \geq 0$ and k is a simple function:

$$k = a_1 \chi_{[0,N_1)} + \cdots + a_k \chi_{[N_{k-1},N_k)}, \quad a_1 > a_2 > \cdots > a_k > 0.$$

(the general k can be approximated by a sequence of simple functions as above)

We have (denote $N_0 = a_{k+1} = 0$):

$$\left(f * K_\varepsilon\right)(x) = \int_{\mathbb{R}^n} f(x - \varepsilon y) K(y) dy$$

$$
\begin{aligned}
&= \sum_{j=1}^{k} a_j \int_{N_{j-1} \le |y| < N_j} f(x - \varepsilon y) dy \\
&= \sum_{j=1}^{k} a_j \left(\int_{|y| < N_j} f(x - \varepsilon y) dy - \int_{|y| < N_{j-1}} f(x - \varepsilon y) dy \right) \\
&= \sum_{j=1}^{k} (a_j - a_{j+1}) \int_{|y| < N_k} f(x - \varepsilon y) dy \\
&\le \left(\sum_{j=1}^{k} (a_j - a_{j+1}) N_k^n \nu_n \right) (Mf)(x) \\
&= \left(\sum_{j=1}^{k} (a_j - a_{j+1}) |B_{N_k}| \right) (Mf)(x) \\
&= \|K\|_1 \cdot (Mf)(x).
\end{aligned}
$$

$\square$

1.3 Fourier transform

Now we recall contents of Fourier transform. we are limited to list some basic properties of Fourier transform for further use without offering any proof. The basic knowledge of Schwartz space, distribution (as well as tempered distribution) is assumed, see Functional Analysis (written by Rudin) for a thorough introduction to these contents.

On Schwartz space, all kinds of limit progressions can be conducted without any difficulty, thus Fourier transform on $\mathscr{S}$ enjoys perfect properties:

(1) Fourier transform is a linear homomorphism from $\mathscr{S}$ to $\mathscr{S}$;

(2) $\|\hat{f}\|_\infty \le \|f\|_1$;

(3) $\hat{\tilde{f}} = \tilde{\hat{f}}$;

(4) $\hat{\bar{f}} = \bar{\tilde{\hat{f}}}$;

(5) $\widehat{\tau^y f} = e^{-2\pi i y \cdot (\cdot)} \hat{f}, \quad \left(e^{2\pi i y \cdot (\cdot)} f\right)^\wedge = \tau^y \hat{f}$;

(6) $\widehat{\delta^t f} = (\hat{f})_t, \quad (f_t)^\wedge = \delta^t \hat{f}$;

(7) $\widehat{D^\alpha f} = (2\pi i(\cdot))^\alpha \hat{f}, \quad D^\alpha \hat{f} = ((-2\pi i(\cdot))^\alpha f)^\wedge$;

(8) $\widehat{f * g} = \hat{f}\hat{g}, \quad \widehat{fg} = \hat{f} * \hat{g}$;

(9) If A is orthogonal transform, then $\widehat{f \circ A} = \hat{f} \circ A$;

(Thus, if f is radical, then $\hat{f}$ is radical)

(10) $\int f\hat{g} = \int \hat{f} g, \int fh = \int \hat{f}\check{h}, \int f\bar{h} = \int \hat{f}\overline{\hat{h}}$,

(Thus, $\|f\|_2 = \|\check{f}\|_2 = \|\hat{f}\|_2$)

(11) $(\hat{f})^\vee = (\check{f})^\wedge = f, \quad \hat{\hat{f}} = \tilde{f}$.

(12) Fourier transform on $\mathscr{S}$ has fix points such as Gaussian $e^{-\pi|x|^2}$.

Fourier transform can be defined on L^1 via a specific formula, and is injective from L^1 to C_0 . When $f, \hat{f}$ are both in L^1 , we have the Fourier inversion formula to recover f from $\hat{f}$.

By Plancherel, Fourier transform can be extended to a L^2- isometry .

When $1 < p < 2$, $f \in L^p$ can be decompose into $f_1 + f_2, f \in L^1, f_2 \in L^2$ as is known before, and thus the Fourier transform can be defined reasonably as $\widehat{f_1} + \widehat{f_2}$. It is obviously that such definition does not depend on specific decomposition. Besides, the Fourier transform on L^p is strong type (p, p') with norm at most 1 by The Riesz-Thorin interpolation theorem.

When $p > 2$, the Fourier transform of $f \in L^p$ should be understood in the sense of tempered distribution. Generally, the Fourier transform of tempered distribution has the following properties dual to those of Fourier transform on Schwartz space:

Remark. Recall that we can conduct a few operators on tempered distributions just as on Schwartz functions, such as: derivation, multiplying by appropriate function, translation, dilation, reflection(more generally, composed with an invertible matrix), Fourier transform, Fourier inverse transform, convolution with Schwartz function. More specifically(in the following, $u \in \mathscr{S}', \varphi, \phi \in \mathscr{S}$, h is a C^∞ function with at most polynomial growth and the same is true for all of its derivatives):

$$\langle D^\alpha u, \phi\rangle := (-1)^{|\alpha|} \langle u, D^\alpha \phi\rangle, \quad \langle hu, \phi\rangle := \langle u, h\phi\rangle;$$
$$\langle \tau^t u, \phi\rangle := \langle u, \tau^{-t}\phi\rangle, \quad \langle \delta^a u, \phi\rangle := \langle u, \phi_a\rangle;$$
$$\langle \tilde{u}, \phi\rangle := \langle u, \tilde{\phi}\rangle, \quad \langle u \circ A, \phi\rangle := |\det A|^{-1} \langle u, \phi \circ A^{-1}\rangle, A \in GL(\mathbb{R}^n);$$
$$\langle \hat{u}, \phi\rangle := \langle u, \hat{\phi}\rangle, \quad \langle \check{u}, \phi\rangle := \langle u, \check{\phi}\rangle, \quad \langle u * \varphi, \phi\rangle := \langle u, \tilde{\varphi} * \phi\rangle.$$

we know that $u * \varphi$ can also be view as a C^∞ function with at most polynomial growth at infinity : $(u * \varphi)(x) = u(\tau^x \tilde{\varphi})$.

Besides, when u has compact support(hence u is in fact continuous functional on C^∞), the Fourier transform of u can also be viewed as an real analytic function defined on $\mathbb{R}^n$,that is , $\hat{u}(x) := u(e^{-2\pi i x})$. □

(in the following, $u \in \mathscr{S}'$ and $f \in \mathscr{S}$)

(1) Fourier transform is a linear homomorphism from $\mathscr{S}'$ to $\mathscr{S}'$;

(2) $\hat{\tilde{u}} = \tilde{\hat{u}}$;

(3) $\widehat{\tau^y u} = e^{-2\pi i(\cdot)\cdot y}\widehat{u}, \quad \left(e^{2\pi i(\cdot)\cdot y}u\right)^\wedge = \tau^y \hat{u}$;

(4) $\widehat{\delta^t u} = (\widehat{u})_t, \quad (u_t)^\wedge = \delta^t \widehat{u}$;

(5) $\widehat{D^\alpha u} = (2\pi i\xi)^\alpha \hat{u}, \quad D^\alpha \hat{u} = ((-2\pi i x)^\alpha u)^\wedge$;

(6) $\widehat{u * f} = \hat{f}\hat{u}, \quad \widehat{fu} = \hat{f} * \hat{u}$;

(7) $(\hat{u})^\vee = (\check{u})^\wedge = u, \quad \hat{\hat{u}} = \tilde{u}$.

2 Singular integrals of convolution type

2.1 Hilbert transform

Hilbert transform is the simplest model for the theory of singular integral.

Definition 2.1.1. Denoted by W_0 the following (1-dimensional) tempered distribution:

$$W_0(\varphi) := \lim_{\varepsilon\to 0}\int_{|x|\geq\varepsilon}\frac{\varphi(x)}{x}dx, \quad \varphi \in \mathscr{S}(\mathbb{R})$$

(W_0 is called the principal value of $\frac{1}{x}$, and can be denoted by $p.v.\frac{1}{x}$.)

The Hilbert transform of $f \in \mathscr{S}(\mathbb{R})$ is defined to be the convolution of W_0 with f (as a smooth function with polynomial growth at infinity), slightly differs in an absolute constant $\frac{1}{\pi}$. That is:

$$Hf := \frac{1}{\pi}W_0 * f$$

Of course, we can give explicit formula of Hf :

$$Hf := \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|y|\geq\varepsilon}\frac{f(x-y)}{y}dy = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|x-y|\geq\varepsilon}\frac{f(y)}{x-y}dy \tag{2.1.1}$$

Inspired by (2.1.1), we have the following conception of "truncated Hilbert transform":

$$H^{(\varepsilon)}(f)(x) := \frac{1}{\pi}\int_{|y|\geq\varepsilon}\frac{f(x-y)}{y}dy = \frac{1}{\pi}\int_{|x-y|\geq\varepsilon}\frac{f(y)}{x-y}dy$$

Thus $Hf = \lim_{\varepsilon\to 0} H^{(\varepsilon)}(f)$ in the pointwise sense.

Remark. (a) W_0 is really a tempered distribution, since

$$W_0(\varphi) = \lim_{\varepsilon\to 0}\int_{1\geq|x|\geq\varepsilon}\frac{\varphi(x)-\varphi(0)}{x}dx + \int_{|x|\geq 1}\frac{\varphi(x)}{x}dx \leq \|\varphi'\|_\infty + C\|x\varphi\|_\infty.$$

(b) The Fourier transform of W_0 can be computed directly (using Fubini and dominate convergence theorem):

$$\widehat{W_0} = -i \cdot \mathrm{sgn}$$

Hence $\widehat{Hf} = \widehat{W_0 * f} = (-i \cdot \mathrm{sgn})\hat{f}$, which implies H is an isometry on $\mathscr{S}$. Thus we can naturally extend the definition of the Hilbert transform to $L^2(\mathbb{R})$.

(c) if we use $\lim_{\varepsilon \to 0} H^{(\varepsilon)}(f)$ to define Hilbert transform, we can extend naturally extend the definition of the Hilbert transform to some other functions such as integrable functions satisfying local Holder condition.

Here, say f satisfies local Holder condition means: for almost all x , there exists a neighborhood δ_x and constant C_x,$\varepsilon_x > 0$, such that

$$|f(y) - f(x)| \le C_x |y - x|^{\varepsilon_x}, \quad \forall y \in \delta_x.$$

In particular, we can conduct Hilbert transform on characteristic function $f = \chi_{[0,1]}$, and by a direct computation, $(Hf)(x) = \frac{1}{\pi} \log \left| \frac{x}{x-1} \right|$. This example shows Hilbert transform does not map L^1into L^1, or L^∞ into L^∞. □

Next, we shall show the L^p boundness of Hilbert transform when $1 \le p < \infty$. To explore the weak type (1,1) of Hilbert transform, we need a useful technique named Calderon-Zygmund decomposition.

Theorem 2.1.2 (Calderon-Zygmund decomposition of L^1)**.** We can associate with $f \in L^1$ and $\alpha > 0$ a collection of disjoint dyadic cubes $\mathcal{Q}(f, \alpha) = \{Q_j\}_j$ such that:

(i) For any $Q_j \in \mathcal{Q}(f, \alpha)$, we have $\alpha < \frac{1}{|Q_j|} \int_{Q_j} |f| \le 2^n \alpha$;

(ii) For almost all $x \notin U_j Q_j$, we have $|f(x)| \le \alpha$.

$\mathcal{Q}(f, \alpha)$ is in fact uniquely determined by f and α (since $\mathcal{Q}(f, \alpha)$ can be described as the collection of maximal dyadic cubes on which the integral mean of f is $> \alpha$), thus we can associate with each $Q_j \in \mathcal{Q}(f, \alpha)$ a function supported on Q_j with mean zero as follows:

$$b_j := \left(f - \frac{1}{|Q_j|} \int_{Q_j} f \right) \chi_{Q_j}.$$

Furthermore, we denote by

$$b := \sum_j b_j, \quad g := f - b = \begin{cases} f, & x \notin \bigcup_j Q_j \\ \frac{1}{|Q_j|} \int_{Q_j} f, & x \in Q_j \end{cases}.$$

It is easy to check the following properties of the decomposition of $f = g + b$ (called the Calderon-Zygmund decomposition of f at height α):

(i) $g \in L^1 \cap L^\infty, \|g\|_1 \leq \|f\|_1, \|g\|_\infty \leq 2^n \alpha$;

(Thus, by interpolation formula of L^p norm, $\|g\|_p \leq 2^{\frac{n}{p'}} \alpha^{\frac{1}{p'}} \|f\|_1^{\frac{1}{p}}$)

(ii) $\|b_j\|_1 \leq 2^{n+1}\alpha \left|Q_j\right|, \|b\|_1 \leq 2\|f\|_1$;

(iii) $|\cup_j Q_j| \leq \alpha^{-1}\|f\|_1$.

Theorem 2.1.3. Hilbert transform is weak type (1,1) and strong type (p, p) for $1 < p < \infty$.

Proof. (i) Given $\alpha > 0$, let $f = g + b$ be the Calderon-Zygmund decomposition of f at height α. By the linearity of Hilbert transform, we have:

$$|\{Hf > \alpha\}| \leq \left|\left\{Hg > \frac{\alpha}{2}\right\}\right| + \left|\left\{Hb > \frac{\alpha}{2}\right\}\right|.$$

For the term concerning g , we have (via the isometry of Hilbert transform):

$$\left|\left\{Hg > \frac{\alpha}{2}\right\}\right| \leq \frac{4}{\alpha^2}\|Hg\|_2^2 = \frac{4}{\alpha^2}\|g\|_2^2 \leq \frac{8}{\alpha}\|f\|_1.$$

For the term concerning b , we shall use some geometry tricks. Denote by $\{I_j\}$ the collection of disjoint dyadic intervals associated with f and α as in Theorem 2.1.2, let $2I_j$ be the interval with the same center c_j as I_j and twice the length, $\Omega^* = \cup_j 2I_j$. We can write:

$$\left|\left\{Hb > \frac{\alpha}{2}\right\}\right| \leq |\Omega^*| + \left|\left\{x \notin \Omega^* : (Hb)(x) > \frac{\alpha}{2}\right\}\right|$$

The term $|\Omega^*|$ is easy to deal with:

$$|\Omega^*| \leq 2\left|\bigcup_j I_j\right| \leq \frac{2}{\alpha}\|f\|_1.$$

The second term need some careful geometry analysis:

$$\begin{aligned}
\left|\left\{x \notin \Omega^* : (Hb)(x) > \frac{\alpha}{2}\right\}\right| &\leq \frac{2}{\alpha}\int_{(\cup_j 2I_j)^c} |Hb|(x)dx \\
&\leq \frac{2}{\alpha}\int_{(\cup_j 2I_j)^c} \sum_j \left|Hb_j\right|(x)dx \\
&\leq \frac{2}{\alpha}\sum_j \int_{(2I_j)^c} \left|Hb_j\right|(x)dx \\
&= \frac{2}{\alpha}\sum_j \int_{(2I_j)^c} \left|\lim_{\varepsilon\to 0}\int_{|x-y|\geq\varepsilon, y\in I_j} \frac{b_j(y)}{x-y}dy\right| dx
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{\alpha}\sum_j \int_{(2I_j)^c} \left| \int_{I_j} \frac{b_j(y)}{x-y} dy \right| dx \\
&= \frac{2}{\alpha}\sum_j \int_{(2I_j)^c} \left| \int_{I_j} b_j(y) \left(\frac{1}{x-y} - \frac{1}{x-c_j} \right) dy \right| dx \\
&\le \frac{2}{\alpha}\sum_j \int_{I_j} |b_j(y)|\, dy \int_{(2I_j)^c} \frac{|y-c_j|}{|x-y|\,|x-c_j|} dx \\
&\le \frac{C}{\alpha}\sum_j \int_{I_j} |b_j(y)|\, dy \int_{(2I_j)^c} \frac{|I_j|}{|x-c_j|^2} dx \\
&\le \frac{C}{\alpha}\sum_j \int_{I_j} |b_j(y)|\, dy \int_{(2I_j)^c} \frac{|I_j|}{|x-c_j|^2} dx \\
&= \frac{C}{\alpha}\sum_j \int_{I_j} |b_j(y)|\, dy = \frac{C}{\alpha}\|b\|_1 \le \frac{C}{\alpha}\|f\|_1.
\end{aligned}
$$

To sum up, $|\{Hf > \alpha\}| \le \frac{C}{\alpha}\|f\|_1$ for some absolute constant C, which proves the weak type (1,1) estimate.

(ii) Since Hilbert transform is an isometry (hence strong type (2,2)), by the Marcinkiewicz interpolation theorem, it is strong type (p,p) for $1 < p < 2$. As for $p > 2$ we just use the argument of duality.

Note that the adjoint operator H^* is induced by the multiplier $i \cdot \mathrm{sgn}$, hence

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

So, for $p > 2$,

$$
\begin{aligned}
\|Hf\|_p &= \sup\left\{ \left| \int (Hf)g \right| : g \in \mathscr{S}, \|g\|_{p'} \le 1 \right\} \\
&= \sup\left\{ \left| \int f(Hg) \right| : g \in \mathscr{S}, \|g\|_{p'} \le 1 \right\} \\
&\le \|f\|_p \cdot \sup\{\|Hg\|_{p'} : g \in \mathscr{S}, \|g\|_{p'} \le 1\} \\
&\le C\|f\|_p.
\end{aligned}
$$

Which proves the strong type (p,p) estimate for $2 < p < \infty$. □

Thanks to Theorem 2.1.3, we can extend the Hilbert transform to functions of L^p when $1 \le p < \infty$ as a L^p bounded operator.

2.2 Riesz transform

Definition 2.2.1. Denoted by W_j the following ($n-$ dimensional) tempered distribution:

$$\langle W_j, \varphi \rangle := \lim_{\varepsilon \to 0} \int_{|y| \geq \varepsilon} \frac{y_j}{|y|^{n+1}} \varphi(y) dy, \quad \varphi \in \mathscr{S}(\mathbb{R}^n)$$

The Riesz transform of $f \in \mathscr{S}(\mathbb{R}^n)$ is defined to be the convolution of W_j with f, slightly differs in an absolute constant depend on the dimension n. That is:

$$R_j f := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} W_j * f$$

Of course, we can give explicit formula of $R_j f$ as well:

$$(R_j f)(x) := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \lim_{\varepsilon \to 0} \int_{|x-y| \geq \varepsilon} \frac{x_j - y_j}{|x-y|^{n+1}} f(y) dy \tag{2.2.1}$$

Remark. (a) W_j is really a tempered distribution, mainly because

$$\begin{aligned} \int_{1 \geq |y| \geq \varepsilon} \frac{y_j}{|y|^{n+1}} (\varphi(y) - \varphi(0)) dy &= \int_\varepsilon^1 r^{n-1} dr \int_{\mathbb{S}^{n-1}} \frac{r\theta_j}{r^{n+1}} (\varphi(r\theta) - \varphi(0)) d\theta \\ &\leq \int_\varepsilon^1 dr \int_{\mathbb{S}^{n-1}} \frac{|\varphi(r\theta) - \varphi(0)|}{r} d\theta \leq C_n \|\varphi'\|_\infty . \end{aligned}$$

(b) The Fourier transform of W_j can also be computed directly similarly:

$$\widehat{W_j} = -i \cdot \frac{\xi_j}{|\xi|}$$

Hence $\widehat{R_j f} = \widehat{W_j * f} = \left(-i \frac{\xi_j}{|\xi|}\right) \hat{f}$, which implies R_j is L^2 bounded. Thus, we can naturally extend the definition of the Riesz transform to $L^2(\mathbb{R}^n)$.

(c) if we use (2.2.1) to define Riesz transform, we can as well extend naturally the definition of the Hilbert transform to integrable functions satisfying local Holder condition. $\square$

2.3 Calderon-Zygmund operator

Definition 2.3.1. $K : \mathbb{R}^n \backslash \{0\} \to \mathbb{C}$ is called a Calderon-Zygmund kernel if and only if K satisfy the following three conditions for some constant B possibly depend on dimension n:

(i) Size condition:

$$|K(x)| \leq B|x|^{-n}, \quad \forall x \neq 0. \tag{2.3.1}$$

(ii) Smoothness condition:

$$\int_{|x|\geq 2|y|} |K(x-y)-K(x)|dx \leq B, \quad \forall y \neq 0. \tag{2.3.2}$$

Or more strongly (called Hormander condition):

$$|\nabla K(x)| \leq B|x|^{-n-1}, \quad \forall x \neq 0. \tag{2.3.3}$$

(iii) Cancellation condition:

$$\int_{r<|x|<s} K(x)dx = 0, \quad \forall 0 < r < s < \infty. \tag{2.3.4}$$

The Calderon-Zygmund operator with kernel K (satisfying (2.3.1), (2.3.2),(2.3.4) is defined as:

$$(Tf)(x) = \lim_{\varepsilon\to 0} \int_{|x-y|\geq\varepsilon} K(x-y)f(y)dy, \quad \forall f \in \mathscr{S}(\mathbb{R}^n). \tag{2.3.5}$$

If K satisfying (2.3.1), (2.3.3), (2.3.4) (that is, satisfying stronger smoothness condition), then the operator defined as in (2.3.5) is called strong Calderon-Zygmund operator.

Remark. (2.3.3) is truly stronger than (2.3.2), since (assume (2.3.3) is satisfied):

$$\begin{aligned}
\int_{|x|\geq 2|y|} |K(x-y)-K(x)|dx &= \int_{|x|\geq 2|y|} dx \int_0^1 |\nabla K(x-ty)\cdot(-y)|dt \\
&\leq B\int_0^1 dt \int_{|x|\geq 2|y|} \frac{|y|}{|x-ty|^{n+1}}dx \\
&\leq 2^{n+1}B\int_0^1 dt \int_{|x|\geq 2|y|} \frac{|y|}{|x|^{n+1}}dx \\
&= C_n B.
\end{aligned}$$

(note that if $|x| \geq 2|y|, 0 \leq t \leq 1$, then $|x-ty| \geq \frac{|x|}{2}$) $\square$

We shall proof the L^p boundness of Calderon-Zygmund operator:

Theorem 2.3.2. If T is a Calderon-Zygmund operator with kernel K , then T is weak type (1,1), and strong type (p, p) for any $1 < p < \infty$.

Proof. Similar to what we have done to Hilbert transform, we shall first prove the L^2 boundness of T by the method of Fourier transform, then L^1 weak-type boundness of T by CalderonZygmund decomposition, then the general L^p boundness of T by interpolation and duality.

(i) L^2 boundness of T, $\|T\|_{L^2\to L^2} \le C_n B$.

Denoted by $K^{(r,s)} = K\chi_{\{r<|x|<s\}}$. It suffices to prove that

$$\sup_{0<r<s<\infty} \left\|\widehat{K^{(r,s)}}\right\|_\infty = \sup_{0<r<s<\infty,\xi} \left|\int_{r<|x|<s} K(x)e^{-2\pi i x\cdot\xi}\right| \le C_n B.$$

Since

$$\begin{aligned}\|Tf\|_2 &= \left\|\lim_{r\to 0, s\to\infty} K^{(r,s)} * f\right\|_2 \le \liminf_{r\to 0, s\to\infty} \left\|K^{(r,s)} * f\right\|_2 \\ &= \liminf_{r\to 0, s\to\infty} \left\|\widehat{K^{(r,s)}}\hat{f}\right\|_2 \le C_n B\|f\|_2.\end{aligned}$$

$\xi = 0$ is of no consideration by cancellation condition, we only need to consider the following three cases:

Case I: $r < |\xi|^{-1} < s$.

$$\begin{aligned}\int_{r<|x|<s} K(x)e^{-2\pi i x\cdot\xi} &= \int_{r<|x|<|\xi|^{-1}} K(x)e^{-2\pi i x\cdot\xi}dx + \int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i x\cdot\xi}dx \\ &:= I_1 + I_2.\end{aligned}$$

For I_1 , we use cancellation condition, size condition and an elementary inequality to obtain:

$$\begin{aligned}|I_1| &= \left|\int_{r<|x|<|\xi|^{-1}} K(x)\left(e^{-2\pi i x\cdot\xi} - 1\right)dx\right| \le 2\pi|\xi| \int_{r<|x|<|\xi|^{-1}} |x|^{-n+1}dx \\ &\le 2\pi B|\xi| \int_{|x|<|\xi|^{-1}} |x|^{-n+1}dx = C_n B.\end{aligned}$$

For the second term, we first use an averaging technical to split it into several easy-to-tackle terms (put $z_\xi = \frac{\xi}{2|\xi|^2}$) :

$$I_2 = -\int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i\left(x+z_\xi\right)\cdot\xi}dx = -\int_{|\xi|^{-1}<\left|x-z_\xi\right|<s} K\left(x-z_\xi\right)e^{-2\pi i x\cdot\xi}dx.$$

Hence:

$$\begin{aligned}2I_2 &= \int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i x\cdot\xi}dx - \int_{|\xi|^{-1}<\left|x-z_\xi\right|<s} K\left(x-z_\xi\right)e^{-2\pi i x\cdot\xi}dx \\ &= \int_{|\xi|^{-1}<\left|x-z_\xi\right|<s} \left[K(x) - K\left(x-z_\xi\right)\right]e^{-2\pi i x\cdot\xi}dx \\ &\quad + \int_{|\xi|^{-1}<|x|<s} K(x)e^{-2\pi i x\cdot\xi}dx - \int_{|\xi|^{-1}<\left|x-z_\xi\right|<s} K(x)e^{-2\pi i x\cdot\xi}dx\end{aligned}$$

$$
\begin{aligned}
&= \int_{|\xi|^{-1}\le\left|x-z_\xi\right|\le s}\left[K(x)-K\left(x-z_\xi\right)\right]e^{-2\pi i x\cdot\xi}dx \\
&\quad+\int_{\substack{|\xi|^{-1}<|x|<s\\ \left|x-z_\xi\right|<|\xi|^{-1}}}K(x)e^{-2\pi i x\cdot\xi}dx+\int_{\substack{|\xi|^{-1}<|x|<s\\ \left|x-z_\xi\right|>s}}K(x)e^{-2\pi i x\cdot\xi}dx \\
&\quad-\int_{\substack{|\xi|^{-1}<|x-z_\xi|<s\\ |x|<|\xi|^{-1}}}K(x)e^{-2\pi i x\cdot\xi}dx-\int_{\substack{|\xi|^{-1}<|x-z_\xi|<s\\ |x|>s}}K(x)e^{-2\pi i x\cdot\xi}dx
\end{aligned}
$$

By size condition, smooth condition (note that $|z_\xi| = 2|\xi|^{-1}$) and some geometry observations, we obtain (for example):

$$
\left|\int_{\substack{|\xi|^{-1}<|x|<s\\ \left|x-z_\xi\right|>s}}K(x)e^{-2\pi i x\cdot\xi}dx\right| \le B\int_{\frac{s}{2}<|x|<s}|x|^{-n}dx = C_nB.
$$

By now, we have completed the prove of case I.

Case II: $|\xi|^{-1} \le r$.

$$
\begin{aligned}
\left|\int_{r<|x|<s}K(x)e^{-2\pi i x\cdot\xi}\right| &= \left|\int_{|\xi|^{-1}<|x|<s}K(x)e^{-2\pi i x\cdot\xi}dx-\int_{|\xi|^{-1}<|x|<r}K(x)e^{-2\pi i x\cdot\xi}dx\right| \\
&\le \left|\int_{|\xi|^{-1}<|x|<s}K(x)e^{-2\pi i x\cdot\xi}dx\right|+\left|\int_{|\xi|^{-1}<|x|<r}K(x)e^{-2\pi i x\cdot\xi}dx\right|.
\end{aligned}
$$

However, we have known from the estimate of I_2 in case I that:

$$
\sup_s\left|\int_{|\xi|^{-1}<|x|<s}K(x)e^{-2\pi i x\cdot\xi}dx\right| \le C_nB,
$$

which complete the proof of case II.

Case III: $|\xi|^{-1} \ge s$.

$$
\begin{aligned}
\left|\int_{r<|x|<s}K(x)e^{-2\pi i x\cdot\xi}\right| &\le \left|\int_{r<|x|<|\xi|^{-1}}K(x)e^{-2\pi i x\cdot\xi}dx-\int_{s<|x|<|\xi|^{-1}}K(x)e^{-2\pi i x\cdot\xi}dx\right| \\
&\le \left|\int_{r<|x|<|\xi|^{-1}}K(x)e^{-2\pi i x\cdot\xi}dx\right|+\left|\int_{s<|x|<|\xi|^{-1}}K(x)e^{-2\pi i x\cdot\xi}dx\right|
\end{aligned}
$$

However, we have known from the estimate of I_1 in case I that:

$$
\sup_r\left|\int_{r<|x|<|\xi|^{-1}}K(x)e^{-2\pi i x\cdot\xi}dx\right| \le C_nB,
$$

which complete the proof of case III.

(ii) weak type (1,1) of T, $\|T\|_{L^1\to L^{1,\infty}} \le C_nB$.

Given $\alpha > 0$. let $f = g + b$ be the Calderon-Zygmund decomposition of f at height α. By the linearity of T, we have:

$$
|\{Tf>\alpha\}| \le \left|\left\{Tg>\frac{\alpha}{2}\right\}\right|+\left|\left\{Tb>\frac{\alpha}{2}\right\}\right|.
$$

For the term concerning g, we have (via the L^2 boundness of T):

$$\left|\left\{Tg > \frac{\alpha}{2}\right\}\right| \leq \frac{4}{\alpha^2}\|Tg\|_2^2 \leq \frac{C_nB}{\alpha^2}\|g\|_2^2 \leq \frac{C_nB}{\alpha}\|f\|_1.$$

For the term concerning b , we shall use some geometry tricks. Denote by $\{Q_j\}$ the collection of disjoint dyadic cubes associated with f and α as in the C-Z decomposition, let $Q_j^* = c_nQ_j$ (explained as in the proof of the C-Z decomposition, here c_n is some large constant, depending on dimension n , denote by d_j the center of Q_j) , $\Omega^* = \cup_j Q_j^*$. We can write:

$$\left|\left\{Tb > \frac{\alpha}{2}\right\}\right| \leq |\Omega^*| + \left|\left\{x \notin \Omega^* : (Tb)(x) > \frac{\alpha}{2}\right\}\right|$$

The term $|\Omega^*|$ is easy to deal with:

$$|\Omega^*| \leq \left|\bigcup_j Q_j^*\right| \leq \frac{C_n}{\alpha}\|f\|_1.$$

The second term need some careful geometry analysis:

$$\begin{aligned}
&\left|\left\{x \notin \Omega^* : (Tb)(x) > \frac{\alpha}{2}\right\}\right| \\
&\leq \frac{2}{\alpha}\int_{\left(\cup_j Q_j^*\right)}{}^{c}|Tb|(x)dx \leq \frac{2}{\alpha}\int_{\left(\cup_j Q_j^*\right)}\sum_j |Tb_j|\,(x)dx \\
&\leq \frac{2}{\alpha}\sum_j \int_{\left(Q_j^*\right)} |Tb_j|\,(x)dx = \frac{2}{\alpha}\sum_j \int_{\left(Q_j^*\right)^c}\left|\int_{Q_j} K(x-y)b_j(y)dy\right| dx \\
&= \frac{2}{\alpha}\sum_j \int_{\left(Q_j^*\right)^c}\left|\int_{Q_j} b_j(y)\left(K(x-y) - K\left(x-d_j\right)\right)dy\right| dx \\
&\leq \frac{2}{\alpha}\sum_j \int_{Q_j} |b_j(y)|\,dy \int_{\left(Q_j^*\right)} |K(x-y) - K\left(x-d_j\right)|\,dx \\
&\leq \frac{C_nB}{\alpha}\sum_j \int_{Q_j} |b_j(y)|\,dy = \frac{C_nB}{\alpha}\|b\|_1 \leq \frac{C_nB}{\alpha}\|f\|_1.
\end{aligned}$$

To ensure the correctness of the above argument, we need c_n to be large enough such that $|x-d_j| \geq 2\,|y-d_j|$ when $x \notin Q_j^*, y \in Q_j$, besides, $|Tb_j|\,(x)$ has the explicit formula when $x \notin Q_j^*$:

$$|Tb_j|\,(x) = \int_{Q_j} K(x-y)b_j(y)dy. \tag{2.3.6}$$

Note that b_j is bounded (since f is bounded) and (2.3.6) holds when b_j is replaced by smooth function supported on (say) $2Q_j$. So, it is not hard to prove (2.3.6) by the process of function approximating.

To sum up, $|\{Tf > \alpha\}| \leq \alpha^{-1} C_n B \|f\|_1$, which proves the weak type (1,1) estimate.

(iii) The general L^p boundness of T when $1 < p < \infty$.

By the Marcinkiewicz interpolation theorem, T is strong type (p, p) for $1 < p < 2$. As for $p > 2$ we just use the argument of duality.

Note that the adjoint operator T^* is induced by the kernel $\bar{\tilde{K}}$ (which is still a Calderon-Zygmund kernel), So, for $p > 2$,

$$\begin{aligned}
\|Tf\|_p &= \sup\left\{\left|\int (Tf)g\right| : g \in \mathscr{S}, \|g\|_{p'} \leq 1\right\} \\
&= \sup\left\{\left|\int f\,(T^*g)\right| : g \in \mathscr{S}, \|g\|_{p'} \leq 1\right\} \\
&\leq \|f\|_p \cdot \sup\left\{\|T^*g\|_{p'} : g \in \mathscr{S}, \|g\|_{p'} \leq 1\right\} \\
&\leq C_{n,p} B \|f\|_p.
\end{aligned}$$

Which proves the strong type (p, p) estimate for $2 < p < \infty$. □

Remark. Review the proof of Theorem 2.3.2, we notice that the weak type (1,1) of T depend only on the smoothness condition of K and the "represent property" illustrated by (2.3.6). □

2.4 Fractional integration

It is known that the Poisson equation $\Delta u = f$ in $\mathbb{R}^n$ has a solution of the form when $n \geq 3$:

$$u(x) = c_n \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-2}} dy$$

Such solution is an example of the so-called fractional integration.

Definition 2.4.1. Associate $0 < \alpha < n$ and $f \in \mathscr{S}(\mathbb{R}^n)$ with the following function:

$$(I_\alpha f)(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}} dy, \quad x \in \mathbb{R}^n \tag{2.4.1}$$

I_α is also called the Riesz potential operator. (Note that (2.4.1) is well-defined for all $x \in \mathbb{R}^n$)

Remark. (a) Let us explain the confusing name "fractional integral". The Laplace operator Δ can be rewritten via the Fourier transform as follows:

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

We can formally exchange the exponent 2 in to a general exponent β, and define

$$\left((-\Delta)^{\frac{\beta}{2}} f\right)^{\wedge} := (2\pi|\xi|)^{\beta} \hat{f}.$$

Equivalently:

$$(-\Delta)^{\frac{\beta}{2}} f := \left((2\pi|\xi|)^{\beta}\right)^{\vee} * f = C_{n,\beta}|\xi|^{-n-\beta} * f$$

Here we use the computational result of the Fourier transform of $|x|^{\beta}$.

Thus, I_α is just $C_{n,\alpha}(-\Delta)^{-\frac{\alpha}{2}}$, the generation of the Laplace operator.

(b) It is easy to check the following basic properties of I_α :

$$I_\alpha \left(I_\beta f\right) = I_{\alpha+\beta} f, \quad \alpha, \beta > 0, \alpha + \beta < n;$$

$$\Delta \left(I_\alpha f\right) = I_\alpha(\Delta f) = -I_{\alpha-2}(f), \quad n \geq 3, 2 \leq \alpha < n.$$

□

We hope to establish the following estimate:

$$\|I_\alpha f\|_q \precsim \|f\|_p \tag{2.4.2}$$

By the standard dilation argument, the necessary condition under which (2.4.2) holds is

$$\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n} \tag{2.4.3}$$

We show that (2.4.3) is also the sufficient condition for (2.4.2) .

Theorem 2.4.2. Let $0 < \alpha < n$, $1 \leq p < q < \infty$ satisfying (2.4.3).

(i) if $f \in L^p$, then the integral defined as in (2.4.1) converge absolutely for a.e. x;

(ii) I_α is strong type (p, q) when $p \neq 1$, and weak type (p, q) when $p = 1$.

Proof. Denoted by $K(x) = |x|^{\alpha-n}, K_1 = K\chi_{\{|x| \leq R\}}, K_\infty = K\chi_{\{|x|>R\}}$ for R to be chosen later.

(i) Since $K_1 \in L^1, K_\infty \in L^{p'}$, we have $I_\alpha f = K * f = K_1 * f + K_\infty * f$ converge absolutely a.e. whenever $f \in L^p$.

(ii) since $\|K_\infty\|_{p'} = C_{p,n,\alpha} R^{-\frac{n}{q}}, \|K_1\|_1 = C_{p,n,\alpha} R^{\alpha}$, we have the following pointwise estimate of $I_\alpha f$ via (1.2.4) (note that K_1 is non-increasing and radical):

$$\left|\left(I_\alpha f\right)(x)\right| \leq \left|\left(K_1 * f\right)(x)\right| + \left|\left(K_\infty * f\right)(x)\right|$$

$$\leq \|K_1\|_1 (Mf)(x) + \|K_\infty\|_{p'} \|f\|_p$$
$$\leq C_{p,n,\alpha} \left(R^{-\frac{n}{q}} (Mf)(x) + R^\alpha \|f\|_p \right).$$

Choose $R = R(x)$ to satisfy:

$$R^{-\frac{n}{q}} (Mf)(x) = R^\alpha \|f\|_p.$$

We finally obtain

$$|I_\alpha f| \leq C_{p,n,\alpha} |Mf|^{\frac{p}{q}} \|f\|_p^{1-\frac{p}{q}}.$$

Then (ii) follows easily from the above pointwise estimate and the L^p boundness of the Hardy-Littlewood maximal operator. $\square$

2.5 Homogenous kernels

We have known from Theorem 2.3.2 that Calderon-Zygmund operator can be defined for any $f \in L^p$. However, this is just an abstract extension process, and we do not have any explicit formula for the L^p case. Note that when K is a Calderon-Zygmund kernel, for any fixed $\varepsilon > 0$,

$$T^{(\varepsilon)}(f)(x) := \int_{|x-y|\geq\varepsilon} K(x-y) f(y) dy$$

is well-defined everywhere when $f \in L^p, 1 \leq p < \infty$.

So, it is natural to ask whether $T^{(\varepsilon)}(f) \to Tf$ a. e . when $f \in L^p$.

Recall the standard argument concerning converge pointwise in harmonic analysis shown in Theorem 1.2.3, we need to consider the L^p boundness of the relevant maximal operator defined as

$$(T^* f)(x) := \sup_{\varepsilon>0} \left| T^{(\varepsilon)}(f)(x) \right| = \sup_{\varepsilon>0} \left| \int_{|x-y|\geq\varepsilon} K(x-y) f(y) dy \right|$$

For simplicity, we are restricted to consider the so-called homogeneous kernels".

Definition 2.5.1. Call $K : \mathbb{R}^n \backslash \{0\} \to \mathbb{C}$ is a homogeneous kernel if and only if K has the form of

$$K(x) = \frac{\Omega\left(\frac{x}{|x|}\right)}{|x|^n} \tag{2.5.1}$$

Where $\Omega : \mathbb{S}^{n-1} \to \mathbb{C}$ is an integrable function with mean zero.

It is not hard to check that homogeneous kernels defined as in (2.5.1) is Calderon-Zygmund kernels (Really?). Next, we shall establish the L^p boundness of maximal operator T^* related to homogeneous kernels defined as in (2.5.1).

Theorem 2.5.2. *T^* is weak type (1,1) and strong type (p,p) when $1<p<\infty$.*

Proof. We shall prove the following pointwise estimate:

$$T^*f \le C_n(M(Tf)+Mf)$$

Here M is the Hardy-Littlewood maximal operator Then the L^p boundness follows easily from that of M and T .

Let $\tilde{K}(x) := K(x)\mathbb{1}_{\{|x|\ge 1\}}$, it's easy to check

$$\tilde{K}_\epsilon(x) = \epsilon^{-n}\tilde{K}\left(\frac{x}{\epsilon}\right) = K(x)\mathbb{1}_{\{|x|\ge\epsilon\}}.$$

Pick a smooth bump function $\varphi \in \mathscr{C}_0^\infty(\mathbb{R}^n)$, $\operatorname{supp}(\varphi)\subset B_1(0)$, $\varphi\ge 0$, $\int_{\mathbb{R}^n}\varphi\,\mathrm{d}x = 1$. Set

$$\Phi := \varphi * K - \tilde{K}$$

where

$$\varphi * K(x) = \lim_{\delta\searrow 0}\int_{|x-y|>\delta} K(x-y)\varphi(y)\,\mathrm{d}y,$$

thus

$$\Phi_\epsilon = (\varphi * K)_\epsilon - \tilde{K}_\epsilon = \varphi_\epsilon * K_\epsilon - \tilde{K}_\epsilon = \varphi_\epsilon * K - \tilde{K}_\epsilon.$$

We claim that for any $f\in L^p(\mathbb{R}^n)$, $1<p<\infty$,

$$(\varphi * K) * f(x) = Tf * \varphi_\epsilon(x)$$

for every $x\in\mathbb{R}^n$. One can formally verify it and we will prove it latter. Assume the claim is right, plugging $\varphi * K = \Phi_\epsilon + \tilde{K}_\epsilon$ into the claim gives

$$\tilde{K}_\epsilon * f = \int_{|y|>\epsilon} K(y) f(x-y)\,\mathrm{d}y = Tf*\varphi_\epsilon - \Phi_\epsilon * f,$$

Notice that $\{\varphi_\epsilon\}_{\epsilon>0}$ and $\{\Phi_\epsilon\}_{\epsilon>0}$ are approximate identities and taking the supremum over $\epsilon>0$, we get

$$T^*f(x) \le \sup_{\epsilon>0}|Tf*\varphi_\epsilon|(x) + \sup_{\epsilon>0}|\Phi_\epsilon * f|(x)$$

and

$$T^{*}f(x) \leq C_n \left(Mf(x) + M(Tf)(x)\right),$$

Now, it's suffice to verify the claim above. We notice first that

$$\left(\varphi_\epsilon * \tilde{K}_\delta\right) * f(x) = T_\delta(f) * \varphi_\epsilon(x)$$

for every $\delta > 0$ because both sides are equal to

$$\int_{\mathbb{R}^n} \int_{|y|>\delta} K(y) f(z-y) \phi_\epsilon(x-z) \, \mathrm{d}y\mathrm{d}z.$$

Moreover $\varphi * \tilde{K}_\delta \to \phi_\epsilon * K$ in L^q norm as $\delta \downarrow 0$, $1 < q < \infty$, and $T_\delta f \to Tf$ in L^p norm as $\delta \downarrow 0$, $1 < p < \infty$. Note that $\varphi \in L^q(\mathbb{R}^n)$ with $1 < q < \infty$, $1/q + 1/p = 1$, by *Hölder*'s inequality, we let $\delta \downarrow 0$, then we get the proof. □

Remark. The proof of the weak-L^1 boundedness of T^* is a variation of that for the same property of T, we will use $Calder\acute{o}n - Zygmund$ decomposition to prove it and you can find that in Elias M. Stein ¡¡S.I.¿¿ page 43-44. □

Notice that $\mathscr{S}(\mathbb{R}^n)$ is a density subspace of $L^p(\mathbb{R}^n)$ and we have proved that for any $f \in \mathscr{S}(\mathbb{R}^n)$, $1 \leq p < \infty$,

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

For any $f \in L^p$, $1 < p < \infty$, choose a Cauchy sequence $\{f_n\}_{n=1}^{\infty}$ in $\mathscr{S}(\mathbb{R}^n)$ satisfies $\|f - f_n\|_{L^p} \to 0$ as $n \to \infty$, then $\{Tf_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^p(\mathbb{R}^n)$, set the limit as Tf, which is not depended on Cauchy sequence we choose. A natural question is that whether we have the following integration representation

$$Tf(x) = p.v. \int_{\mathbb{R}^n} K(x-y) f(y) \, \mathrm{d}y$$

for $\mathcal{L}^n - a.e.\ x \in \mathbb{R}^n$. We can answer this question just using the maximal function T^*f. For any $g \in \mathscr{S}(\mathbb{R}^n)$ and $\alpha > 0$, we have

$$\begin{aligned}
&\left|\left\{x \in \mathbb{R}^n : \left|\limsup_{\epsilon \searrow 0} T_\epsilon f(x) - \liminf_{\epsilon \searrow 0} T_\epsilon f(x)\right| > \alpha\right\}\right| \\
&\leq \left|\left\{x \in \mathbb{R}^n : 2T^*(f-g)(x) > \alpha\right\}\right| \\
&\lesssim_n \frac{\|f-g\|_{L^p}^p}{\alpha^p},
\end{aligned}$$

hence we get

$$\left|\left\{x \in \mathbb{R}^n : \left|\limsup_{\epsilon \searrow 0} T_\epsilon f(x) - \liminf_{\epsilon \searrow 0} T_\epsilon f(x)\right| > 0\right\}\right| = 0,$$

so we get the proof for $1 < p < \infty$, case $p = 1$ is similiar.

2.6 BMO and singular integrals

We have known that s.i.o. T is weak-type $(1,1)$ and strong-type (p,p), $1 < p < \infty$, and Hardy-Littlewood maximal operator M is weak-type $(1,1)$, weak-type (∞,∞) and strong-type (p,p), $1 < p < \infty$, a natural question is that s.i.o. T bounded from $L^\infty \to L^\infty$? The answer is NO! We will introduce $BMO\,(\mathbb{R}^n)$, and one can see s.i.o. T is bounded from $L^\infty \to BMO\,(\mathbb{R}^n)$. Also, one can easily check $L^\infty\,(\mathbb{R}^n) \subsetneq BMO\,(\mathbb{R}^n)$.

Remark. In fact, it turns out that $BMO\,(\mathbb{R}^n)$ is the smallest space that contains $T\,(L^\infty\,(\mathbb{R}^n))$ for every s.i.o. T. You can see it in Schlag Vol.1. page 183 for this fact and see the notes from page 193. $\square$

$BMO\,(\mathbb{R}^n)$, is the space of functions having *bounded mean oscillation* on $\mathbb{R}^n$. The definition about $BMO\,(\mathbb{R}^n)$ as follows.

Definition 2.6.1. $BMO\,(\mathbb{R}^n) := \{f \in L^1_{loc}\,(\mathbb{R}^n) : \|f\|_* < \infty\}$, *where*

$$\|f\|_* = \|f\|_{BMO(\mathbb{R}^n)} := \sup_{Q \subset \mathbb{R}^n} |Q|^{-1} \int_Q |f - f_Q|\,, \quad f_Q = |Q|^{-1} \int_Q f,$$

the supremum is taken over all cubes $Q \subset \mathbb{R}^n$.

By definition and $\triangle-$inequality, we can easily check the following properties,

$$\begin{cases} \|f+g\|_* \leq \|f\|_* + \|g\|_*, \\ \|\lambda f\|_* = |\lambda|\,\|f\|_*,\ \forall \lambda \in \mathbb{C}, \end{cases}$$

hence $BMO\,(\mathbb{R}^n)$ is a linear space.

Remark. $\|\cdot\|_*$ is not a norm but a semi-norm, if $\|f\|_* = 0$, $f = const.\ \mathcal{L}^n - a.e.$ $x \in \mathbb{R}^n$, the constant can be given arbitrarily. Also, the product of two $BMO\,(\mathbb{R}^n)$ functions may not be in $BMO\,(\mathbb{R}^n)$. $\square$

For any $c \in \mathbb{C}$, $(f+c)_Q = f_Q + c$, thus $\|f+c\|_* = \|f\|_*$. If $f, g \in BMO\,(\mathbb{R}^n)$ and $f - g = const.\ \mathcal{L}^n - a.e.\ x \in \mathbb{R}^n$, we identify f with g (think f and g are the same element in $BMO\,(\mathbb{R}^n)$). $BMO\,(\mathbb{R}^n)$ consists of equivalent classes $[f] = \{g \in BMO\,(\mathbb{R}^n) : f - g = const.\ \mathcal{L}^n - a.e.\}$.

Definition 2.6.2 (Sharp function)**.** Let $f \in L^1_{loc}(\mathbb{R}^n)$. Define

$$f^{\#}(x) := \sup_{Q \ni x} |Q|^{-1} \int_Q |f(y) - f_Q| \, \mathrm{d}y,$$

where the supremum is taken over all cubes Q contains x.

The following proposition shows the relationship between $L^\infty(\mathbb{R}^n)$-norm and $BMO(\mathbb{R}^n)$-norm.

Proposition 2.6.3. $L^\infty(\mathbb{R}^n) \subset BMO(\mathbb{R}^n)$ and $\|f\|_{BMO(\mathbb{R}^n)} \leq 2\|f\|_{L^\infty(\mathbb{R}^n)}$.

Proof. Note that $|f_Q| \leq \|f\|_{L^\infty(\mathbb{R}^n)}$, then for any cube $Q \subset \mathbb{R}^n$,

$$|Q|^{-1} \int_Q |f - f_Q| \leq 2\|f\|_{L^\infty(\mathbb{R}^n)},$$

hence $\|f\|_{BMO(\mathbb{R}^n)} \leq 2\|f\|_{L^\infty(\mathbb{R}^n)}$. $\square$

The next proposition gives a equivalent definition about $BMO(\mathbb{R}^n)$.

Proposition 2.6.4. Suppose that there exists an $A > 0$, s.t. for any cube $Q \subset \mathbb{R}^n$, there exists a constant c_Q, s.t.

$$\sup_Q |Q|^{-1} \int_Q |f(y) - c_Q| \, \mathrm{d}y \leq A,$$

then $f \in BMO(\mathbb{R}^n)$ and $\|f\|_* \leq 2A$.

Proof. For any cube $Q \subset \mathbb{R}^n$,

$$\begin{aligned} |f(x) - f_Q| &\leq |f(x) - c_Q| + |f_Q - c_Q| \\ &\leq |f(x) - c_Q| + \left| |Q|^{-1} \int_Q (f(y) - c_Q) \, \mathrm{d}y \right| \\ &\leq |f(x) - c_Q| + |Q|^{-1} \int_Q |f(y) - c_Q| \, \mathrm{d}y \\ &\leq |f(x) - c_Q| + A, \end{aligned}$$

hence

$$|Q|^{-1} \int_Q |f(y) - c_Q| \, \mathrm{d}y \leq 2A.$$

$\square$

Definition 2.6.5 (Equivalent $BMO(\mathbb{R}^n)$-norm)**.** Let $f \in L^1_{loc}(\mathbb{R}^n)$. Define

$$\|f\|'_* := \sup_Q \inf_{c \in \mathbb{C}} |Q|^{-1} \int_Q |f(x) - c| \, \mathrm{d}x,$$

and we call it *equivalent* $BMO(\mathbb{R}^n) - $ *norm*.

In fact, the next proposition shows that $BMO\left(\mathbb{R}^n\right)$-norm has certain equivalent norms.

Proposition 2.6.6. For all $f \in L^1_{loc}\left(\mathbb{R}^n\right)$, we have

$$\frac{1}{2}\|f\|_* \leq \sup_Q \inf_{c\in\mathbb{C}} |Q|^{-1} \int_Q |f(x) - c| \,\mathrm{d}x \leq \|f\|_*.$$

Proof. The right side inequality is obvious, for the left side inequality, for any cube $Q \subset \mathbb{R}^n$ and $c \in \mathbb{C}$, we have

$$|Q|^{-1} \int_Q |f(x) - f_Q| \,\mathrm{d}x \leq 2\,|Q|^{-1} \int_Q |f(x) - c| \,\mathrm{d}x,$$

thus

$$\frac{1}{2}\,|Q|^{-1} \int_Q |f(x) - f_Q| \,\mathrm{d}x \leq \inf_{c\in\mathbb{C}} |Q|^{-1} \int_Q |f(x) - c| \,\mathrm{d}x,$$

so we get the proof. □

Theorem 2.6.7. $f \in BMO\left(\mathbb{R}^n\right)$ iff $\forall Q$, $\exists c_Q > 0$, s.t.

$$\sup_Q |Q|^{-1} \int_Q |f(x) - c_Q| \,\mathrm{d}x < \infty.$$

Proof. Just set $c_Q = f_Q$ we can prove one side, on the other side, we can see

$$\|f\|_* \leq 2 \sup_Q |Q|^{-1} \int_Q |f(x) - c_Q| \,\mathrm{d}y < \infty.$$

□

Proposition 2.6.8. If $f \in BMO\left(\mathbb{R}^n\right)$ and $h \in \mathbb{R}^n$, $\tau^h f(x) := f(x-h) \in BMO\left(\mathbb{R}^n\right)$ and $\|\tau^h f\|_* = \|f\|_*$.

Proof. For any cube $Q \subset \mathbb{R}^n$, let $Q_h := Q - h$, then

$$|Q|^{-1} \int_Q \left|\tau^h f - \left(\tau^h f\right)_Q\right| = |Q_h|^{-1} \int_{Q_h} |f - f_{Q_h}|\,,$$

taking the supremum over all cubes we get the proof. □

Proposition 2.6.9. If $f \in BMO\left(\mathbb{R}^n\right)$ and $\lambda > 0$, $\delta^\lambda f(x) := f(\lambda x) \in BMO\left(\mathbb{R}^n\right)$ and $\|\delta^\lambda f\|_* = \|f\|_*$.

Proof. For any cube $Q \subset \mathbb{R}^n$, let λQ be the cube which has the same center of Q but $l(\lambda Q) = \lambda l(Q)$, then

$$|Q|^{-1} \int_Q \left|\delta^\lambda f - \left(\delta^\lambda f\right)_Q\right| = |\lambda Q|^{-1} \int_{\lambda Q} |f - f_{\lambda Q}|\,,$$

taking the supremum over all cubes we get the proof. □

Actually, one can use balls to define a similar linear space like $BMO(\mathbb{R}^n)$, we call it $BMO_{balls}(\mathbb{R}^n)$ and it's defined as follows.

Definition 2.6.10. For $f \in L^1_{loc}(\mathbb{R}^n)$, define

$$\|f\|_{BMO_{balls}} := \sup_B |B|^{-1} \int_B |f - f_B|,$$

where the supremum is taken over all balls B in $\mathbb{R}^n$, and $f_B = |B|^{-1} \int_B f$.

Theorem 2.6.11. For $f \in L^1_{loc}(\mathbb{R}^n)$, then $\exists c_n,\ C_n > 0$, s.t.

$$c_n \|f\|_{BMO} \leq \|f\|_{BMO_{balls}} \leq C_n \|f\|_{BMO}.$$

Proof. Given any cube $Q \subset \mathbb{R}^n$, let B be the smallest ball that contains it. Then $|B| / |Q| = 2^{-n} |B_1(0)| \sqrt{n^n}$ and

$$\begin{aligned} |Q|^{-1} \int_Q |f(x) - f_B| \, \mathrm{d}x &\leq \frac{|B|}{|Q|} \frac{1}{|B|} \int_B |f(x) - f_B| \, \mathrm{d}x \\ &= 2^{-n} |B_1(0)| \sqrt{n^n} \frac{1}{|B|} \int_B |f(x) - f_B| \, \mathrm{d}x \\ &\leq C_n \|f\|_{BMO_{balls}}, \end{aligned}$$

here we set $C_n := 2^{-n} |B_1(0)| \sqrt{n^n}$. Taking the supremum over all cubes we have

$$\|f\|_{BMO} \leq 2 \sup_Q |Q|^{-1} \int_Q |f(x) - f_B| \, \mathrm{d}x \leq 2C_n \|f\|_{BMO_{balls}}.$$

Given any ball $B \subset \mathbb{R}^n$, similiarly we can prove the right side inequality. □

Proposition 2.6.12. $L^\infty(\mathbb{R}^n) \subsetneq BMO(\mathbb{R}^n)$, the function $\log|x|$ is a typical element of $BMO(\mathbb{R}^n)$.

Proof. Notice that $\log|x|$ is not essentially bounded, so it's suffice to verify $\log|x|$ is in $BMO(\mathbb{R}^n)$. To prove this, for every $x_0 \in \mathbb{R}^n$ and $R > 0$, define

$$C_{x_0,R} = \begin{cases} \log|x_0| & if\ \ |x_0| > 2R, \\ \log R & if\ \ |x_0| \leq 2R. \end{cases}$$

Indeed if $|x_0| > 2R$, then

$$|B_R(x_0)|^{-1} \int_{B_R(x_0)} |\log|x| - C_{x_0,R}| \, \mathrm{d}x = |B_R(x_0)|^{-1} \int_{B_R(x_0)} \left| \log \frac{|x|}{|x_0|} \right| \mathrm{d}x \leq \log 2.$$

Also, if $|x_0| \leq 2R$, then

$$\begin{aligned}
|B_R(x_0)|^{-1} \int_{B_R(x_0)} |\log|x| - C_{x_0,R}| \,\mathrm{d}x &= |B_R(x_0)|^{-1} \int_{B_R(x_0)} \left|\log \frac{|x|}{R}\right| \mathrm{d}x \\
&\leq |B_R(x_0)|^{-1} \int_{B_{3R}(0)} \left|\log \frac{|x|}{R}\right| \mathrm{d}x \\
&= |B_1(0)|^{-1} \int_{B_3(0)} |\log|x|| \,\mathrm{d}x.
\end{aligned}$$

Thus $\log|x|$ is in $BMO(\mathbb{R}^n)$. $\square$

Proposition 2.6.13. For $f \in L^1_{loc}(\mathbb{R}^n)$, then $\||f|\|_* \leq 2\|f\|_*$.

Proof. Given any cube $Q \subset \mathbb{R}^n$,

$$\left||f|_Q - |f_Q|\right| = |Q|^{-1} \left|\int_Q (|f| - |f_Q|)\right| \leq |Q|^{-1} \int_Q |f - f_Q| \leq \|f\|_*,$$

hence we have

$$|Q|^{-1} \int_Q \left||f|_Q - |f|\right| = |Q|^{-1} \int_Q \left||f|_Q - |f_Q| + |f_Q| - |f|\right| \leq 2\|f\|_*.$$

$\square$

Remark. Let $f_n(x) = n\mathbb{1}_{(0,\infty)} - n\mathbb{1}_{(-\infty,0)}$, then we have $\|f_n\|_* \geq n$ and $\||f_n|\|_* = 0$, so it's impossible to get inequality like $\underline{\|f\|_* \lesssim \||f|\|_*}$. $\square$

Theorem 2.6.14. For $f, g \in BMO(\mathbb{R}^n)$ and are real-valued, then $\max(f,g)$, $\min(f,g) \in BMO(\mathbb{R}^n)$. Moreover, we have

$$\|\max(f,g)\|_*,\ \|\min(f,g)\|_* \leq \frac{3}{2}(\|f\|_* + \|g\|_*).$$

Proof. Notice that

$$\begin{cases} \max(f,g) &= \frac{(f+g)+|f-g|}{2}, \\ \min(f,g) &= \frac{(f+g)-|f-g|}{2}, \end{cases}$$

thus we have

$$\|\min(f,g)\|_*,\ \|\max(f,g)\|_* \leq \frac{1}{2}(\|f\|_* + \|g\|_*) + \frac{1}{2}\||f-g|\|_* \leq \frac{3}{2}(\|f\|_* + \|g\|_*).$$

$\square$

Theorem 2.6.15. $(BMO(\mathbb{R}^n), \|\cdot\|_*)$ is completed.

Proof. For any given Cauchy sequence $\{f_n\} \subset BMO(\mathbb{R}^n)$ and for any cube $Q \subset \mathbb{R}^n$, we have

$$|Q|^{-1} \int_Q \left|(f_n - f_m) - \left((f_n)_Q - (f_m)_Q\right)\right| \leq \|f_n - f_m\|_* \rightarrow 0 \ as \ n, m \rightarrow \infty,$$

so $\left\{f_n - (f_n)_Q\right\}$ is a Cauchy sequence in $L^1(Q)$. Notice that $L^1(Q)$ is a Banach space, there exists a function $f^Q \in L^1(Q)$ s.t.

$$\left\|f^Q - \left(f_n - (f_n)_Q\right)\right\|_{L^1(Q)} \rightarrow 0 \ as \ n \rightarrow \infty.$$

By Riesz theorem, there exists a subsequence of $\left\{f_n - (f_n)_Q\right\}$ s.t. this subsequence converges to f^Q in Q pointwisely. WLOG, take $\left\{f_n - (f_n)_Q\right\}$ as this subsequence. By D.C.T., we have $\int_Q f^Q = 0$. Now, fix Q and for any cube $Q' \subset \mathbb{R}^n$ satisfies $Q \subset Q'$, we have

$$\left|(f_n)_{Q'} - (f_n)_Q\right| \leq \frac{|Q'|}{|Q|} \sup_{n \in \mathbb{N}^+} \|f_n\|_* \lesssim \frac{|Q'|}{|Q|},$$

hence by Weierstrass theorem there exists $c_{Q,Q'} \in \mathbb{C}$ and a subsequence of $\{(f_n)_{Q'} - (f_n)_Q\}$ s.t. this subsequence converges to $c_{Q,Q'}$, also, WLOG, take $\left\{(f_n)_{Q'} - (f_n)_Q\right\}$ as this subsequence. Thus we have

$$f_n - (f_n)_Q = f_n - (f_n)_{Q'} + (f_n)_{Q'} - (f_n)_Q \ in \ Q,$$

and on the left side we have

$$f_n - (f_n)_Q \rightarrow f^Q \quad \mathcal{L}^n - a.e. \ as \ n \rightarrow \infty,$$

and on the right side we have

$$f_n - (f_n)_{Q'} + (f_n)_{Q'} - (f_n)_Q \rightarrow f^{Q'} + c_{Q,Q'} \quad \mathcal{L}^n - a.e. \ as \ n \rightarrow \infty,$$

hence for $\mathcal{L}^n - a.e. \ x \in Q$, we have

$$f^Q(x) = f^{Q'}(x) + c_{Q,Q'}.$$

Choose $\{Q_j\}$ satisfy $Q_j \nearrow \mathbb{R}^n$, we have defined $\{f^{Q_j}\}$ and for $j < j'$, $c_{Q_j,Q_{j'}} = f^{Q_j} - f^{Q_{j'}}$ $\mathcal{L}^n - a.e.$ on Q_j.(denote $c_{Q_1,Q_1} = 0$) Notice that

$$\begin{cases} f^{Q_1} = f^{Q_j} + c_{Q_1,Q_j} & \mathcal{L}^n - a.e. \text{ on } Q_1; \\ f^{Q_1} = f^{Q_{j'}} + c_{Q_1,Q_{j'}} & \mathcal{L}^n - a.e. \text{ on } Q_1; \\ f^{Q_j} = f^{Q_{j'}} + c_{Q_j,Q_{j'}} & \mathcal{L}^n - a.e. \text{ on } Q_j, \end{cases}$$

hence we get

$$c_{Q_1,Q_{j'}} - c_{Q_1,Q_j} = c_{Q_j,Q_{j'}}.$$

Thus if we define

$$f(x) = f^{Q_j}(x) + c_{Q_1,Q_j} \quad if\ x \in Q_j\ for\ some\ j,$$

we get a well-defined L^1_{loc} function f, and for any $Q \subset \mathbb{R}^n$, take $j >> 1$ s.t. $Q \subset Q_j$, hence there exists a constant c_{Q,Q_j} s.t. $f^Q = f^{Q_j} + c_{Q,Q_j}$ $\mathcal{L}^n - a.e.$ on Q and $f^Q = f - c_{Q_1,Q_j} + c_{Q,Q_j}$ $\mathcal{L}^n - a.e.$ on Q. This means f is the limit of Cauchy sequence $\{f_n\}$ in $BMO(\mathbb{R}^n)$ because by Fatou's lemma, we have

$$\begin{aligned}
|Q|^{-1}\int_Q \left|(f_n - f) - (f_n - f)_Q\right| &= |Q|^{-1}\int_Q \left|\left(f_n - f^Q\right) - \left(f_n - f^Q\right)_Q\right| \\
&\leq \limsup_{m\to\infty} |Q|^{-1}\int_Q \left|(f_n - f_m) - \Big((f_n)_Q - (f_m)_Q\Big)\right| \\
&\leq \limsup_{m\to\infty} \|f_n - f_m\|_* \to 0\ as\ n \to \infty.
\end{aligned}$$

□

Theorem 2.6.16. Let T be a Calderón-Zygmund s.i.o.. Then

$$\|Tf\|_* \leq CB\|f\|_{L^\infty(\mathbb{R}^n)},\ \forall f \in L^2(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n).$$

Proof. WLOG, assmue $B = 1$. Fix some f and a ball $B_0 = B_R(x_0) \subset \mathbb{R}^n$ and define

$$C_{B_0} := \int_{|y-x_0|>2R} K(x_0 - y) f(y)\,\mathrm{d}y,$$

it's well-defined by Hölder inequality and $f \in L^2(\mathbb{R}^n)$. Then, with $B_0^* = B_{2R}(x_0)$, we have

$$\begin{aligned}
&\int_{B_0} |Tf(x) - C_{B_0}|\,\mathrm{d}x \\
&= \int_{B_0} \left|T\left(\mathbb{1}_{(B_0^*)^c} f + \mathbb{1}_{B_0^*} f\right)(x) - C_{B_0}\right| \mathrm{d}x \\
&\leq \int_{B_0} \left|T\left(\mathbb{1}_{(B_0^*)^c} f\right)(x) - C_{B_0}\right| \mathrm{d}x + \int_B \left|T\left(\mathbb{1}_{B_0^*} f\right)(x)\right| \mathrm{d}x \\
&= \int_{B_0} \left|\int_{|y-x_0|>2R} K(x-y) f(y)\,\mathrm{d}y - \int_{|y-x_0|>2R} K(x_0-y) f(y)\,\mathrm{d}y\right| \mathrm{d}x \\
&\quad + \int_{B_0} \left|T\left(\mathbb{1}_{B_0^*} f\right)(x)\right| \mathrm{d}x \\
&\leq \int_{B_0}\int_{|y-x_0|>2R} |K(x-y) - K(x_0-y)|\,|f(y)|\,\mathrm{d}y\mathrm{d}x + \int_{B_0} \left|T\left(\mathbb{1}_{B_0^*} f\right)(x)\right| \mathrm{d}x \\
&\leq |B_0|\,\|f\|_{L^\infty(\mathbb{R}^n)} + c\,|B_0|^{1/2}\,\|\mathbb{1}_{B_0^*} f\|_{L^2(\mathbb{R}^n)} \\
&\leq C\,|B_0|\,\|f\|_{L^\infty(\mathbb{R}^n)},
\end{aligned}$$

this is because

$$T\left(\mathbb{1}_{\left(B_0^*\right)^c}f\right)(x)=\lim_{\epsilon\searrow 0}\int_{|x-y|>\epsilon}K\left(x-y\right)f\left(y\right)\mathbb{1}_{\left(B_0^*\right)^c}\left(y\right)\mathrm{d}y$$

exists in L^2-norm, hence

$$T\left(\mathbb{1}_{\left(B_0^*\right)^c}f\right)(x)=\int_{|y-x_0|>2R}K\left(x-y\right)f\left(y\right)\mathrm{d}y\quad if\ x\in B_0.$$

Thus we get

$$\|Tf\|_*\leq C\|f\|_{L^\infty(\mathbb{R}^n)}.$$

□

Remark. See more details in *Elias M. Stein* ¡¡S.I.¿¿ page 35 theorem 2. □

2.7 John-Nirenberg inequalities

First, we give two versions about the John-Nirenberg inequalities, they are equivalent.

Theorem 2.7.1 (John-Nirenberg inequality, version A)**.** For all $f\in BMO\left(\mathbb{R}^n\right)$, for all cube Q and $\lambda>0$, we have

$$\left|\left\{x\in Q:\left|f\left(x\right)-f_Q\right|>\lambda\right\}\right|\leq c_1\left|Q\right|e^{-c_2\frac{\lambda}{\|f\|_*}},$$

where $c_1,\ c_2$ are absolute constants. In particular, we can take $c_1=e,\ c_2=\left(2^ne\right)^{-1}$.

Theorem 2.7.2 (John-Nirenberg inequality, version B)**.** $\exists c_1,\ c_2>0$, s.t. $\forall f\in BMO\left(\mathbb{R}^n\right)$,

$$\sup_Q\frac{1}{|Q|}\int_Q e^{\frac{c_2}{\|f\|_*}\left|f(x)-f_Q\right|}\mathrm{d}x\leq c_1.$$

Proof of Version B $\Rightarrow$ Version A. Let $E_\lambda=\left\{x\in Q:\left|f\left(x\right)-f_Q\right|>\lambda\right\}$. By version B, we have

$$e^{c_2\frac{\lambda}{\|f\|_*}}E_\lambda=\int_{E_\lambda}e^{c_2\frac{\lambda}{\|f\|_*}}\mathrm{d}x\leq\int_{E_\lambda}e^{\frac{c_2}{\|f\|_*}\left|f(x)-f_Q\right|}\mathrm{d}x\leq\int_Q e^{\frac{c_2}{\|f\|_*}\left|f(x)-f_Q\right|}\mathrm{d}x\leq c_1\left|Q\right|.$$

□

Proof of Version A $\Rightarrow$ Version B. Set $a = \frac{c_2}{2}$, for any cube Q, by Fubini-Tonelli theorem, we have

$$\begin{aligned}
\int_Q e^{\frac{a}{\|f\|_*}\left|f(x)-f_Q\right|}\mathrm{d}x &= \int_{(0,\infty)} \left|\left\{x \in Q : e^{\frac{a}{\|f\|_*}\left|f(x)-f_Q\right|} > t\right\}\right| \mathrm{d}t \\
&= \left(\int_{(0,1)} + \int_{(1,\infty)}\right) \left|\left\{x \in Q : \frac{a}{\|f\|_*}\left|f(x) - f_Q\right| > \log t\right\}\right| \mathrm{d}t \\
&\leq |Q| + \int_{(1,\infty)} c_1 |Q| \, e^{-c_2 \frac{\lambda}{\|f\|_*} \frac{\|f\|_*}{a} \log t} \mathrm{d}t \\
&= |Q| + c_1 |Q| \int_{(1,\infty)} t^{-2} \mathrm{d}t \\
&\leq (1 + c_1) |Q| .
\end{aligned}$$

□

Proof of the theorem 2.7.2 John – Nirenberg inequality, version B. $\forall N > 0$, define

$$f_N(x) = \begin{cases} N & if \ f(x) \geq N \\ f(x) & if \ |f(x)| < N \\ -N & if \ f(x) \leq -N, \end{cases}$$

then $f_N(x) = \max\{-N, \min\{N, f(x)\}\} = \min\{N, \max\{-N, f(x)\}\}$, thus we have

$$\|f_N\|_* \leq \frac{3}{2}\left(\|-N\|_* + \|\min\{N, f\}\|_*\right) = \frac{3}{2}\|\min\{N, f\}\|_* \leq \frac{9}{4}\left(\|N\|_* + \|f\|_*\right) = \frac{9}{4}\|f\|_*.$$

Here we divide the proof into two steps.

First step: Assume $f \in L^\infty(\mathbb{R}^n)$, let Q_0 be any given cube in $\mathbb{R}^n$ and set

$$\triangle(Q_0) := \{all\ children\ of\ Q_0,\ produced\ by\ bisectly\ sides\ of\ Q_0\}.$$

$\forall Q,\ \tilde{Q} \in \triangle(Q_0)$ with $Q \subset \tilde{Q}$ and $l(Q) = \frac{1}{2} l\left(\tilde{Q}\right)$,

$$\left|f_Q - f_{\tilde{Q}}\right| = \left|\frac{1}{|Q|}\int_Q f - f_{\tilde{Q}}\right| = \left|\frac{1}{|Q|}\int_Q \left(f - f_{\tilde{Q}}\right)\right| \leq \frac{2^n}{\left|\tilde{Q}\right|}\int_{\tilde{Q}} \left|f - f_{\tilde{Q}}\right| \leq 2^n \|f\|_*.$$

Let $\alpha = 2\|f\|_*$, since $(f - f_{Q_0}) \mathbb{1}_{Q_0} \in L^1(\mathbb{R}^n)$ and

$$\frac{1}{|Q_0|}\left|(f - f_{Q_0}) \mathbb{1}_{Q_0}\right| \leq \|f\|_* < \alpha,$$

apply the Calderón-Zygmund decomposition to $(f - f_{Q_0}) \mathbb{1}_{Q_0}$ and level α, we get disjoint cube $\{Q_i\}$, $Q_i \in \triangle(Q_0)$, s.t.

1. $\forall i$, $2\|f\|_*= \alpha \leq \frac{1}{|Q_i|}\int_{Q_i}|(f-f_{Q_0})\mathbb{1}_{Q_0}|$;

2. $\mathcal{L}^n - a.e.\ x \in Q_0 \setminus \bigcup_i Q_i$, $|(f-f_{Q_0})\mathbb{1}_{Q_0}|(x) \leq \alpha = 2\|f\|_*$;

3. $|\bigcup_i Q_i| \leq \frac{1}{\alpha}\sum_i \int_{Q_i}|(f-f_{Q_0})\mathbb{1}_{Q_0}| \leq \frac{1}{\alpha}\int_{Q_0}|f-f_{Q_0}| \leq \frac{|Q_0|}{\alpha}\|f\|_*= \frac{|Q_0|}{2}$.

If $\tilde{Q}_i \in \triangle(Q_0)$ satisfying $Q_i \subset \tilde{Q}_i$ and $l(Q_i) = \frac{1}{2}l\left(\tilde{Q}_i\right)$, by Calderón-Zygmund decomposition,

$$2\|f\|_*= \alpha \geq \left|\tilde{Q}_i\right|^{-1}\int_{\tilde{Q}_i}|(f-f_{Q_0})\mathbb{1}_{Q_0}|.$$

Since

$$\begin{aligned}|f_{Q_i}-f_{Q_0}| &\leq \left|f_{Q_i}-f_{\tilde{Q}_i}\right| + \left|f_{\tilde{Q}_i}-f_{Q_0}\right| \\ &\leq 2^n\|f\|_*+\left|\tilde{Q}_i\right|^{-1}\int_{\tilde{Q}_i}|f-f_{Q_0}| \\ &\leq (2^n+2)\|f\|_*,\end{aligned}$$

denote $A(\lambda) = \sup_Q |Q|^{-1}\int_Q \exp\left(\frac{\lambda}{\|f\|_*}|f(x)-f_Q|\right)\mathrm{d}x \leq \exp\left(\frac{2\lambda\|f\|_{L^\infty(\mathbb{R}^n)}}{\|f\|_*}\right) < \infty$, this is because we assume $f \in L^\infty(\mathbb{R}^n)$. For any given cube $Q_0 \subset \mathbb{R}^n$,

$$\begin{aligned}&\frac{1}{|Q_0|}\int_{Q_0}\exp\left(\frac{\lambda}{\|f\|_*}|f(x)-f_{Q_0}|\right)\mathrm{d}x \\ &\leq \frac{1}{|Q_0|}\left(\int_{Q_0\setminus(\bigcup_i Q_i)} + \int_{\bigcup_i Q_i}\right)\exp\left(\frac{\lambda}{\|f\|_*}|f(x)-f_{Q_0}|\right)\mathrm{d}x \\ &=: I + II.\end{aligned}$$

For I, we have

$$|I| \leq \frac{1}{|Q_0|}\int_{Q_0\setminus(\bigcup_i Q_i)} e^{\frac{\lambda}{\|f\|_*}\alpha}\mathrm{d}x \leq e^{2\lambda},$$

for II, we have

$$\begin{aligned}|II| &\leq \frac{1}{|Q_0|}\sum_i \int_{Q_i} e^{\frac{c_2}{\|f\|_*}\left|f(x)-f_{Q_i}+f_{Q_i}-f_{Q_0}\right|}\mathrm{d}x \\ &\leq \frac{1}{|Q_0|}\sum_i |Q_i|^{-1}\int_{Q_i} e^{\frac{c_2}{\|f\|_*}\left|f(x)-f_{Q_i}\right|}\mathrm{d}x\,|Q_i|\,e^{\lambda(2^n+2)} \\ &\leq e^{\lambda(2^n+2)}\frac{1}{|Q_0|}\sum_i |Q_i|\,A(\lambda) \\ &\leq \frac{e^{\lambda(2^n+2)}}{2}A(\lambda).\end{aligned}$$

Hence,

$$\frac{1}{|Q_0|}\int_{Q_0}\exp\left(\frac{\lambda}{\|f\|_*}|f(x)-f_{Q_0}|\right)\mathrm{d}x \leq e^{2\lambda} + \frac{e^{\lambda(2^n+2)}}{2}A(\lambda),$$

take "sup", we get

$$A(\lambda) \leq e^{2\lambda} + \frac{e^{\lambda(2^n+2)}}{2} A(\lambda),$$

so

$$\left(1 - \frac{e^{\lambda(2^n+2)}}{2}\right) A(\lambda) \leq e^{2\lambda},$$

by taking λ sufficiently small, then

$$A(\lambda) \leq \frac{e^{2\lambda}}{1 - \frac{1}{2} e^{\lambda(2^n+2)}}.$$

Notice that the right side is independent of $\|f\|_*$ and $\|f\|_{L^\infty(\mathbb{R}^n)}$.

Second step: $\forall f \in BMO(\mathbb{R}^n)$ and $\forall N \in \mathbb{N}$, let

$$f_N(x) = \min\{N, \max\{-N, f(x)\}\},$$

then $f_N \in L^\infty(\mathbb{R}^n)$. So $\exists \lambda > 0,\ c > 0$, s.t.

$$\sup_Q \frac{1}{|Q|} \int_Q e^{\frac{\lambda}{\|f_N\|_*} \left|f_N(x) - (f_N)_Q\right|} \mathrm{d}x \leq c,$$

and

$$LHS \geq \frac{1}{|Q|} \int_Q e^{\frac{\lambda}{\frac{9}{4}\|f\|_*} \left|f_N(x) - (f_N)_Q\right|} \mathrm{d}x$$

for all cube $Q \subset \mathbb{R}^n$. $\forall$ fixed $Q \subset \mathbb{R}^n$, let $N \nearrow \infty$ and by Fatou's lemma, we have

$$\frac{1}{|Q|} \int_Q e^{\frac{\frac{4}{9}\lambda}{\|f\|_*} \left|f(x) - f_Q\right|} \mathrm{d}x \leq c.$$

□

As a direct consequence of the John-Nirenberg inequalities, we get the following corollary.

Corollary 2.7.3. If $f \in BMO(\mathbb{R}^n)$ and $1 < p < \infty$,

$$\|f\|_* \asymp \sup_Q \left(\frac{1}{|Q|} \int_Q |f(x) - f_Q|^p \,\mathrm{d}x\right)^{\frac{1}{p}} =: \|f\|_{*,p}.$$

Proof. On the one side, by Hölder's inequality, we have

$$\|f\|_* = \sup_Q \frac{1}{|Q|} \int_Q |f(x) - f_Q| \,\mathrm{d}x \leq \sup_Q \left(\frac{1}{|Q|} \int_Q |f(x) - f_Q|^p \,\mathrm{d}x\right)^{\frac{1}{p}}.$$

On the other side, by Fubini-Tonelli theorem and the John-Nirenberg inequalities, we have

$$\int_Q |f(x) - f_Q|^p \,\mathrm{d}x = p \int_{(0,\infty)} |\{x \in Q : |f(x) - f_Q| > \lambda\}| \, \lambda^{p-1} \mathrm{d}\lambda$$

$$\leq pc_1 |Q| \int_{(0,\infty)} \lambda^{p-1} e^{-\frac{c_2\lambda}{\|f\|_*}} \mathrm{d}\lambda,$$

let $y = c_2\lambda/\|f\|_*$, we get

$$\int_Q |f(x) - f_Q|^p \, \mathrm{d}x \leq pc_1 |Q| \left(\frac{\|f\|_*}{c_2}\right)^p \int_{(0,\infty)} y^{p-1} e^{-y} \mathrm{d}\lambda,$$

thus

$$\left(\frac{1}{|Q|} \int_Q |f(x) - f_Q|^p \, \mathrm{d}x\right)^{\frac{1}{p}} \leq c_p \|f\|_*.$$

□

Next theorem shows that BMO is useful in interpolation theory, and we will not prove that. One can see it in Schlag Vol.1. page 183 for the proof.

Theorem 2.7.4. Assume $1 < p_0 < \infty$. T is a linear operator bounded from $L^{p_0}(\mathbb{R}^n)$ to $L^{p_0}(\mathbb{R}^n)$ and bounded from $L^\infty(\mathbb{R}^n)$ to $BMO(\mathbb{R}^n)$. Then T is bounded from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$, $\forall p_0 < q < \infty$.

2.8 Almost orthogonality

We have proved L^2-boundedness of Calderón-Zygmund operators by Fourier transforms. Actually, one can avoid Fourier transforms by Cotlar's lemma(or Cotlar-Stein lemma). Before we introduce it, let's review some results in functional analysis.

Proposition 2.8.1. $\mathscr{H}$ is a Hilbert space. $T \in \mathcal{L}(\mathscr{H}, \mathscr{H})$. $\exists$ a unique $T^* \in \mathcal{L}(\mathscr{H}, \mathscr{H})$, adjoint of T, s.t.

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

$\forall x, y \in \mathscr{H}$, and $\|T^*\| = \|T\|$, $\|T^*T\| = \|T\|^2$, $T^{**} = T$. Moreover, for any $a, b \in \mathbb{C}$, $S, T \in \mathcal{L}(\mathscr{H}, \mathscr{H})$,

$$\begin{cases} (aS + bT)^* = \overline{a}S^* + \overline{b}T^*, \\ (ST)^* = T^*S^*. \end{cases}$$

Here we talk about a special Hilbert space $L^2(\mathbb{R}^n)$.

Lemma 2.8.2 (The Cotlar-Stein lemma)**.** Let $\{T_j\}_{j=1}^N$ be finitely many bounded operetors on $L^2(\mathbb{R}^n)$, s.t. for some function $\gamma : \mathbb{Z} \to \mathbb{R}^+$, one has $\|T_j^*T_k\| \leq \gamma(j-k)^2$, $\|T_jT_k^*\| \leq \gamma(j-k)^2$ for any $1 \leq j, k \leq N$. Let $\sum_{l\in\mathbb{Z}} \gamma(l) = A < \infty$. Then

$$\left\|\sum_{j=1}^N T_j\right\| \leq A.$$

Remark. Because $\|T_j^* T_k\| \leq \gamma(j-k)^2$ for any $1 \leq j,k \leq N$. Let $j=k$, then we have $\|T_j\|^2 = \|T_j^* T_j\| = \gamma(0)^2$, thus $\|T_j\| \leq \gamma(0) \leq A$, $\forall 1 \leq j \leq N$. Consider the following trivial estimate

$$\|\sum_{j=1}^{N} T_j\| \leq \sum_{j=1}^{N} \|T_j\| \leq N\gamma(0) \leq AN.$$

But by Cotlar-Stein lemma, we have

$$\left\| \sum_{j=1}^{N} T_j \right\| \leq A.$$

□

Remark. Back to s.i.o. kernel $K(x)$, by a dyadic decomposition of the kernel K, for its related s.i.o. T, which defined as follows,

$$Tf(x) = p.v. \int_{\mathbb{R}^n} f(x-y) K(y) \, \mathrm{d}y,$$

we can decompose it like $T = \sum_{j \in \mathbb{Z}} T_j$ and T_j defined as follows

$$\begin{aligned} T_j : L^2(\mathbb{R}^n) &\to L^2(\mathbb{R}^n) \\ f \mapsto T_j f(x) &:= \int_{\mathbb{R}^n} f(x-y) K(y) \phi_j(y) \, \mathrm{d}y, \end{aligned}$$

there ϕ_j satisfies $\operatorname{supp}(\phi_j) \subset annulus$. We may can use the $Cotlar-Stein$ lemma for $\sum_{j=-N}^{N} T_j$ if we have some nice estimates about $\|T_j^* T_k\|$ for any $-N \leq j,k \leq N$. □

Proof of the Cotlar-Stein lemma 2.8.2. Denote $T = \sum_{j=1}^{N} T_j$. For any positive integer n, we have

$$(T^* T)^n = \sum_{\substack{j_1,\ldots,j_n = 1 \\ k_1,\ldots,k_n = 1}}^{N} T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n},$$

hence

$$\|(T^* T)^n\| \leq \sum_{\substack{j_1,\ldots,j_n = 1 \\ k_1,\ldots,k_n = 1}}^{N} \|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}\|.$$

Here we divide $T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}$ into parts as follows

$$T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n} = \left(T_{j_1}^* T_{k_1}\right) \cdots \left(T_{j_n}^* T_{k_n}\right),$$

or

$$T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n} = \left(T_{j_1}^*\right) \left(T_{k_1} T_{j_2}^*\right) \cdots \left(T_{k_{n-1}} T_{j_n}^*\right) \left(T_{k_n}\right).$$

So we have

$$\|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}\| \leq \|T_{j_1}^* T_{k_1}\| \cdots \|T_{j_n}^* T_{k_n}\|$$

and

$$\|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}\| \leq \|T_{j_1}^*\| \|T_{k_1} T_{j_2}^*\| \cdots \|T_{k_{n-1}} T_{j_n}^*\| \|T_{k_n}\|.$$

Multiply the above inequalities we get

$$\|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}\| \leq \left(\|T_{j_1}^*\| \|T_{k_n}\|\right)^{\frac{1}{2}} \prod_{i=1}^{n-1} \|T_{k_i} T_{j_{i+1}}^*\|^{\frac{1}{2}} \prod_{i=1}^{n} \|T_{j_i}^* T_{k_i}\|^{\frac{1}{2}},$$

thus we have

$$\|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}\| \leq A \prod_{i=1}^{n-1} \gamma\left(k_i - j_{i+1}\right) \prod_{i=1}^{n} \gamma\left(j_i - k_i\right).$$

Then

$$\begin{aligned}
& \|\left(T^* T\right)^n\| \\
\leq & \sum_{j_1,\ldots,j_n=1,k_1,\ldots,k_n=1}^{N} A \prod_{i=1}^{n-1} \gamma\left(k_i - j_{i+1}\right) \prod_{i=1}^{n} \gamma\left(j_i - k_i\right) \\
= & \sum_{j_1,\ldots,j_n=1,k_1,\ldots,k_n=1}^{N} A \gamma\left(j_1 - k_1\right) \gamma\left(k_1 - j_2\right) \cdots \gamma\left(k_{n-1} - j_n\right) \gamma\left(j_n - k_n\right) \\
\leq & \sum_{j_2,\ldots,j_n=1,k_1,\ldots,k_n=1}^{N} A^2 \gamma\left(k_1 - j_2\right) \gamma\left(j_2 - k_2\right) \cdots \gamma\left(k_{n-1} - j_n\right) \gamma\left(j_n - k_n\right) \\
& \cdots \\
\leq & \sum_{k_n=1}^{N} A^{2N} 1 = A^{2N} N,
\end{aligned}$$

so we have

$$\|T\|^{2n} = \|T^* T\|^n = \|\left(T^* T\right)^n\| \leq A^{2N} N,$$

and

$$\|T\| \leq N^{\frac{1}{2n}} A,$$

let $n \to \infty$ we get $\|T\| = \|\sum_{j=1}^{N} T_j\| \leq A$. $\square$

Lemma 2.8.3 (A partition of unity over geometric scale)**.** There exists $\psi \in \mathscr{C}_0^\infty\left(\mathbb{R}^n\right)$ with the property that $supp\left(\psi\right) \subset \mathbb{R}^n \setminus \{0\}$ and that

$$\sum_{j \in \mathbb{Z}} \psi\left(2^{-j} x\right) \equiv 1 \ \ for\ any\ x \neq 0.$$

Moreover, ψ can be chosen to be a radical function, which satisfies $supp\left(\psi\right) \subset \left\{x \in \mathbb{R}^n : \frac{1}{2} \leq |x| \leq 2\right\}$ and $\psi \geq 0$.

Remark. Here we often denote $\psi_j(\cdot) = \psi(2^{-j}\cdot)$, so we have

$$\sum_{j\in\mathbb{Z}} \psi_j(x) \equiv 1 \ \ \forall x \in \mathbb{R}^n \setminus \{0\}.$$

Hence

$$supp\,(\psi_j) \subset \left\{x \in \mathbb{R}^n : 2^{j-1} \leq |x| \leq 2^{j+1}\right\} \quad \forall j \in \mathbb{Z}.$$

$\square$

Proof of lemma 2.8.3. Let $\chi \in \mathscr{C}_0^\infty(\mathbb{R}^n)$, s.t. $\chi(x) = 1$ for all $|x| \leq 1$ and $\chi(x) = 0$ for $|x| \geq 2$. Set $\psi(x) := \chi(x) - \chi(2x)$. Clearly for any positive N,

$$\sum_{j=-N}^{N} \psi\left(2^{-j}x\right) = \chi\left(2^{-N}x\right) - \chi\left(2^{N+1}x\right).$$

If $x \neq 0$ is given, then we take N so large that

$$\chi\left(2^{N+1}x\right) = 0 \ \ \textit{and} \ \ \chi\left(2^{-N}x\right) = 1,$$

Hence we get the proof. $\square$

Corollary 2.8.4. Let K be a Calderón-Zygmund kernel with the additional assumption that $|\nabla K(x)| \leq B\,|x|^{-n-1}$. Then

$$\|T\|_{2\to2} \leq CB$$

with $C = C(n)$.

Proof. Assume $B = 1$. By a dyadic decomposition, we set

$$K_j(x) = K(x)\,\psi\left(2^{-j}x\right).$$

It's easy to see that these kernels have the following properties:

1. $\int K_j(x)\,\mathrm{d}x = 0\,\forall j \in \mathbb{Z}$;
2. $\|\nabla K_j\|_{L^\infty(\mathbb{R}^n)} \leq C2^{-j}2^{-jn}$;
3. $\sup_{j\in\mathbb{Z}} \int |K_j(x)|\,\mathrm{d}x < \infty$;
4. $\sup_{j\in\mathbb{Z}} 2^{-j} \int |x|\,|K_j(x)|\,\mathrm{d}x < \infty$.

One can prove 1 in the sense of Riemann integral because K satisfies for any $0 < a < b < \infty$, $\int_{\{a<|x|<b\}} K(x)\,\mathrm{d}x = 0$ and ψ is radial. For 2, note that

$$\frac{\partial K_j}{\partial x_i}(x) = \psi\left(2^{-j}x\right)\frac{\partial K}{\partial x_i}(x) + K(x)\,\frac{\partial \psi\left(2^{-j}\cdot\right)}{\partial x_i}(x),$$

so we have

$$\left|\frac{\partial K_j}{\partial x_i}(x)\right| \le 2^{n+1}\|\psi\|_{L^\infty(\mathbb{R}^n)}2^{-j-jn} + 2^n 2^{-j}2^{-jn}\|\nabla\psi\|_{L^\infty(\mathbb{R}^n)},$$

thus

$$\|\nabla K_j\|_{L^\infty(\mathbb{R}^n)} \le C(n)\,2^{-j}2^{-jn}.$$

For 3 and 4, for any $j \in \mathbb{Z}$, it's easy to verify

$$\int |K_j(x)|\,\mathrm{d}x \le \int_{\{2^{j-1}\le|x|\le 2^{j+1}\}} |x|^{-n}\,\mathrm{d}x \le 2^{-(j-1)n}\,|B_1(0)|\,2^{(j+1)n} \le C(n),$$

and

$$2^{-j}\int |x|\,|K_j(x)|\,\mathrm{d}x \le 2\int |K_j(x)|\,\mathrm{d}x \le 2C(n).$$

Define

$$T_j f(x) = \int_{\mathbb{R}^n} K_j(x-y)\,f(y)\,\mathrm{d}y \quad \forall j \in \mathbb{Z},$$

It's bounded from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$. This integral is absolutely convergent for any $f \in L^1_{loc}(\mathbb{R}^n)$. Let $\tilde{K}_j(x) = \overline{K_j(-x)}$. Then

$$T_j^* T_k(x) = \int_{\mathbb{R}^n}\left(\tilde{K}_j * K_k\right)(y)\,f(x-y)\,\mathrm{d}y$$

and

$$T_j T_k^*(x) = \int_{\mathbb{R}^n}\left(K_j * \tilde{K}_k\right)(y)\,f(x-y)\,\mathrm{d}y.$$

Hence, by Young's inequaity,

$$\begin{cases} \|T_j^*T_k\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \le \|\tilde{K}_j * K_k\|_{L^1(\mathbb{R}^n)}, \\ \|T_jT_k^*\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \le \|K_j * \tilde{K}_k\|_{L^1(\mathbb{R}^n)}. \end{cases}$$

It suffices to consider the case $j \ge k$. Then, using $\int K_k(y)\,\mathrm{d}y = 0$, one obtains

$$\begin{aligned} \left|\left(\tilde{K}_j * K_k\right)(x)\right| &= \left|\int_{\mathbb{R}^n} \overline{K_j(y-x)}K_k(y)\,\mathrm{d}y\right| \\ &= \left|\int_{\mathbb{R}^n}\left(\overline{K_j(y-x)} - \overline{K_j(x)}\right)K_k(y)\,\mathrm{d}y\right| \\ &\le \int_{\mathbb{R}^n}\|\nabla K_j\|_{L^\infty(\mathbb{R}^n)}|y|\,|K_k(y)|\,\mathrm{d}y \\ &\le C2^{-j}2^{-jn}2^k, \end{aligned}$$

since

$$\operatorname{supp}\left(\tilde{K}_j * K_k\right) \subset \overline{\operatorname{supp}\left(\tilde{K}_j\right) + \operatorname{supp}\left(K_k\right)} \subset B_{c2^j}\left(0\right)$$

for some constant $c > 0$, thus

$$\|\tilde{K}_j * K_k\|_{L^1(\mathbb{R}^n)} \leq C2^{-(j-k)} = C2^{-|j-k|}.$$

By the Cotlar's lemma, denote $\gamma\left(l\right)^2 = C2^{-|l|}$, we get

$$\left\|\sum_{j=-N}^{N} T_j\right\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \leq C.$$

For any $f \in \mathscr{S}\left(\mathbb{R}^n\right)$, one has $\sum_{j=-N}^{N} T_j \to Tf$ pointwise and thus by Fatou's lemma,

$$\|Tf\|_{L^2(\mathbb{R}^n)} = \left\|\lim_{N\to\infty}\left|\sum_{j=-N}^{N} T_j f\right|\right\|_{L^2(\mathbb{R}^n)} \leq \liminf_{N\to\infty}\left\|\sum_{j=-N}^{N} T_j\right\|_{L^2(\mathbb{R}^n)} \leq C\|f\|_{L^2(\mathbb{R}^n)}.$$

$\square$

2.9 Multiplier theory

Let m be a bounded measurable function on $\mathbb{R}^n$. One can define a linear trandformation T_m, whose domain is $L^2\left(\mathbb{R}^n\right) \cap L^p\left(\mathbb{R}^n\right)$, by

$$\left(T_m f\right)^{\wedge}\left(\xi\right) = m\left(\xi\right)\widehat{f}\left(\xi\right), \quad \forall f \in L^2\left(\mathbb{R}^n\right) \cap L^p\left(\mathbb{R}^n\right).$$

We say that m is a multiplier(or symbol) is whenever $f \in L^2\left(\mathbb{R}^n\right) \cap L^p\left(\mathbb{R}^n\right)$, then $T_m f$ is also in $L^p\left(\mathbb{R}^n\right)$, and T_m is bounded, that is

$$\|T_m f\|_{L^p(\mathbb{R}^n)} \leq A\|f\|_{L^p(\mathbb{R}^n)} \tag{2.9.1}$$

for any $f \in L^2\left(\mathbb{R}^n\right) \cap L^p\left(\mathbb{R}^n\right)$.(with constant A independent of f) The smallest A for which inequality 2.9.1 holds will be called the norm of the multiplier,

$$\|T_m\| = \sup_{f\in L^p(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)} \frac{\|T_m f\|_{L^p(\mathbb{R}^n)}}{\|f\|_{L^p(\mathbb{R}^n)}}.$$

If inequality 2.9.1 is satisfied and $p < \infty$, then T_m has a unique bounded extension to $L^p\left(\mathbb{R}^n\right)$ to $L^p\left(\mathbb{R}^n\right)$, which again satisfies the inequality 2.9.1,

$$T_m f\left(x\right) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} m\left(\xi\right)\widehat{f}\left(\xi\right) \mathrm{d}\xi$$

for "good" f and m. Denote by $\mathcal{M}_p$ the class of multipliers with the indicated norm.

Example.

1. Hilbert transform $\widehat{Hf}(\xi) = -i\mathrm{sgn}(\xi)\widehat{f}(\xi)$;

2. Riesz transform $\widehat{R_j f}(\xi) = -i\frac{\xi_i}{|\xi|}\widehat{f}(\xi)$;

3. $\mathcal{M}_2$ is the class of all bounded measurable functions and the multiplier-norm is $L^\infty(\mathbb{R}^n)$-norm,

$$norm\ of\ m \in \mathcal{M}_2 = \|T_m\| = \|m\|_{L^\infty(\mathbb{R}^n)}.$$

Proposition 2.9.1. Suppose $\frac{1}{p}+\frac{1}{p'}=1$, $1 \le p \le \infty$. Then $\mathcal{M}_p = \mathcal{M}_{p'}$, with an identity of norms.

Proof. Let $\sigma(f)(x) = \overline{f(-x)}$. It's easy to verify $\sigma T_m \sigma = T_{\overline{m}}$. Indeed, note that $\forall f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$, we have $\sigma(f) \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ and

$$\begin{aligned}
((\sigma T_m \sigma) f)^\wedge(\xi) &= \int_{\mathbb{R}^n} \overline{(T_m\sigma(f))(-x)} e^{-2\pi i x\cdot\xi}\mathrm{d}x \\
&= \overline{\int_{\mathbb{R}^n} (T_m\sigma(f))(x)\, e^{-2\pi i x\cdot\xi}\mathrm{d}x} \\
&= \overline{\widehat{T_m\sigma(f)}}(\xi) = \overline{m(\xi)\,\widehat{\sigma f}(\xi)} = \overline{m(\xi)\,\overline{\overline{\widehat{f}(\xi)}}} = \overline{m(\xi)}\widehat{f}(\xi).
\end{aligned}$$

Therefore, if $m \in \mathcal{M}_p$, so does $\overline{m}$. Moreover, $\overline{m}$ has the same norm as m. By Plancherel's formula,

$$\int_{\mathbb{R}^n} T_m f\overline{g} = \int_{\mathbb{R}^n} m\widehat{f}\widehat{\overline{g}} = \int_{\mathbb{R}^n} \widehat{f}\overline{\overline{m}\widehat{g}} = \int_{\mathbb{R}^n} f\overline{(\overline{m}\widehat{g})^\vee} = \int_{\mathbb{R}^n} f\overline{T_{\overline{m}}g},$$

whenever $f, g \in L^2(\mathbb{R}^n)$. Assume $f \in L^2(\mathbb{R}^n) \cap L^{p'}(\mathbb{R}^n)$, $g \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ and $\|g\|_{L^p(\mathbb{R}^n)} \le 1$. Then

$$\left|\int_{\mathbb{R}^n} T_m f\overline{g}\right| \le \|f\|_{L^{p'}(\mathbb{R}^n)}\|T_{\overline{m}}g\|_{L^p(\mathbb{R}^n)} \le A\|f\|_{L^{p'}(\mathbb{R}^n)},$$

where A is the norm m in $\mathcal{M}_p$. Taking the supermum over all indecated g gives

$$\|T_m f\|_{L^{p'}(\mathbb{R}^n)} \le A\|f\|_{P'}, \quad f \in L^2(\mathbb{R}^n) \cap L^{p'}(\mathbb{R}^n).$$

Therefore, $m \in \mathcal{M}_{p\prime}$ and its $\mathcal{M}_{p\prime}$-norm $\le$ its $\mathcal{M}_p$-norm. Since the situation is symmetric in p and $p\prime$, the two norms are identical. $\square$

Theorem 2.9.2 (Mikhlin)**.** Let $m : \mathbb{R}^n \setminus \{0\} \to \mathbb{C}$ satisfy for any multi-index γ of length $|\gamma| \le n+2$,

$$|D^\gamma m(\xi)| \le B|\xi|^{-|\gamma|} \quad for\ all\ \xi \ne 0.$$

Then for any $1 < p < \infty$, there is a constant $C = C(n,p)$, s.t.

$$\left\| \left(m\widehat{f} \right)^{\vee} \right\|_{L^p(\mathbb{R}^n)} \leq CB \|f\|_{L^p(\mathbb{R}^n)} \quad for\ all\ f \in \mathscr{S}(\mathbb{R}^n).$$

Proof. Let ψ and ψ_j form a partition of unity over a geometric scale. For any $j \in \mathbb{Z}$, define

$$m_j(\xi) = \psi\left(2^{-j}\xi\right) m(\xi) \quad and \quad K_j = m_j^{\vee},$$

hence $m_j \in \mathscr{C}_c^{n+2}(\mathbb{R}^n)$. For any fixed $N \in \mathbb{N}$, denote $K(x) = \sum_{j=-N}^{N} K_j(x)$. Claim

$$|K(x)| \leq \frac{CB}{|x|^n} \quad and \quad |\nabla K(x)| \leq \frac{CB}{|x|^{n+1}} \tag{2.9.2}$$

for any $x \neq 0$ with $C = C(n)$.

Proof of the claim. Here we just verify the second inequality in 2.9.2. Note that

$$\|D^{\gamma}(\xi_i m_j)\|_{L^{\infty}(\mathbb{R}^n)} \leq CB 2^{-j(|\gamma|-1)}$$

for $|\gamma| \leq n+2$. Hence

$$\|D^{\gamma}(\xi_i m_j)\|_{L^1(\mathbb{R}^n)} \leq CB 2^{-j(|\gamma|-1)} 2^{jn},$$

and

$$\|x^{\gamma} D_i(\check{m}_j)\|_{L^{\infty}(\mathbb{R}^n)} \lesssim \|D^{\gamma}(\xi_i m_j)\|_{L^1(\mathbb{R}^n)} \leq CB 2^{j(n+1-|\gamma|)}.$$

Since

$$|x|^k = \left(\sum_{i=1}^{n} x_1^2 \right)^{\frac{k}{2}} \lesssim (|x_1| + \cdots + |x_n|)^k \leq C(k,n) \sum_{|\gamma|=k} |x^{\gamma}|,$$

we get

$$|D_i \check{m}_j(x)| \leq CB 2^{j(n+1-k)} |x|^{-k} = CB 2^{j(n+1)} \left(2^j |x|\right)^{-k}$$

for any $0 \leq k \leq n+2$, $j \in \mathbb{Z}$ and $x \neq 0$. Thus

$$\begin{aligned}
|\nabla K(x)| &\leq \sum_{j=-N}^{N} |\nabla(\check{m}_j(x))| \\
&\leq \sum_{2^j|x|\leq 1,\ |j|\leq N} |\nabla(\check{m}_j(x))| + \sum_{2^j|x|>1,\ |j|\leq N} |\nabla(\check{m}_j(x))| \\
&\leq \sum_{2^j|x|\leq 1,\ |j|\leq N} CB 2^{j(n+1)} + \sum_{2^j|x|>1,\ |j|\leq N} CB 2^{j(n+1)} \left(2^j |x|\right)^{-(n+2)} \\
&\leq CB |x|^{-(n+1)} + CB |x|^{-(n+2)} |x| \\
&\leq CB |x|^{-(n+1)}.
\end{aligned}$$

Set $Tf(x) = K * f$ for $f \in L^2(\mathbb{R}^n)$. Note that since $\|m\|_{L^\infty(\mathbb{R}^n)} \leq B$, we have

$$\|Tf\|_{L^2(\mathbb{R}^n)} = \|\widehat{K}\widehat{f}\|_{L^2(\mathbb{R}^n)} = \left\| \left(\sum_{j=-N}^{N} \psi_j m \right) \widehat{f} \right\|_{L^2(\mathbb{R}^n)} \leq B\|\widehat{f}\|_{L^2(\mathbb{R}^n)} = B\|f\|_{L^2(\mathbb{R}^n)},$$

i.e. T is bounded on $L^2(\mathbb{R}^n)$. Note also that $Tf(x) = \int_{\mathbb{R}^n} K(x-y) f(y) \, \mathrm{d}y$ and K satisfies the Hörmander condition. By Calderón-Zygmund theory,

$$\|Tf\|_{L^p(\mathbb{R}^n)} \leq CB\|f\|_{L^p(\mathbb{R}^n)}, \quad C = C(p,n), \quad 1 < p < \infty.$$

Then by a limiting argument, we get

$$\left\| \left(m\widehat{f} \right)^\vee \right\|_{L^p(\mathbb{R}^n)} \leq CB\|f\|_{L^p(\mathbb{R}^n)} \quad \forall f \in \mathscr{S}(\mathbb{R}^n).$$

□

Remark. Denote $K_N(x) = \sum_{j=-N}^{N} K_j(x)$ and $T_N f(x) = K_N * f(x)$ for all $f \in \mathscr{S}(\mathbb{R}^n)$. Observe that

$$\left\| T_N f - \left(m\widehat{f} \right)^\vee \right\|_{L^2(\mathbb{R}^n)} \to 0.$$

Indeed, by D.C.T. and Plancherel formula, we have

$$\begin{aligned}
0 &= \left\| \lim_{N\to\infty} \left\{ \left(\sum_{j=-N}^{N} \psi_j - 1 \right) m\widehat{f} \right\} \right\|_{L^2(\mathbb{R}^n)} \\
&= \lim_{N\to\infty} \left\| \left(\sum_{j=-N}^{N} \psi_j - 1 \right) m\widehat{f} \right\|_{L^2(\mathbb{R}^n)} \\
&= \lim_{N\to\infty} \left\| \sum_{j=-N}^{N} m_j \widehat{f} - \widehat{f} \right\|_{L^2(\mathbb{R}^n)} \\
&= \lim_{N\to\infty} \left\| T_N f - \left(m\widehat{f} \right)^\vee \right\|_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

By Riesz theorem, $\exists \{N_k\}_{k=1}^{\infty}$, s.t. $T_{N_K} f \to \left(m\widehat{f} \right)^\vee$ $\mathcal{L}^n - a.e.$. By Fatou's lemma,

$$\left\| \left(m\widehat{f} \right)^\vee \right\|_{L^p(\mathbb{R}^n)} = \left\| \lim_{k\to\infty} T_{N_K} f \right\|_{L^p(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|T_{N_K} f\|_{L^p(\mathbb{R}^n)} \leq CB \|f\|_{L^p(\mathbb{R}^n)}.$$

□

2.10 Littlewood-Paley theory

Lemma 2.10.1 (Khinchin's inequality)**.** Let $\{\omega_n\}_{n=1}^N$ be independent random variables, taking ± 1 with equal probability $\frac{1}{2}$. Then

$$C_1 \left(\sum_{n=1}^N |a_n|^2 \right)^{\frac{p}{2}} \leq \mathbb{E} \left(\left| \sum_{n=1}^N a_n \omega_n \right|^p \right) \leq C_2 \left(\sum_{n=1}^N |a_n|^2 \right)^{\frac{p}{2}}.$$

Here, $1 < p < \infty$, C_1, C_2 are constant depending on p only.

Remark. $\left(\mathbb{E} \left(\left| \sum_{n=1}^N a_n \omega_n \right|^p \right) \right)^{\frac{1}{p}} \asymp \left(\sum_{n=1}^N |a_n|^2 \right)^{\frac{1}{2}} = \left\| \{a_n\}_{n=1}^N \right\|_{l^2}.$ □

Proof of Khinchin's inequality 2.10.1. One can clearly assume the $\{a_n\}$ are real. Let $t > 0$, we have

$$\mathbb{E} \left(e^{t \sum_{n=1}^N a_n \omega_n} \right) = \prod_{n=1}^N \mathbb{E} \left(e^{t a_n \omega_n} \right) = \prod_{n=1}^N \frac{1}{2} \left(e^{-a_n t} + e^{a_n t} \right) \leq \prod_{n=1}^N e^{\frac{t^2}{2} a_n^2} = e^{\frac{t^2}{2} \sum_{n=1}^N a_n^2},$$

therefore

$$\begin{aligned} \mathbb{P} \left(\left\{ \sum_{n=1}^N a_n \omega_n \geq \lambda \right\} \right) &\leq e^{-t\lambda} \int_{\left\{ \sum_{n=1}^N a_n \omega_n \geq \lambda \right\}} e^{t \sum_{n=1}^N a_n \omega_n} \mathrm{d}\mathbb{P} \\ &\leq e^{-t\lambda} \mathbb{E} \left(e^{t \sum_{n=1}^N a_n \omega_n} \right) \\ &\leq e^{-t\lambda} e^{\frac{t^2}{2} \sum_{n=1}^N a_n^2} \\ &\leq e^{-\lambda^2 / \left(2 \sum_{n=1}^N a_n^2 \right)}. \end{aligned}$$

Similiarly,

$$\mathbb{P} \left(\left\{ \sum_{n=1}^N a_n \omega_n \leq -\lambda \right\} \right) \leq e^{-\lambda^2 / \left(2 \sum_{n=1}^N a_n^2 \right)},$$

hence

$$\mathbb{P} \left(\left\{ \left| \sum_{n=1}^N a_n \omega_n \right| \geq \lambda \right\} \right) \leq 2 e^{-\lambda^2 / \left(2 \sum_{n=1}^N a_n^2 \right)}.$$

Thus we get

$$\begin{aligned} \mathbb{E} \left(\left| \sum_{n=1}^N a_n \omega_n \right|^p \right) &= p \int_{(0,\infty)} \lambda^{p-1} \mathbb{P} \left(\left\{ \left| \sum_{n=1}^N a_n \omega_n \right| > \lambda \right\} \right) \mathrm{d}\lambda \\ &\leq 2p \int_0^\infty \lambda^{p-1} e^{-\lambda^2 / \left(2 \sum_{n=1}^N a_n^2 \right)} \mathrm{d}\lambda \\ &\leq C_p \left(\sum_{n=1}^N |a_n|^2 \right)^{\frac{p}{2}}. \end{aligned}$$

As to the lower bound, by Hölder's inequality, we get

$$\begin{aligned}
\sum_{n=1}^{N} |a_n|^2 &= \mathbb{E}\left|\sum_{n=1}^{N} a_n\omega_n\right|^2 \\
&\leq \left(\mathbb{E}\left|\sum_{n=1}^{N} a_n\omega_n\right|^p\right)^{\frac{1}{p}} \left(\mathbb{E}\left|\sum_{n=1}^{N} a_n\omega_n\right|^{p'}\right)^{\frac{1}{p'}} \\
&\leq C_p \left(\sum_{n=1}^{N} |a_n|^2\right)^{\frac{1}{2}} \left(\mathbb{E}\left|\sum_{n=1}^{N} a_n\omega_n\right|^p\right)^{\frac{1}{p}},
\end{aligned}$$

thus

$$\left(\sum_{n=1}^{N} |a_n|^2\right)^{\frac{1}{2}} \lesssim \left(\mathbb{E}\left|\sum_{n=1}^{N} a_n\omega_n\right|^p\right)^{\frac{1}{p}}.$$

□

With $\{\psi_j\}$ the partition of unity over a geometric scale, define

$$P_j f = \left(\psi_j \widehat{f}\right)^{\vee} = f * \check{\psi}_j$$

and Littlewood-Paley square function

$$Sf = \left(\sum_{j\in\mathbb{Z}} |P_j f|^2\right)^{\frac{1}{2}}.$$

Remark. Formally, we have

$$f = \left(\widehat{f}\right)^{\vee} = \left(\sum_{j\in\mathbb{Z}} \psi_j \widehat{f}\right)^{\vee} = \sum_{j\in\mathbb{Z}} \left(\psi_j \widehat{f}\right)^{\vee} = \sum_{j\in\mathbb{Z}} P_j f$$

with

$$\widehat{P_j f} = \psi_j \widehat{f}, \quad \mathrm{supp}\widehat{P_j f} \subset \textit{annulus} \sim 2^j.$$

□

By Plancherel formula, we have

$$C^{-1} \|f\|^2_{L^2(\mathbb{R}^n)} \leq \|Sf\|^2_{L^2(\mathbb{R}^n)} = \sum_{j\in\mathbb{Z}} \|P_j f\|^2_{L^2(\mathbb{R}^n)} \leq \|f\|^2_{L^2(\mathbb{R}^n)}$$

for any $f \in \mathscr{S}(\mathbb{R}^n)$. Indeed, note that $0 \leq \psi_j \leq 1$, we have

$$\|Sf\|^2_{L^2(\mathbb{R}^n)} = \sum_{j\in\mathbb{Z}} \int \left|\left(\psi_j \widehat{f}\right)^{\vee}\right|^2 = \int \left(\sum_{j\in\mathbb{Z}} \psi_j^2\right) \left|\widehat{f}\right|^2 \leq \|f\|^2_{L^2(\mathbb{R}^n)}.$$

But $\sum_{j\in\mathbb{Z}} \psi_j^2 \geq 1/3$, this is because

$$\begin{aligned}
1 &= \sum_{j\in\mathbb{Z}} \psi_j = \sum_{j\in\mathbb{Z}} \psi_j \mathbb{1}_{\{x\in\mathbb{R}^n : 2^{j-1}\leq |x| \leq 2^{j+1}\}} \\
&\leq \left(\sum_{j\in\mathbb{Z}} \psi_j^2\right)^{\frac{1}{2}} \left(\sum_{j\in\mathbb{Z}} \mathbb{1}_{\{x\in\mathbb{R}^n : 2^{j-1}\leq |x| \leq 2^{j+1}\}}\right)^{\frac{1}{2}} \leq 3^{-\frac{1}{2}} \left(\sum_{j\in\mathbb{Z}} \psi_j^2\right)^{\frac{1}{2}}.
\end{aligned}$$

Therefore,

$$\|Sf\|_{L^2(\mathbb{R}^n)} \asymp \|f\|_{L^2(\mathbb{R}^n)}.$$

Theorem 2.10.2 (Littlewood-Paley)**.** For any $1 < p < \infty$, there is a constant $C = C(p, n)$, s.t.

$$C^{-1} \|f\|_{L^p(\mathbb{R}^n)} \leq \|Sf\|_{L^p(\mathbb{R}^n)} \leq C \|f\|_{L^p(\mathbb{R}^n)}$$

for any $f \in \mathscr{S}(\mathbb{R}^n)$.

Remark. Actually, one can use a rough partition of unity to define a analogous Littlewood-Paley square function, here we denote $\triangle_j = \{x \in \mathbb{R}^n : 2^j \leq |x| < 2^{j+1}\}$, thus we have

$$f = \left(\widehat{f}\right)^\vee = \left(\sum_{j\in\mathbb{Z}} \mathbb{1}_{\triangle_j} \widehat{f}\right)^\vee = \sum_{j\in\mathbb{Z}} \left(\mathbb{1}_{\triangle_j} \widehat{f}\right)^\vee = \sum_{j\in\mathbb{Z}} f_j$$

where $f_j = \left(\mathbb{1}_{\triangle_j} \widehat{f}\right)^\vee$. By Plancherel formula, we also have

$$\begin{aligned}
\|f\|_{L^2(\mathbb{R}^n)}^2 &= \langle f, f\rangle = \langle \sum_{j\in\mathbb{Z}} f_j, \sum_{j\in\mathbb{Z}} f_j\rangle = \sum_{j,l\in\mathbb{Z}} \langle f_j, f_l\rangle = \sum_{j\in\mathbb{Z}} \|f_j\|_{L^2(\mathbb{R}^n)}^2 + \sum_{j,\neq l} \langle f_j, f_l\rangle \\
&= \sum_{j\in\mathbb{Z}} \|f_j\|_{L^2(\mathbb{R}^n)}^2 + \sum_{j\neq l} \langle \mathbb{1}_{\triangle_j} \widehat{f}, \mathbb{1}_{\triangle_l} \widehat{f}\rangle = \sum_{j\in\mathbb{Z}} \|f_j\|_{L^2(\mathbb{R}^n)}^2,
\end{aligned}$$

So

$$\begin{aligned}
\|f\|_{L^2(\mathbb{R}^n)} &= \left(\sum_{j\in\mathbb{Z}} \|f_j\|_{L^2(\mathbb{R}^n)}^2\right)^{\frac{1}{2}} = \left(\sum_{j\in\mathbb{Z}} \int |f_j|^2\right)^{\frac{1}{2}} \\
&= \left(\int \left(\sum_{j\in\mathbb{Z}} |f_j|^2\right)^{\frac{1}{2}\cdot 2}\right)^{\frac{1}{2}} = \left\| \|\{f_j\}_{j\in\mathbb{Z}}\|_{l^2} \right\|_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

Naturally we will think whether we have

$$\|f\|_{L^p(\mathbb{R}^n)} \asymp \left\| \left(\sum_{j\in\mathbb{Z}} |f_j|^2\right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)} \tag{2.10.1}$$

for $p \neq 2$, in fact, for $n \geq 2$ and $p \neq 2$, we don't have inequality 2.10.1. One can see this in Grafakos Vol.1. §6.1.4. However, by Littlewood-Paley theorem, we have $\|Sf\|_{L^p(\mathbb{R}^n)} \asymp \|f\|_{L^p(\mathbb{R}^n)}$. In a sense, Littlewood-Paley square function Sf is a good replacement of $\left(\sum_{j\in\mathbb{Z}} |f_j|^2\right)^{\frac{1}{2}}$, which reveals some orthogonality on $L^p(\mathbb{R}^n)$. $\square$

Proof of Littlewood-Paley theorem 2.10.2. Let $\{\omega_j\}$ be a sequence of independent random variables with

$$\mathbb{P}(\omega_j = 1) = \mathbb{P}(\omega_j = -1) = \frac{1}{2}$$

for all $j \in \mathbb{Z}$. Then

$$m(\xi) := \sum_{j=-N}^{N} \omega_j \psi_j(\xi)$$

satisfies the conditions of the Mikhlin multiplier theorem uniformly N and uniformly in $\{\omega_j\}$. Indeed, for any γ,

$$\begin{aligned} |D^\gamma m(\xi)| &\leq \sum_{j=-N}^{N} |\omega_j D^\gamma \psi_j(\xi)| \leq \sum_{j=-N}^{N} 2^{-j|\gamma|} \left|(D^\gamma \psi)\left(2^{-j}\xi\right)\right| \\ &\lesssim \sum_{|j|\leq N} |\xi|^{-|\gamma|} \left|(D^\gamma \psi)\left(2^{-j}\xi\right)\right| \lesssim \|D^\gamma \psi\|_{L^\infty(\mathbb{R}^n)} |\xi|^{-|\gamma|}. \end{aligned}$$

By Khinchin's inequality and Mikhlin multiplier theorem,

$$\begin{aligned} \|Sf\|_{L^p(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} |(Sf)(x)|^p \,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} \lim_{N\to\infty} \left(\sum_{|j|\leq N} |P_j f(x)|^2\right)^{\frac{p}{2}} \mathrm{d}x \\ &\leq \liminf_{N\to\infty} \int_{\mathbb{R}^n} \left(\sum_{|j|\leq N} |P_j f(x)|^2\right)^{\frac{p}{2}} \mathrm{d}x \\ &\leq C_p \liminf_{N\to\infty} \int_{\mathbb{R}^n} \mathbb{E}\left|\sum_{|j|\leq N} \omega_j P_j f(x)\right|^p \mathrm{d}x \\ &= C_p \liminf_{N\to\infty} \mathbb{E}\left(\int_{\mathbb{R}^n} \left|\sum_{|j|\leq N} \omega_j P_j f(x)\right|^p \mathrm{d}x\right) \\ &= C_p \liminf_{N\to\infty} \mathbb{E}\left(\int_{\mathbb{R}^n} \left|\left(\sum_{|j|\leq N} \omega_j \psi_j \widehat{f}\right)^\vee (x)\right|^p \mathrm{d}x\right) \\ &\leq C_{p,n,\psi} \liminf_{N\to\infty} \mathbb{E}\left(\|f\|_{L^p(\mathbb{R}^n)}^p\right) \\ &= C_{p,n,\psi} \|f\|_{L^p(\mathbb{R}^n)}^p. \end{aligned}$$

As to the lower bound, choose a function $\tilde{\psi}$ s.t. $\tilde{\psi} \equiv 1$ on suppψ, $\tilde{\psi} \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ and supp$\tilde{\psi} \subset \mathbb{R}^n \setminus \{0\}$. Hence $\tilde{\psi}\psi = \psi$. Set $\tilde{P}_j f = \left(\tilde{\psi}_j \widehat{f}\right)^\vee$ with $\tilde{\psi}_j(\cdot) = \tilde{\psi}(2^{-j}\cdot)$. Then $\tilde{P}_j P_j = P_j$, indeed, note that

$$\tilde{P}_j P_j f = \left(\tilde{\psi}_j \widehat{P_j f}\right)^\vee = \left(\tilde{\psi}_j \widehat{\left(\psi_j \widehat{f}\right)^\vee}\right)^\vee = \left(\tilde{\psi}_j \psi_j \widehat{f}\right)^\vee = \left(\psi_j \widehat{f}\right)^\vee = P_j f.$$

For $f, g \in \mathscr{S}(\mathbb{R}^n)$ and any $1 < p < \infty$,

$$\begin{aligned}
|\langle f, g\rangle| &= \left|\langle \sum_{j\in\mathbb{Z}} P_j f, g\rangle\right| = \left|\sum_{j\in\mathbb{Z}} \langle P_j f, g\rangle\right| = \left|\sum_{j\in\mathbb{Z}} \langle \tilde{P}_j P_j f, g\rangle\right| = \left|\sum_{j\in\mathbb{Z}} \langle P_j f, \tilde{P}_j g\rangle\right| \\
&= \left|\int \sum_{j\in\mathbb{Z}} P_j f \overline{\tilde{P}_j g}\right| \le \int \left(\sum_{j\in\mathbb{Z}} |P_j f|^2\right)^{\frac{1}{2}} \left(\sum_{j\in\mathbb{Z}} \left|\tilde{P}_j g\right|^2\right)^{\frac{1}{2}} \\
&\le \|Sf\|_{L^p(\mathbb{R}^n)} \left\|\tilde{S}g\right\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|Sf\|_{L^p(\mathbb{R}^n)} \|g\|_{L^{p'}(\mathbb{R}^n)},
\end{aligned}$$

here we denote $Sf := \left(\sum_{j\in\mathbb{Z}} |P_j f|^2\right)^{1/2}$ and $\tilde{S}g := \left(\sum_{j\in\mathbb{Z}} \left|\tilde{P}_j g\right|^2\right)^{1/2}$ and we can also prove for $\tilde{\psi}$ and $\tilde{\psi}_j(\cdot) = \psi(2^{-j}\cdot)$, multiplier $\tilde{m}(\xi) := \sum_{j=-N}^{N} \omega_j \tilde{\psi}_j(\xi)$ also satisfies the condition of Mikhlin multiplier theorem uniformly N and uniformly in $\{\omega_j\}$. Hence, we have

$$\frac{|\langle f, g\rangle|}{\|g\|_{L^{p'}(\mathbb{R}^n)}} \lesssim \|Sf\|_{L^p(\mathbb{R}^n)},$$

taking supremum we get the proof. □

Corollary 2.10.3. Let $1 < p < \infty$. Then for any $f \in L^p(\mathbb{R}^n)$, one has $Sf \in L^p(\mathbb{R}^n)$ and

$$C_{p,n}^{-1} \|f\|_{L^p(\mathbb{R}^n)} \le \|Sf\|_{L^p(\mathbb{R}^n)} \le C_{p,n} \|f\|_{L^p(\mathbb{R}^n)}.$$

Remark. Since $f \in L^p(\mathbb{R}^n)$, $f \in \mathscr{S}'(\mathbb{R}^n)$, then Sf is defined pointwisely since $P_j f = f * \check{\psi}_j \in \mathscr{C}^\infty(\mathbb{R}^n)$. □

Proof. Let $f_k \in \mathscr{S}(\mathbb{R}^n)$, s.t. $f_k \to f$ in $L^p(\mathbb{R}^n)$. We claim that

$$\lim_{k\to\infty} \|Sf_k - Sf\|_{L^p(\mathbb{R}^n)} = 0.$$

Since $C_{p,n}^{-1} \|f_k\|_{L^p(\mathbb{R}^n)} \le \|Sf_k\|_{L^p(\mathbb{R}^n)} \le C_{p,n} \|f_k\|_{L^p(\mathbb{R}^n)}$, let $k \to \infty$, we have

$$C_{p,n}^{-1} \|f\|_{L^p(\mathbb{R}^n)} \le \|Sf\|_{L^p(\mathbb{R}^n)} \le C_{p,n} \|f\|_{L^p(\mathbb{R}^n)}.$$

To prove the claim, note

$$\begin{aligned}
|Sf_k(x) - Sf(x)| &= \left| \left\| \{P_j f_k(x)\}_{j\in\mathbb{Z}} \right\|_{l^2} - \left\| \{P_j f(x)\}_{j\in\mathbb{Z}} \right\|_{l^2} \right| \\
&\leq \left\| \{P_j (f_k - f)(x)\}_{j\in\mathbb{Z}} \right\|_{l^2} \\
&= S(f_k - f)(x) \leq \liminf_{m\to\infty} S(f_k - f_m)(x).
\end{aligned}$$

Indeed, we have

$$\begin{aligned}
S^2(f - f_k)(x) &= \lim_{N\to\infty} \sum_{j=-N}^{N} |P_j (f - f_k)(x)|^2 \\
&= \lim_{N\to\infty} \sum_{j=-N}^{N} \left|\check{\psi}_j * (f - f_k)(x)\right|^2 \\
&= \lim_{N\to\infty} \lim_{m\to\infty} \sum_{j=-N}^{N} \left|\check{\psi}_j * (f_m - f_k)(x)\right|^2 \\
&\leq \liminf_{m\to\infty} \lim_{N\to\infty} \sum_{j=-N}^{N} \left|\check{\psi}_j * (f_m - f_k)(x)\right|^2 \\
&= \liminf_{m\to\infty} S^2(f_m - f_k)(x).
\end{aligned}$$

Therefore, by Fatou's lemma and Littlewood-Paley theorem, we get

$$\|Sf_k - Sf\|_{L^p(\mathbb{R}^n)} \leq \liminf_{m\to\infty} \|Sf_k - Sf_m\|_{L^p(\mathbb{R}^n)} \leq \liminf_{m\to\infty} C_{p,n} \|f_k - f_m\|_{L^p(\mathbb{R}^n)},$$

let $k \to \infty$, then we get $\limsup_{k\to\infty} \|Sf_k - Sf\|_{L^p(\mathbb{R}^n)} \leq 0$. $\square$

What about rough cut-off function $\mathbb{1}_{\{1\leq|x|<2\}}(2^{-j}\cdot)$ instead of ψ_j? Do we have analogous Littlewood-Paley theorem? Let $\triangle_j = \{x \in \mathbb{R}^n : 2^j \leq |x| < 2^{j+1}\}$, define

$$S_{new} f := \left(\sum_{j\in\mathbb{Z}} \left| \left(\mathbb{1}_{\triangle_j} \widehat{f} \right)^{\vee} \right|^2 \right)^{\frac{1}{2}} \quad \textit{and} \quad S_j^{new} f := \left(\mathbb{1}_{\triangle_j} \widehat{f} \right)^{\vee}.$$

By Plancherel formula, for $p = 2$, $\|S_{new} f\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}$. In fact, we have the following theorem.

Theorem 2.10.4 ($\mathbb{R}^1$ case)**.** Let $f \in \mathscr{S}(\mathbb{R})$. $1 < p < \infty$. Then $\exists c_p,\ C_p > 0$, s.t.

$$c_p \|f\|_{L^p(\mathbb{R})} \leq \|S_{new} f\|_{L^p(\mathbb{R})} \leq C_p \|f\|_{L^p(\mathbb{R})}.$$

We will prove it latter. Now, let's introduce an application of Khinchin's inequality.

Theorem 2.10.5. Let T be a linear operator, s.t.

$$\|Tf\|_{L^p(\mathbb{R}^n)} \leq C_p \|f\|_{L^p(\mathbb{R}^n)}$$

for $1 < p < \infty$ and for all $f \in L^p(\mathbb{R}^n)$. Then,

$$\left\| \left(\sum_{j\in\mathbb{Z}} |Tf_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)} \leq C_p' \left\| \left(\sum_{j\in\mathbb{Z}} |f_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)}$$

for some $C_p' > 0$ depending on C_p and p only.

Proof. Let $\{\omega_j\}_{j=1}^N$ be i.r.v. taking ± 1 with equal probability. Then by Fubini-Tonelli theorem and Khinchin's inequality, we have

$$\begin{aligned}
\left\| \left(\sum_{|j|\leq N} |Tf_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^n)}^p &\leq C_1^{-1} \int_{\mathbb{R}^n} \mathbb{E}\left(\left| \sum_{|j|\leq N} \omega_j Tf_j(x) \right|^p \right) \mathrm{d}x \\
&= C_1^{-1} \mathbb{E}\left(\int_{\mathbb{R}^n} \left| \sum_{|j|\leq N} \omega_j Tf_j(x) \right|^p \mathrm{d}x \right) \\
&= C_1^{-1} \int \left\| \sum_{|j|\leq N} \omega_j Tf_j \right\|_{L^p(\mathbb{R}^n)}^p \mathrm{d}\mathbb{P} \\
&= C_1^{-1} \int \left\| T\left(\sum_{|j|\leq N} \omega_j f_j \right) \right\|_{L^p(\mathbb{R}^n)}^p \mathrm{d}\mathbb{P} \\
&\leq C_1^{-1} C_p^p \int \left\| \sum_{|j|\leq N} \omega_j f_j \right\|_{L^p(\mathbb{R}^n)}^p \mathrm{d}\mathbb{P} \\
&= C_1^{-1} C_p^p \int_{\mathbb{R}^n} \mathbb{E}\left(\left| \sum_{|j|\leq N} \omega_j f_j(x) \right|^p \right) \mathrm{d}x \\
&\leq C_1^{-1} C_2 C_p^p \int_{\mathbb{R}^n} \left(\sum_{|j|\leq N} |f_j|^2 \right)^{\frac{p}{2}} \mathrm{d}x \\
&\leq C_1^{-1} C_2 C_p^p \int_{\mathbb{R}^n} \left(\sum_{j\in\mathbb{Z}} |f_j|^2 \right)^{\frac{p}{2}} \mathrm{d}x.
\end{aligned}$$

By Fatou's lemma, we complete the proof. $\square$

Lemma 2.10.6. Set $S_j^{new} f = \left(\mathbb{1}_{\{x\in\mathbb{R}: 2^j \leq |x| < 2^{j+1}\}} \widehat{f} \right)^{\vee}$, we have

$$\left\| \left(\sum_{j\in\mathbb{Z}} \left| S_j^{new} f_j \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})} \leq C_p \left\| \left(\sum_{j\in\mathbb{Z}} |f_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})}$$

for $1 < p < \infty$ and $f_j \in \mathscr{S}(\mathbb{R})$.

Proof. Note that if $\widehat{S_{[a,b]}f}(\xi) := \mathbb{1}_{[a,b]}(\xi)\,\widehat{f}(\xi)$ for $f \in \mathscr{S}(\mathbb{R})$, then

$$S_{[a,b]}f = \frac{i}{2}\left(M_a H M_{-a} f - M_b H M_{-b} f\right) \quad \mathcal{L}^1 - a.e.,$$

here H is the Hilbert transform, $M_a f(\cdot) = e^{2\pi i a \cdot} f(\cdot)$. Indeed, we note that

$$\left(\widehat{M_a H M_{-a} f}\right)(\xi) = -i\mathrm{sgn}(\xi - a)\,\widehat{M_{-a}f}(\xi - a) = -i\mathrm{sgn}(\xi - a)\,\widehat{f}(\xi),$$

hence we get

$$\begin{aligned}\left[\frac{i}{2}\left(M_a H M_{-a} f - M_b H M_{-b} f\right)\right]^{\wedge} &= \frac{i}{2}\left(-i\mathrm{sgn}(\xi - a) + i\mathrm{sgn}(\xi - b)\right)\widehat{f}(\xi) \\ &= \mathbb{1}_{[a,b]}(\xi)\,\widehat{f}(\xi) \quad \mathcal{L}^1 - a.e..\end{aligned}$$

Thus by the Calderón–Zygmund theory we know that the Hilbert transform H is bounded from $L^p(\mathbb{R})$ to $L^p(\mathbb{R})$ for $1 < p < \infty$ and then this lemma follows from Theorem 2.10.5. □

Now we complete the proof of theorem 2.10.4.

Proof of theorem 2.10.4. Let $\psi \in \mathscr{S}(\mathbb{R})$, $\psi \geq 0$ and $\mathrm{supp}\psi \subset \left\{\frac{1}{2} \leq |x| \leq 4\right\}$, $\psi \equiv 1$ on $1 \leq |x| \leq 2$. By Littlewood-Paley theorem and lemma 2.10.6, we have

$$\begin{aligned}\|S_{new}f\|_{L^p(\mathbb{R})} &= \left\| \left(\sum_{j\in\mathbb{Z}} \left| \left(\mathbb{1}_{\triangle_j}\widehat{f}\right)^{\vee} \right|^2\right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})} \\ &= \left\| \left(\sum_{j\in\mathbb{Z}} \left| \left(\mathbb{1}_{\triangle_j}\psi_j\widehat{f}\right)^{\vee} \right|^2\right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})} \\ &= \left\| \left(\sum_{j\in\mathbb{Z}} \left| S_j^{new} S_j f \right|^2\right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})} \\ &\leq C_p \left\| \left(\sum_{j\in\mathbb{Z}} \left| S_j f_j \right|^2\right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})} \\ &\leq C\,\|f\|_{L^p(\mathbb{R})}.\end{aligned}$$

From the polorization identity 2.10 and $\|S_{new}f\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}$, we get

$$\int_{\mathbb{R}} \sum_{j\in\mathbb{Z}} S_j^{new} f\,\overline{S_j^{new} g} = \int_{\mathbb{R}} f\overline{g}.$$

Hence

$$\begin{aligned}
\|f\|_{L^p(\mathbb{R})} &= \sup_{\|g\|_{L^{p'}(\mathbb{R})}\le 1} \left| \int_{\mathbb{R}} f\overline{g} \right| = \sup_{\|g\|_{L^{p'}(\mathbb{R})}\le 1} \left| \int_{\mathbb{R}} \sum_{j\in\mathbb{Z}} S_j^{new} f \overline{S_j^{new} g} \right| \\
&\le \sup_{\|g\|_{L^{p'}(\mathbb{R})}\le 1} \left\| \left(\sum_{j\in\mathbb{Z}} \left| S_j^{new} f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R})} \left\| \left(\sum_{j\in\mathbb{Z}} \left| S_j^{new} g \right|^2 \right)^{\frac{1}{2}} \right\|_{L^{p'}(\mathbb{R})} \\
&\le C \left\| S_{new} f \right\|_{L^p(\mathbb{R})}.
\end{aligned}$$

□

Remark. Let $\mathscr{H}$ be a Hilbert space. We have the following *polorization identity*:

$$\langle f, g\rangle = \frac{1}{4} \left(\|f+g\|^2 - \|f-g\|^2 + i\,\|f+ig\|^2 - i\,\|f-ig\|^2 \right).$$

□

Remark. For $\mathbb{R}^n (n \ge 2)$ case. When $p = 2$, we have $\|S_{new} f\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}$. For $p \ne 2$, L^p-equivalence does not hold. The reason is that when $n \ge 2$ and $p \ne 2$, indicator function $\mathbb{1}_{B_1(0)}$ is not in $\mathcal{M}_p$, i.e. $\mathbb{1}_{B_1(0)}$ is not a L^p-multiplier on $\mathbb{R}^n$. In fact, Chareles Fefferman proved the following argument:

$$For\ n \ge 2,\ \left\| \left(\mathbb{1}_{B_1(0)} \widehat{f} \right)^{\vee} \right\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^p(\mathbb{R}^n)}\ iff\ p = 2.$$

One can see some details in Grafakos Vol.1. §5.1.3. □

3 Weighted inequalities

In this section, we shall briefly introduce the content of weighted inequalities, or namely, A_p weights.

It is known that the (uncentered) Hardy-Littlewood maximal operator is weak type (1,1) and strong type (p,p) . More specifically,

$$|\{Mf > \alpha\}| \lesssim \alpha^{-1} \int |f|, \quad f \in L^1, \alpha > 0; \tag{3.0.1}$$

$$\int |Mf|^p \precsim \int |f|^p, \quad f \in L^p. \tag{3.0.2}$$

Note that in (3.0.1)(3.0.2), the measure under discussion is Lebesgue measure only. For practical use, we hope to obtain something like (3.0.1)(3.0.2) but with Lebesgue measure replaced by some weighted measure (for example, the elliptic measure in PDE theory).

Definition 3.0.1. We say that $w : \mathbb{R}^n \to [0, \infty]$ is a weight, if and only if $w \in L^1_{loc}$ and $0 < w < \infty$ a.e. (which ensures that we can talk about w^{-1} , however, w^{-1} is not necessarily locally integrable.). w can automatically be viewed as a positive measure: $w(E) = \int_E w$.

Remark. In this section, we shall always use w to refer to a weight in $\mathbb{R}^n$, M the HardyLittlewood maximal operator, and the component p in the following context is understood to take values in $(1, \infty)$ unless otherwise claimed. □

We shall mainly study the following problems below: under what assumptions on w can we establish

$$w(\{Mf > \alpha\}) \precsim \alpha^{-1} \int |f| w, \quad f \in L^1(w), \alpha > 0; \tag{3.0.3}$$

$$\int |Mf|^p w \precsim \int |f|^p w, \quad f \in L^p(w). \tag{3.0.4}$$

Or equivalently, for what kinds of w can we have

$$\|M\|_{L^1(w) \to L^{1,\infty}(w)}, \|M\|_{L^p(w) \to L^p(w)} < \infty$$

(Note that the Lebesgue measure in hidden in M)

Usually, it is hard to find the sufficient conditions, so we attempted to find the necessary condition first which means that we will take some special form of f, α in (3.0.3)(3.0.4) to find the property that w should satisfy.

The hard part in (3.0.3)(3.0.4) is the term concerning M . Note that M relies on balls in $\mathbb{R}^n$. So, we first localize everything on balls.

(In this section, we shall denote B a general open ball of $\mathbb{R}^n$)

Let us first consider (3.0.4) . Given B , replace f with $f\chi_B$, we have

$$\int_B |f|^p w \gtrsim \int_{\mathbb{R}^n} |M(f\chi_B)|^p w \geq \int_B |M(f\chi_B)|^p w.$$

Note that we have uniform lower estimate for M on B :

$$Mf \geq \frac{1}{|B|} \int_B |f| \text{ on } B$$

Hence

$$\int_B |f|^p w \gtrsim w(B)\left(\frac{1}{|B|}\int_B |f|\right)^p, \quad \text{i.e. } \frac{1}{|B|}\int_B |f|^p w \gtrsim \frac{w(B)}{|B|}\left(\frac{1}{|B|}\int_B |f|\right)^p. \tag{3.0.5}$$

A nature idea is to balance the integrand, thus, we take $f = w^{-\frac{1}{p-1}}$ to obtain:

$$\left(\frac{1}{|B|}\int_B w\right)\left(\frac{1}{|B|}\int_B w^{-\frac{1}{p-1}}\right)^{p-1} \lesssim 1, \quad \forall B \tag{3.0.6}$$

we thus obtain the necessary condition for (3.0.4).

NOTE: we must be careful when dividing something in analysis to make sure that the divisor is not 0 or ∞ . Here, if we take $f = w^{-\frac{1}{p-1}}$ directly, the divisor $\frac{1}{|B|}\int_B w^{-\frac{1}{p-1}} \neq 0$ is ensured by the definition of weight. However, we have no reason to ensure it is not ∞ (take $p = 2$ to see more explicitly). However, we can make a simply correction to this obstacle: take $f = (w+\varepsilon)^{-\frac{1}{p-1}}$ and let $\varepsilon \to 0$ via the monotonic convergence theorem.

Definition 3.0.2. Denote by

$$[w]_{A_p} := \sup_B \left(\frac{1}{|B|}\int_B w\right)\left(\frac{1}{|B|}\int_B w^{-\frac{1}{p-1}}\right)^{p-1}.$$

Say w is a A_p weight or simply $w \in A_p$ if $[w]_{A_p} < \infty$.

Remark. We have known that

$$\|M\|_{L^p(w)\to L^p(w)} < \infty \Rightarrow w \in A_p.$$

And note that $[w]_{A_p} < \infty$ implies $w \in L^1_{loc}$. □

As to the case when $p = 1$, we can guess from the L^p norm limit the condition under which (3.0.3) holds. Let $p \to 1$ in (3.0.6), we have:

$$\left(\frac{1}{|B|}\int_B w\right)\left\|w^{-1}\right\|_{L^\infty(B)} \lesssim 1, \quad \forall B \tag{3.0.7}$$

Of course, we should check that whether (3.0.7) is really a necessary condition of (3.0.3). Given B as usual, let $\alpha < \frac{1}{|B|}\int_B |f|$ in (3.0.3) (hence $\{Mf > \alpha\}$ contains B), we have:

$$\int |f| w \gtrsim \frac{w(B)}{|B|}\int_B |f|$$

Furthermore, replace f with $f\chi_B$ to get:

$$\frac{1}{|B|}\int_B |f| \lesssim \frac{1}{w(B)}\int_B |f| w. \tag{3.0.8}$$

(3.0.8) says the integral means if f over B respect to Lebesgue measure is controlled by that of f over B respect to w)

To give a more explicit insight, take $f = \chi_S$ where S is an arbitrary subset of B of positive measure. Hence,

$$\frac{1}{|B|}\int_B w \preccurlyeq \frac{1}{|S|}\int_S w.$$

Since S is arbitrary, once if $x \in B$ is the Lebesgue point of w , we have

$$\frac{1}{|B|}\int_B w \precsim w(x).$$

We thus recover (3.0.7) from (3.0.3) since almost every point of $\mathbb{R}^n$ is the Lebesgue point of w .

Definition 3.0.3. Denote by

$$[w]_{A_1} := \sup_B \left(\frac{1}{|B|}\int_B w\right) \left\|w^{-1}\right\|_{L^\infty(B)}.$$

Say w is a A_1 weight or simply $w \in A_1$ if $[w]_{A_1} < \infty$.

Remark. We have known that

$$\|M\|_{L^p(w)\to L^p(w)} < \infty \Rightarrow w \in A_p.$$

And note that $[w]_{A_p} < \infty$ implies $w \in L^1_{loc}$.

□

Next, we shall show the surprising fact that the necessary condition we obtained before is also sufficient. That is, we shall prove:

Theorem 3.0.4.

$$w \in A_1 \Rightarrow \|M\|_{L^1(w)\to L^{1,\infty}(w)} < \infty$$

$$w \in A_p \Rightarrow \|M\|_{L^p(w)\to L^p(w)} < \infty$$

Before giving the proof of Theorem 4, we make some preparations first.

1. A_p weight is doubling measure when $p \geq 1$. Take $f = \chi_{\frac{B}{2}}$ in eqrefeq3.0.5(3.0.8) is OK (Note that (3.0.5),(3.0.6) are equivalent and (3.0.7),(3.0.8) are equivalent).

2. the maximal operator induced by doubling measure is L^p bounded. Let $\mu \geq 0$ be a Radon measure in $\mathbb{R}^n$, define the maximal operator induced by μ as

$$(M^\mu f)(x) := \sup_{B \ni x, 0<\mu(B)<\infty} \frac{1}{\mu(B)} \int_B |f| d\mu.$$

If μ is doubling, then

$$\|M^\mu\|_{L^1(\mu)\to L^{1,\infty}(\mu)} < \infty, \quad \|M^\mu\|_{L^p(\mu)\to L^p(\mu)} < \infty$$

(The proof of this result is almost the same with that of M)

3. the duality in A_p weight when $p > 1$. If $w \in A_p$, then $\sigma := w^{-\frac{1}{p-1}} \in A_{p'}$.

Now, we can give the proof of Theorem 3.0.4.

Proof. The idea of the prove is to obtain pointwise estimate for M concerning M^w, M^σ where $\sigma = w^{-\frac{1}{p-1}}$, and the (weighted) L^p boundness of M follows from that of M^w and M^σ .

(i) if $w \in A_1$, then we have from (3.0.8) that $Mf \precsim M^w f$, thus

$$\|Mf\|_{L^{1,\infty}(w)} \precsim \|M^w f\|_{L^{1,\infty}(w)} \precsim \|f\|_{L^1(w)}$$

(ii) reformulate $\frac{1}{|B|}\int_B |f|$ as follows:

$$\frac{1}{|B|}\int_B |f| = \frac{w(B)^{\frac{1}{p-1}}\sigma(B)}{|B|^{\frac{p}{p-1}}} \left(\frac{|B|}{w(B)}\left(\frac{1}{\sigma(B)}\int_B |f|\right)^{p-1}\right)^{\frac{1}{p-1}}$$

Note that the coefficient outside in the righthand is $[w]_{A_p}^{\frac{1}{p-1}}$, the treatment for the integral is:

$$\begin{aligned}
&\frac{1}{\sigma(B)}\int_B |f| \leq M^\sigma\left(|f|\sigma^{-1}\right) \text{ on } B \\
\Rightarrow &\left(\frac{1}{\sigma(B)}\int_B |f|\right)^{p-1} \leq \left(M^\sigma\left(|f|\sigma^{-1}\right)\right)^{p-1} \text{ on } B \\
\Rightarrow &\left(\frac{1}{\sigma(B)}\int_B |f|\right)^{p-1} \leq \frac{1}{|B|}\int_B \left(M^\sigma\left(|f|\sigma^{-1}\right)\right)^{p-1} \\
\Rightarrow &\frac{|B|}{w(B)}\left(\frac{1}{\sigma(B)}\int_B |f|\right)^{p-1} \leq \frac{1}{w(B)}\int_B \left(M^\sigma\left(|f|\sigma^{-1}\right)\right)^{p-1} \\
\Rightarrow &\frac{|B|}{w(B)}\left(\frac{1}{\sigma(B)}\int_B |f|\right)^{p-1} \leq M^w\left(\left(M^\sigma\left(|f|\sigma^{-1}\right)\right)^{p-1} w^{-1}\right)
\end{aligned}$$

Hence

$$Mf \preccurlyeq \left(M^w\left(\left(M^\sigma\left(|f|\sigma^{-1}\right)\right)^{p-1} w^{-1}\right)\right)^{\frac{1}{p-1}}$$

Hence

$$
\begin{aligned}
\|Mf\|_{L^p(w)} &\lesssim \left\| \left(M^w \left(\left(M^\sigma \left(|f|\sigma^{-1} \right) \right)^{p-1} w^{-1} \right) \right)^{\frac{1}{p-1}} \right\|_L (w) \\
&= \left\| M^w \left(\left(M^\sigma \left(|f|\sigma^{-1} \right) \right)^{p-1} w^{-1} \right) \right\|_{L^{p'}(w)}^{\frac{1}{p-1}} \\
&\lesssim \left\| \left(M^\sigma \left(|f|\sigma^{-1} \right) \right)^{p-1} w^{-1} \right\|_{L^{p'}(w)}^{\frac{1}{p-1}} \\
&= \left\| M^\sigma \left(|f|\sigma^{-1} \right) \right\|_{L^p} (\sigma) \\
&\lesssim \left\| |f|\sigma^{-1} \right\|_{L^p(\sigma)} = \|f\|_{L^p(w)}
\end{aligned}
$$

□

In the end, we list some further properties of A_p weights without offering any proof.

1. pair of A_p weights when $1 < p < \infty$.

It is nature to ask whether we can establish inequality like (3.0.4) but setting the weight differently on either side, that is

$$\int |Mf|^p u \preccurlyeq \int |f|^p w, \quad f \in L^p(w).$$

The conclusion is similar to what we obtained before but somehow weaker: if

$$[u, w]_{(A_p, A_p)} := \sup_B \left(\frac{1}{|B|} \int_B u \right) \left(\frac{1}{|B|} \int_B w^{-\frac{1}{p-1}} \right)^{p-1} < \infty.$$

Then

$$\|M\|_{L^p(w) \to L^{p,\infty}(u)} < \infty.$$

2. geometry properties when $1 \leq p < \infty$.

Take $f = \chi_S$, $S \subset B$ in (3.0.5), we have

$$\left(\frac{|S|}{|B|} \right)^p \precsim \frac{w(S)}{w(B)}, 1 \leq p < \infty.$$

To excavate the geometry meaning of this inequality, replace S with $B \backslash S$:

$$\left(1 - \frac{|S|}{|B|} \right)^p \lesssim 1 - \frac{w(S)}{w(B)}$$

Thus, we have:

$$\frac{|S|}{|B|} \leq \alpha \Rightarrow \frac{w(S)}{w(B)} \leq \beta \text{ for some } \beta.$$

That is, if the proportion of S in B is small in the sight of Lebesgue measure, then it is also small in the sight of w .

The reverse is also true: if the proportion of S in B is small in the sight of w, then it is also small in the sight of Lebesgue measure. More specifically, we can prove that: if $w \in A_p$, then for any B and $S \subset B$, we have

$$\frac{w(S)}{w(B)} \precsim \left(\frac{|S|}{|B|}\right)^{\delta} \text{ for some } \delta = \delta(n, p, w)$$

3. reverse Holder's inequality when $1 \leq p < \infty$.

If $w \in A_p$, then for some $\gamma = \gamma(n, p, w) > 0$, we have

$$\left(\frac{1}{|B|}\int_B w^{1+\gamma}\right)^{\frac{1}{1+\gamma}} \lesssim \frac{1}{|B|}\int_B w$$

The reverse holder inequality may be the core property of A_p weights which gives abundant surprising results of A_p weights. For example, it is obvious to see that A_p class is monotonic, we can give a deeper connection between A_p class via reverse holder inequality when $p > 1$:

$$A_p = \bigcup_{1<q<p} A_q, \quad p > 1.$$

That is, if $w \in A_p$, then $w \in A_{p-\varepsilon}$ for some $\varepsilon > 0$.

4 Oscillatory integrals

4.1 Some Fourier transforms calculation

This part is based on T. Wolff's note chapter 4.

Definition 4.1.1. A tempered function is a function $f \in L^1_{loc}(\mathbb{R}^n)$ s.t.

$$\int_{\mathbb{R}^n} (1 + |x|)^{-N} |f(x)| \, \mathrm{d}x < \infty$$

for some constant N.

In Stein's book ¡¡Fourier Analysis¿¿, he called a tempered function as a tempered L^1 function, these definitions are equivalent. Belows are four basic propositions in this section and we don't prove them here.

Proposition 4.1.2. If f is tempered and $\varphi \in \mathscr{S}$, then

$$\int |\varphi f| < \infty$$

and the following mapping

$$\varphi \mapsto \int \varphi f$$

is continuous on $\mathscr{S}$. Also, $\varphi * f$ is well-defined and tempered.

Proposition 4.1.3. If f and g are tempered functions, we say that is the distributional Fourier transform of f if

$$\int g\varphi = \int f\widehat{\varphi}$$

for all $\varphi \in \mathscr{S}$. For given f, such a function g is unique. We note g by $\widehat{f}$.

Proposition 4.1.4. If $f \in L^1 + L^2$, then its $L^1 + L^2$ Fourier transform coincides with its distributional Fourier transform.

Proposition 4.1.5. A tempered function $\in \mathscr{S}'$, its Fourier transform in the sense of tempered distributional coincides with its distributional Fourier transform if the latter exists.

Proposition 4.1.6. Let $h_a(x) = \gamma\left(\frac{a}{2}\right)\pi^{-\frac{a}{2}}|x|^{-a}$, $a \in \mathbb{C}$ and $x \in \mathbb{R}^n$. Then

$$\widehat{h_a} = h_{n-a}$$

in the sense of $L^1 + L^2$ Fourier transform if $\frac{n}{2} < Re\, a < n$, and in the sense of distributional Fourier transform if $0 < Re\, a < n$. Here γ is the gamma function, its definition as follows

$$\gamma(s) = \int_0^\infty e^{-t}t^{s-1}\,\mathrm{d}t.$$

Proof. Here we divide the proof into three steps.

First step. We claim that if $\frac{n}{2} < a < n$, then $\widehat{h_a} = h_{n-a}$ in the sense of $L^1 + L^2$ Fourier transform. Denote $f_\epsilon(\cdot) = f(\epsilon\,\cdot)$. Let $f(x) = |x|^{-a}$, $\frac{n}{2} < a < n$, then $\widehat{f}$ is radial and $\widehat{f_\epsilon} = \epsilon^{-(n-a)}\widehat{f}$. Hence

$$\widehat{f}(x) = \widehat{f}\left(|x|\frac{x}{|x|}\right) = |x|^{-(n-a)}\,\widehat{f}\left(\frac{x}{|x|}\right) = c\,|x|^{-(n-a)}$$

with c to be determined. Since

$$\int_{\mathbb{R}^n} f(x) e^{-\pi|x|^2} \mathrm{d}x = \int_{\mathbb{R}^n} \widehat{f}(x) e^{-\pi|x|^2} \mathrm{d}x,$$

we have

$$\int_{\mathbb{R}^n} |x|^{-a} e^{-\pi|x|^2} \mathrm{d}x = c \int_{\mathbb{R}^n} |x|^{-(n-a)} e^{-\pi|x|^2} \mathrm{d}x.$$

Note that

$$\begin{aligned} LHS &= \left|\mathbb{S}^{n-1}\right| \int_0^\infty r^{-a+n-1} e^{-\pi r^2} \mathrm{d}r = \left|\mathbb{S}^{n-1}\right| \int_0^\infty e^{-t} \left(\frac{t}{\pi}\right)^{\frac{n-a}{2}} \frac{\mathrm{d}t}{2t} \\ &= \left(\frac{\left|\mathbb{S}^{n-1}\right|}{2}\right) \pi^{-\left(\frac{n-a}{2}\right)} \gamma\left(\frac{n-a}{2}\right), \end{aligned}$$

and

$$RHS = c\left(\frac{\left|\mathbb{S}^{n-1}\right|}{2}\right) \pi^{-\frac{a}{2}} \gamma\left(\frac{a}{2}\right).$$

Hence

$$c = \frac{\pi^{\frac{a}{2}} \gamma\left(\frac{n-a}{2}\right)}{\pi^{\frac{n-a}{2}} \gamma\left(\frac{a}{2}\right)}.$$

Second step. For $\varphi \in \mathscr{S}$, it's suffice to show

$$\int h_a \widehat{\varphi} = \int h_{n-a} \varphi$$

for $0 < Re\, a < n$ in the sense of distributional Fourier transform. Fix $\varphi \in \mathscr{S}$, consider

$$\begin{aligned} A(z) &= \int_{\mathbb{R}^n} h_z \widehat{\varphi} = \int_{\mathbb{R}^n} \frac{\gamma\left(\frac{z}{2}\right)}{\pi^{\frac{z}{2}}} |x|^{-z} \widehat{\varphi}(x)\ \mathrm{d}x \\ B(z) &= \int_{\mathbb{R}^n} h_{n-z} \varphi = \int_{\mathbb{R}^n} \frac{\gamma\left(\frac{n-z}{2}\right)}{\pi^{\frac{n-z}{2}}} |x|^{-(n-z)} \varphi(x)\ \mathrm{d}x \end{aligned}$$

where $0 < Re\, z < n$. Both A and B are analytic when $0 < Re\, z < n$ since γ is analytic and $\int_{\mathbb{R}^n} |x|^{-z} \phi(x)\ \mathrm{d}x$ is analytic when $\varphi \in \mathscr{S}$, which is justified by complex differentiation under the integral sign. By first step $A(x) = B(x)$ when $\frac{n}{2} < x < n$ and the uniqueness theorem, we have $A(z) = B(z)$ when $0 < Re\, z < n$.

Third step. If $\frac{n}{2} < Re\, a < n$, then $h_a \in L^1 + L^2$, so that $L^1 + L^2$ Fourier transform and distributional Fourier transform coincide. $\square$

Let T be an invertible $n \times n$ real symmetric matrix. The signature of T is the number of positive eignvalues $-$ the number of negative eignvalues, counted with multiplicities. Define $G_T(x) = e^{-\pi i \langle Tx, x\rangle}$, observe that $|G_T| \equiv 1$, hence G_T is tempered.

Proposition 4.1.7. Let T be an invertible $n \times n$ real symmetric matrix with sigature σ. Then

$$\widehat{G}_T = e^{-\pi i \frac{\sigma}{4}} \left|\det T\right|^{\frac{1}{2}} G_{-T^{-1}}$$

in the sense of distributional Fourier transform.

Proof. We need to show that

$$\int_{\mathbb{R}^n} e^{-\pi i \langle Tx, x\rangle} \widehat{\varphi}(x)\ \mathrm{d}x = e^{-\pi i \frac{\sigma}{4}} \left|\det T\right|^{-\frac{1}{2}} \int_{\mathbb{R}^n} e^{\pi i \langle T^{-1}x, x\rangle} \varphi(x)\ \mathrm{d}x \tag{4.1.1}$$

Here we also divide the proof into steps.

First step. Consider the $n = 1$ case. Let $\sqrt{z}$ be the branch of the square root defined on the complement of the nonpositive real numbers and positive real axis. Thus $\sqrt{\pm i} = e^{\pm \frac{\pi}{4} i}$. Identity 4.1.1 is equivalent to

$$\int_{\mathbb{R}} e^{-\pi z x^2} \widehat{\varphi}(x)\ \mathrm{d}x = \left(\sqrt{z}\right)^{-1} \int_{\mathbb{R}} e^{-\pi \frac{x^2}{z}} \varphi(x)\ \mathrm{d}x. \tag{4.1.2}$$

If $\varphi \in \mathscr{S}$ and z is purly imaginary and not equal to 0, identity 4.1.2 can be proved by complex analysis method from the real case.

1. If $z > 0$, known;
2. If $z > 0$, by scaling;
3. If $Re\, z > 0$, both side of identity 4.1.2 are analytic in z for $Re\, z > 0$. Use the uniqueness theorem, identity 4.1.2 holds for $Re\, z > 0$;
4. If $z = iT$, both sides of identity 4.1.2 are continuous in z when $Re\, z \geq 0,\ z \neq 0$, thus the desired identity 4.1.2 holds.

Second step. Case $n \geq 2$, claim that if identity 4.1.1 is ture for a given T, it is also true when T is replaced by $U^{-1}TU$ for any $U \in SO(n)$. Indeed, denote that $f(x) = e^{-\pi i \langle Tx, x\rangle}$. Hence $\widehat{f}(x) = e^{-\frac{\pi}{4} i \sigma} \left|\det T\right|^{-\frac{1}{2}} e^{\pi i \langle T^{-1}x, x\rangle}$ in the sense of distributional Fourier transform. Note that

$$\widehat{f \circ U} = \left|\det U\right|^{-1} \widehat{f} \circ U^{-T}$$

in the sense of distributional Fourier transform. Hence,

$$\int_{\mathbb{R}^n} f(Ux)\, \widehat{\varphi}(x)\ \mathrm{d}x = \int_{\mathbb{R}^n} \widehat{f \circ U}(x)\, \varphi(x)\ \mathrm{d}x = \int_{\mathbb{R}^n} \widehat{f}(Ux)\, \varphi(x)\ \mathrm{d}x.$$

Plugging the definition of f gives

$$\int_{\mathbb{R}^n} e^{-\pi i\langle TUx,Ux\rangle}\widehat{\varphi}(x)\,\mathrm{d}x = e^{-\pi i\frac{\sigma}{4}}\left|\det T\right|^{-\frac{1}{2}}\int_{\mathbb{R}^n} e^{\pi i\langle T^{-1}Ux,Ux\rangle}\varphi(x)\,\mathrm{d}x,$$

we then proved the claim. It therefore suffices to prove identity 4.1.1 when T is diagonal. Here we use a fact in T. Wolff's book, which is the proposition $2.1'$. It states that finite linear combinations of $\mathscr{C}_0^\infty$ tensor functions are dense in $\mathscr{S}$. If T is diagonal and φ is $\mathscr{C}_0^\infty$ tensor function, then identity 4.1.1 factor as products of one variable integrals and identity 4.1.1 follows from identity 4.1.2. For the general case, $\varphi \in \mathscr{S}$, $\exists$ a sequence of finite combination of $\mathscr{C}_0^\infty$ tensor functions $\{\varphi_n\}$ s.t. $\varphi_n \to \varphi$ in $\mathscr{S}$. We have

$$\int_{\mathbb{R}^n} e^{-\pi i\langle Tx,x\rangle}\widehat{\varphi_n}(x)\,\mathrm{d}x = e^{-\pi i\frac{\sigma}{4}}\left|\det T\right|^{-\frac{1}{2}}\int_{\mathbb{R}^n} e^{\pi i\langle T^{-1}x,x\rangle}\varphi_n(x)\,\mathrm{d}x,$$

let $n \to \infty$, by proposition 4.1.2,

$$\begin{aligned} RHS &\to e^{-\pi i\frac{\sigma}{4}}\left|\det T\right|^{-\frac{1}{2}}\int_{\mathbb{R}^n} e^{\pi i\langle T^{-1}x,x\rangle}\varphi(x)\,\mathrm{d}x; \\ LHS &\to \int_{\mathbb{R}^n} e^{-\pi i\langle Tx,x\rangle}\widehat{\varphi}(x)\,\mathrm{d}x. \end{aligned}$$

$\square$

Remark. A $\mathscr{C}_0^\infty$ tensor function is a function $f : \mathbb{R}^n \to \mathbb{C}$ of the form $f(x) = \prod_{j=1}^n \phi_j(x_j)$ for $\phi_j \in \mathscr{C}_0^\infty(\mathbb{R})$. $\square$

4.2 The Uncertainty Principle

This part is based on T. Wolff's book chapter 5. Before we introduce some results in the Uncertainty Principle, there is an abstract statement which may summary the idea of the Uncertainty Principle:

(*Heuristic principle*) *It's not possible for both f and its Fourier transform $\widehat{f}$ to be localized on small set.*

A rigerous qualitative weaker version: it's impossible for both $f \in L^2(\mathbb{R}^n)$ and $\widehat{f}$ to be compactly supported, unless $f = 0\ \mathcal{L}^n - a.e.$ (since we know that the Fourier transform of tempered distribution with compact support is a real analytic function)

we will give some quantitative theorem illustrating the uncertainty principle.

Theorem 4.2.1 (Amrein-Berthier Uncertainty Principle)**.** Let E and F be sets of finite measure in $\mathbb{R}^n$. Then

$$\|f\|_{L^2(\mathbb{R}^n)} \leq C\left(\|f\|_{L^2(E^c)} + \left\|\widehat{f}\right\|_{L^2(F^c)}\right)$$

for some constant $C = C(E, F, n)$ and any $f \in L^2(\mathbb{R}^n)$.

Proof. see Schlag Vol.1. theorem 10.4. $\square$

Theorem 4.2.2 (Bernstein's inequality)**.** suppose $f \in L^1 + L^2$, supp $f \subset B_R$, then

(i) for any muti-indices α, and $p \in [1, \infty]$, we have

$$\|D^\alpha f\|_p \precsim_{\alpha,n} R^{|\alpha|}\|f\|_p$$

(ii) for any $1 \geq p \geq q \geq \infty$, we have

$$\|f\|_q \precsim_{p,q,n} R^{n\left(\frac{1}{p}-\frac{1}{q}\right)}\|f\|_p.$$

Proof. one can easily check the correctness of the following reformulation of f via Fourier transform:

$$f = f * \varphi_{R^{-1}}$$

Here, φ is a bump function associated with $\overline{B_1} \subset B_2$ (that is, φ is smooth, $0 \leq \varphi \leq 1$ and $\varphi = 1$ on $\overline{B_1}$, supp $\varphi \subset B_2$)

This reformulation illustrates the fact that $f \in C^\infty \cap C_0$, hence both (i),(ii) follow easily from Minkowski's inequality and Young's inequality.

$$\|D^\alpha f\|_p = \|f * D^\alpha(\varphi_{R^{-1}})\|_p \leq \|f\|_p \|D^\alpha(\varphi_{R^{-1}})\|_1 = R^{|\alpha|}\|f\|_p \|D^\alpha\varphi\|_1 \precsim R^{|\alpha|}\|f\|_p$$

$$\|f\|_q = \|f * \varphi_{R^{-1}}\|_q \leq \|\varphi_{R^{-1}}\|_r \|f\|_p = R^{\left(\frac{1}{p}-\frac{1}{q}\right)}\|\varphi\|_r\|f\|_p \precsim R^{\left(\frac{1}{p}-\frac{1}{q}\right)}\|f\|_p$$

Here, $1/q = 1/p - 1/r'$. $\square$

Theorem 4.2.3 (Heisenberg's uncertainty principle)**.** for any $f \in \mathscr{S}$, we have

$$\|f\|_2^2 \leq \frac{4\pi}{n} \inf_{x_0 \in \mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x)|^2 |x - x_0|^2 \, dx\right)^{\frac{1}{2}} \inf_{\xi_0 \in \mathbb{R}^n}\left(\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 |\xi - \xi_0|^2 \, d\xi\right)^{\frac{1}{2}}.$$

and the equality holds if and only if f is Gaussian function $f = Ae^{-B|x|^2}$ for some $B < 0$.

Proof. it suffices to proof the inequality for $x_0 = \xi_0 = 0$ (replacing f with $fe^{2\pi i\xi_0\cdot(\cdot)}$ and $\tau^{-x_0}f$ to obtain the general case). Firstly, we use integration by parts to reformulate $\|f\|_2^2$ as follows:

$$\begin{aligned}\|f\|_2^2 &= \frac{1}{n}\int f\bar{f}\sum_j \frac{\partial}{\partial x_j}(x_j) = -\frac{1}{n}\int \sum_j x_j\left(\bar{f}\frac{\partial}{\partial x^j}f + f\frac{\partial}{\partial x^j}\bar{f}\right) \\ &= -\frac{2}{n}\operatorname{Re}\sum_j \langle x_j f, D_j f\rangle_{L^2}.\end{aligned}$$

The required inequality then follows via Cauchy-Schwarz and Plancherel:

$$\begin{aligned}\frac{n}{2}\|f\|_2^2 &\le \sum_j \left|\langle x_j f, D_j f\rangle_{L^2}\right| \le \sum_j \|x_j f\|_2 \|D_j f\|_2 \\ &\le \left(\sum_j \|x_j f\|_2^2\right)^{\frac{1}{2}}\left(\sum_j \|D_j f\|_2^2\right)^{\frac{1}{2}} = 2\pi\|xf\|_2\|\xi\hat{f}\|_2\end{aligned}$$

We mainly use the Cauchy-Schwarz above (and keep in mind that we use trivial inequality $-\operatorname{Re} z \le |z|$ at the beginning), hence the equality holds if and only if

$$D_j f = -\beta x_j f \text{ for some common constant } \beta > 0, \quad \forall j$$

Which shows that f is a Gaussian function. $\square$

4.3 Stationary phase and non-stationary phase method

The purpose of this chapter is to study the following integral's behavior when $\lambda \to \infty$:

$$I(\lambda) = \int_{\mathbb{R}^n} a(x)e^{-\pi i\lambda\phi(x)}dx \tag{4.3.1}$$

Here $a \in C_c^\infty$ (so-called amplitude function), $\phi \in C^\infty$ and is real-valued (so-called phase function) .

Remark. Via trivial estimate, $|I(\lambda)| \le \|a\|_1$. However, we hope to obtain the estimate as $|I(\lambda)| = O(\lambda^{-\delta})$ for some $\delta > 0$.

Usually, the amplitude function is less important (just beneficial for restricting the domain of integral), what really matters is the property of the phase function ϕ. To get an intuitive sight, let us consider the case when $n = 1$, $\phi = x$ or x^2, and a is supported in a small neighborhood of 0.

In the case when $\phi = x$, note that the graph of $e^{-\pi i \lambda x}$ oscillates quickly as $\lambda \to \infty$, which suggests the occurrence of frequent cancellations in the integral. In fact, we shall proof the following estimate in this case:

$$|I(\lambda)| = O_N\left(\lambda^{-N}\right), \quad \forall N$$

In the case when $\phi = x^2$, the graph of $e^{-\pi i \lambda x^2}$ oscillates much less near 0, which suggests the slow decreasing rate as $\lambda \to \infty$. In fact, we shall proof the following estimate in this case:

$$|I(\lambda)| = O\left(\lambda^{-\frac{1}{2}}\right)$$

□

Before stating the formal theorems concerning the decreasing rate of $I(\lambda)$, let us review two results from Mathematical analysis (They can be found in Evans PDE), which, informally speaking, turn the general phase function into the special case when $\phi = x$ or x^2.

Proposition 4.3.1 (straightening lemma). Suppose Ω is an open set of $\mathbb{R}^n$, $p \in \Omega$, $f : \Omega \to \mathbb{R}$ is smooth. If $\nabla f(p) \neq 0$, then f can be viewed as a linear function near p. More specifically, there exists U, V, which is a neighborhood of 0 and p separately, and a diffeomorphism $G : U \to V$, such that

$$f \circ G = f(p) + x_n \quad \text{on} \quad U.$$

Proposition 4.3.2 (Morse lemma). Suppose Ω is an open set of $\mathbb{R}^n$, $p \in \Omega$, $f : \Omega \to \mathbb{R}$ is smooth. If $\nabla f(p) = 0$, and $H_f(p)$ is invertible , then f can be viewed as real-valued diagonal quadratic form near p. More specifically, there exists U, V, which is a neighborhood of 0 and p separately, and a diffeomorphism $G : U \to V$, such that

$$(f \circ G)(x) = f(p) + x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 \quad \text{on} \quad U.$$

Here k is the number of positive eigenvalues of $(H_f)(p)$. Besides, the derivations of G are bounded by bounds depending on finitely many derivations of f and a lower bound of $|\nabla f(p)|$.

The method of nonstationary phase is nothing other than integration by parts.

Theorem 4.3.3. Suppose $(\nabla\phi)(p) \neq 0$, a is supported in a sufficiently small neighborhood of p such that ϕ can be viewed as a linear function (see Proposition 4.3.1). Then for any N , we have

$$|I(\lambda)| \leq C_N \left(\lambda^{-N}\right)$$

Here C_N depends on bounds for finitely many derivations of ϕ, a and a lower bound for $|\nabla\phi(p)|$ (and on N of course).

Proof. WLOG, we may take $p = \varphi(p) = 0$. As to the special case when $\phi = x_n$ on supp a , what we are going to deal with is in fact a one-dimension problem. Suppose supp $a \subset [-M, M]^n$ for some $M > 0$, then

$$|I(\lambda)| = \left|\int_{[-M,M]^n} a(x)e^{-\pi i\lambda x_n}dx\right| \leq \int_{[-M,M]^{n-1}} dx_1 \dots dx_{n-1} \left|\int_{-M}^{M} a(x)e^{-\pi i\lambda x_n}dx_n\right|.$$

Fixe $x_1, \dots, x_{n-1}$, via integration by parts, we have:

$$\begin{aligned}\left|\int_{-M}^{M} a(x)e^{-\pi i\lambda x_n}dx_n\right| &= \frac{1}{(\pi\lambda)^N}\left|\int_{-M}^{M} \frac{\partial^N a(x)}{\partial x_n^N}e^{-\pi i\lambda x_n}dx_n\right| \\ &\leq \frac{1}{(\pi\lambda)^N}\int_{-M}^{M}\left|\frac{\partial^N a(x)}{\partial x_n^N}\right| dx_n\end{aligned}$$

Hence

$$|I(\lambda)| \precsim \frac{1}{\lambda^N}\left\|\frac{\partial^N a}{\partial x_n^N}\right\|_{L^1} \precsim_{N,a} \frac{1}{\lambda^N}.$$

The implicit constant here depends on N and $\frac{\partial^N a}{\partial x_n^N}$.

For general ϕ , we take a diffeomorphism $G : U \to V \supset \operatorname{supp} a$ via Proposition 4.3.1, such that $\phi \circ G = x_n$ on U . Using formula of changing variables, we have

$$I(\lambda) = \int_V a(x)e^{-\pi i\lambda\phi(x)}dx = \int_U a(G(x))\,|J_G(x)|\,e^{-\pi i\lambda x_n}dx$$

We thus complete the proof if we view $(a \circ G)\,|J_G|$ as the new amplitude function and use the result proved before. $\square$

The method of nonstationary phase consists is little more complicated. A relative easy approach is to combine the Fourier transform and Talyor expansion.

Proposition 4.3.4. If $\phi(x) = \langle Tx,\, x\rangle$, here T is a $n \times n$ real symmetric matrix with signature σ . Then for any N , we have the following asymptotic formula for $I(\lambda)$:

$$I(\lambda) = e^{-\pi i\frac{\sigma}{4}}|\det T|^{-\frac{1}{2}}\lambda^{-\frac{n}{2}}\left(a(0) + \sum_{j=1}^{N}\lambda^{-N}\mathcal{D}_j a(0) + O_{N,T,a}\left(\lambda^{-N-1}\right)\right).$$

Here $\mathcal{D}_j$ is a linear combination of $\{D^\alpha : |\alpha| = 2j\}$ with coefficient depending on T and j , the $O-$term depends on N, T and finitely many Schwartz seminorms of a .

Proof. Via proposition 4.1.7, we have

$$\begin{aligned}
I(\lambda) &= \int a(x) e^{-\pi i \lambda \langle Tx, x\rangle} dx \\
&= \int \check{a}(\xi) \left(e^{-\pi i \lambda \langle Tx, x\rangle}\right)^{\wedge} (\xi) d\xi \\
&= \int \check{a}(\xi) \left(e^{-\pi i \lambda \langle Tx, x\rangle}\right)^{\wedge} (\xi) d\xi \\
&= e^{-\pi i \frac{\sigma}{4}} \det T\Big|^{-\frac{1}{2}} \lambda^{-\frac{n}{2}} \int \check{a}(\xi) e^{\pi i \lambda^{-1} \left\langle T^{-1}\xi, \xi\right\rangle} d\xi \\
&= e^{-\pi i \frac{\sigma}{4}} |\det T|^{-\frac{1}{2}} \lambda^{-\frac{n}{2}} \int \hat{a}(\xi) e^{\pi i \lambda^{-1} \left\langle T^{-1}\xi, \xi\right\rangle} d\xi
\end{aligned}$$

Expanding the term $e^{\pi i \lambda^{-1} \left\langle T^{-1}\xi, \xi\right\rangle}$ via Talyor formula:

$$e^{\pi i \lambda^{-1} \left\langle T^{-1}\xi, \xi\right\rangle} = 1 + \sum_{j=1}^{N} \frac{\left(\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle\right)^j}{j!} + O_{N,T}\left(\frac{|\xi|^{2N+2}}{\lambda^{N+1}}\right)$$

Hence:

$$\begin{aligned}
&\int \hat{a}(\xi) e^{\pi i \lambda^{-1} \left\langle T^{-1}\xi, \xi\right\rangle} d\xi \\
&= \int \hat{a}(\xi) \left(1 + \sum_{j=1}^{N} \frac{\left(\pi i \lambda^{-1} \langle T^{-1}\xi, \xi\rangle\right)^j}{j!} + O_{N,T}\left(\frac{|\xi|^{2N+2}}{\lambda^{N+1}}\right)\right) d\xi \\
&= \int \hat{a}(\xi) d\xi + \sum_{j=1}^{N} \lambda^{-j} \int \hat{a}(\xi) \frac{\left(\pi i \langle T^{-1}\xi, \xi\rangle\right)^j}{j!} d\xi + O_{N,T}\left(\lambda^{-N-1} \int \hat{a}(\xi) |\xi|^{2N+2} d\xi\right).
\end{aligned}$$

Note that: (i) $\int \hat{a}(\xi) d\xi = a(0)$;

(ii) $\langle T^{-1}\xi, \xi\rangle$ is a homogeneous polynomial of $\xi_1, \ldots, \xi_n$ of order $2j$, hence

$$\hat{a}(\xi) \frac{\left(\pi i \langle T^{-1}\xi, \xi\rangle\right)^j}{j!}$$

is a linear combination of $\left\{\widehat{D^\alpha a}(\xi), |\alpha| = 2j\right\}$.

(iii) The error term is dealt with similarly.

We have now complete the proof. $\square$

Theorem 4.3.5. Suppose $(\nabla\phi)(p) = 0$ and $H_\phi(p)$ is invertible, a is supported in a sufficiently small neighborhood of p such that ϕ can be viewed as a real-valued diagonal quadratic form (see Proposition4.3.1). Let σ be the signature of $H_\phi(p)$, $\Delta = 2^{-n}\left|\det H_\phi(p)\right|$. Then for any N , we have

$$I(\lambda) = e^{-\pi i\lambda\phi(p)}e^{-\pi i\frac{\sigma}{4}}\Delta^{-\frac{1}{2}}\lambda^{-\frac{n}{2}}\left(a(p) + \sum_{j=1}^{N}\lambda^{-j}\mathcal{D}_j a(p) + O\left(\lambda^{-N-1}\right)\right).$$

Here $\mathcal{D}_j$ is a linear combination of $\{D^\alpha : |\alpha| \leq 2j\}$ with coefficient depending on ϕ and j , the $O-$ term depends on N , finitely many derivations of ϕ, a and a lower bound of $|\det H_\phi(p)|$.

As an immediate corollary, if all derivatives of order $< 2j$ of a vanish at p, then $|I(\lambda)| = O\left(\lambda^{-\frac{n}{2}-j}\right)$.

Proof. WLOG, we may take $p = \phi(p) = 0$. Let $G : U \to V$ be the diffeomorphism as in Morse lemma, then via the formula of changing variables, we have:

$$I(\lambda) = \int (a\circ G)(x)\left|J_G(x)\right| e^{-\pi i\lambda\left(x_1^2+\cdots+x_k^2-x_{k+1}^2-x_n^2\right)}dx, \quad k = \frac{n+\sigma}{2}.$$

Denote by b the new amplitude function $(a\circ G)\left|J_G\right|$, then use Proposition 4.3.4, we have

$$I(\lambda) = e^{-\pi i\frac{\sigma}{4}}\lambda^{-\frac{n}{2}}\left(b(0) + \sum_{j=1}^{N}\lambda^{-N}\widetilde{\mathcal{D}}_j b(0) + O\left(\lambda^{-N-1}\right)\right)$$

Here $\widetilde{\mathcal{D}}_j$ is a linear combination of $\{D^\alpha : |\alpha| = 2j\}$ with coefficient depending on ϕ and j .

Via a direct computation from Mathematical Analysis, we have

$$H_{\phi\circ G}(0) = DG(0)^T H_\phi(0) DG(0) \Rightarrow \left|\det H_{\phi\circ G}(0)\right| = \left|J_G(0)\right|^2\left|\det H_\phi(0)\right|.$$

Hence (Note that $\phi\circ G = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2$),

$$\left|J_G(0)\right| = \left(\frac{\left|\det H_{\phi\circ G}(0)\right|}{\left|\det H_\phi(0)\right|}\right)^{\frac{1}{2}} = \Delta^{-\frac{1}{2}} \Rightarrow b(0) = a(0)\left|J_G(0)\right| = a(p)\Delta^{-\frac{1}{2}}$$

Thus, we complete the proof if we expand $\widetilde{\mathcal{D}}_j b(0)$ via product rule to obtain a linear combination of derivations of a at 0 with order $\leq 2j$ and coefficients depending on G (thus, ϕ) and j. (The discussion on $O-$ term is similar, so we omit the details.) □

With Theorem 4.3.3and 4.3.5, we can easily obtain the estimate for the derivatives of $I(\lambda)$.

Proposition 4.3.6. (i) under the same assumptions with Theorem 4.3.3, we have

$$\left|\frac{d^j}{d\lambda^j} I(\lambda)\right| = O_{j,N}\left(\lambda^{-N}\right).$$

(ii) under the same assumptions with Theorem 4.3.5, we have

$$\left|\frac{d^j}{d\lambda^j} e^{\pi i\lambda\phi(p)} I(\lambda)\right| = O_j\left(\lambda^{-\frac{n}{2}-j}\right)$$

Proof. (i) via differential under integral sign and Theorem 4.3.3, we have

$$\left|\frac{d^j}{d\lambda^j} I(\lambda)\right| = \left|\int (-\pi i\phi(x))^j a(x) e^{-\pi i\lambda\phi(x)} dx\right| = O_{j,N}\left(\lambda^{-N}\right).$$

(ii) via differential under integral sign, we have

$$\frac{d^j}{d\lambda^j} e^{\pi i\lambda\phi(p)} I(\lambda) = (-\pi i)^j \int (\phi(x) - \phi(p))^j a(x) e^{-\pi i\lambda(\phi(x)-\phi(p))} dx.$$

Expanding the new amplitude function via Taylor's formula (Note that $\nabla\phi(p) = 0$):

$$\tilde{a}(x) := (\phi(x) - \phi(p))^j a(x) = a(x)\left(\frac{(x-p)^T H_\phi(p)(x-p)}{2} + O\left(|x-p|^3\right)\right)^j.$$

Thus, it is easy to see that $D^\alpha \tilde{a}(p) = 0$ for any muti-indices α with $|\alpha| < 2j$, and we complete the proof via Theorem4.3.5.

$\square$

4.4 Application: Gauss circle problem

Before we introduce the classic Gauss circle problem, we first for further use give an estimate for the Fourier transform of $d\sigma$, where σ is surface measure on $\mathbb{S}^{n-1}$.

By the definition of Fourier transform of finite Borel measure and symmetry of $d\sigma$, we have

$$\widehat{d\sigma}(\xi) = \int_{\mathbb{S}^{n-1}} e^{-2\pi i x\cdot\xi} d\sigma(\xi) = \int_{\mathbb{S}^{n-1}} \cos(2\pi i x\cdot\xi) d\sigma(\xi) \tag{4.4.1}$$

It is easily to see from (4.4.1) that $\widehat{d\sigma}$ is radical and real-valued.

Theorem 4.4.1.

$$\widehat{d\sigma}(\xi) = 2|\xi|^{-\frac{n-1}{2}} \cos\left(2\pi\left(|\xi| - \frac{n-1}{8}\right)\right) + O\left(|\xi|^{-\frac{n+1}{2}}\right).$$

Proof. Since $\widehat{d\sigma}(\xi)$ is radical, we only need to consider the case when $\xi = \lambda e_n$. In the following, $x \in \mathbb{R}^n$ will be denoted by (x', x_n) with $x' \in \mathbb{R}^{n-1}$.

Denoted by U_1, U_2 sufficiently small neighborhoods of the north pole and the south pole of $\mathbb{S}^{n-1}$ separately. More specifically, U_1 is the image of the map

$$x' \to (x', \varphi(x')) := \left(x', \sqrt{1 - |x'|^2}\right)$$

defined on $B_\delta \subset \mathbb{R}^{n-1}$ for some small δ to be chosen later, U_2 is the symmetric to U_1 with the origin.

Add an open set U_3 with respect to $\mathbb{S}^{n-1}$ such that U_3 doesn't contain the north pole as well as the south pole and $\{U_1, U_2, U_3\}$ covers $\mathbb{S}^{n-1}$, take $\{q_1, q_2, q_3\}$ to be the partition of unity subordinate to this chart. We thus divide $\widehat{d\sigma}(\lambda e_n)$ into three parts:

$$I_1 := \int_{\mathbb{S}^{n-1}} e^{-2\pi i\lambda x_n} q_1(x) d\sigma = \int_{B_\delta} q_1(x', \varphi(x')) (1 + |\nabla\varphi(x')|)^{\frac{1}{2}} e^{-2\pi i\lambda\varphi(x')} dx'.$$

$$I_2 := \int_{\mathbb{S}^{n-1}} e^{-2\pi i\lambda x_n} q_2(x) d\sigma = \int_{B_\delta} q_2(x', -\varphi(x')) (1 + |\nabla\varphi(x')|)^{\frac{1}{2}} e^{2\pi i\lambda\varphi(x')} dx'.$$

$$I_3 := \int_{\mathbb{S}^{n-1}} e^{-2\pi i\lambda x_n} q_3(x) d\sigma$$

For simplicity, we denote

$$\tilde{q}_1(x) := q_1(x', \varphi(x')) (1 + |\nabla\varphi(x')|)^{\frac{1}{2}}$$

$$\tilde{q}_2(x) := q_2(x', -\varphi(x')) (1 + |\nabla\varphi(x')|)^{\frac{1}{2}}.$$

Via the direct computation, we have

$$\nabla\varphi(0) = 0, \quad H_\varphi(0) = -2I_{n-1}.$$

Hence via Theorem4.3.5 (take δ small enough such that the assumption on amplitude function is satisfied and note that $\tilde{q}_1(0) = 1$),

$$I_1 = e^{-2\pi i\lambda} e^{\pi i\frac{n-1}{4}} \lambda^{-\frac{n-1}{2}} + O\left(\lambda^{-\frac{n+1}{2}}\right).$$

Similarly,

$$I_2 = e^{2\pi i\lambda} e^{-\pi i\frac{n-1}{4}} \lambda^{-\frac{n-1}{2}} + O\left(\lambda^{-\frac{n+1}{2}}\right).$$

For I_3, we can of course turn the integral on $\mathbb{S}^{n-1}$ into integral on $\mathbb{R}^{n-1}$ (maybe into two parts) and to write I_3 as an oscillation integral with nonstationary phase. Hence by Theorem4.3.3, I_3 can be absorbed into the error term $O\left(\lambda^{-\frac{n+1}{2}}\right)$.

Finally, we obtain

$$\widehat{d\sigma}\left(\lambda e_n\right)=I_1+I_2+I_3=2\cos\left(2\pi\left(\lambda-\frac{n-1}{8}\right)\right)\lambda^{-\frac{n-1}{2}}+O\left(\lambda^{-\frac{n+1}{2}}\right).$$

as desired.

$\square$

Remark. (a) $\widehat{d\sigma}(\xi)$ can be reformulate via Bessel function:

$$\widehat{d\sigma}(\xi)=2\pi|\xi|^{-\frac{n-2}{2}}J_{-\frac{n-2}{2}}(2\pi|\xi|).$$

where

$$J_m(r)=\frac{1}{2\pi}\int_0^{2\pi}e^{ir\sin\theta}e^{-im\theta}d\theta$$

is the Bessel function.

(b) suppose S is a smooth surface with Gaussian curvature nonzero everywhere and surface measure denoted by $d\sigma$, let $d\mu=\psi d\sigma,\psi\in C_0^\infty$. Then we also have

$$|\widehat{d\mu}(\xi)|=O\left(|\xi|^{-\frac{n-1}{2}}\right).$$

$\square$

Next, we introduce the classical Gauss circle problem. Gauss first studied the number of lattice point of $\mathbb{Z}^2$ contained in ball B_μ as $\mu\to\infty$. He obtained that $\#\left(B_\mu\cap\mathbb{Z}^2\right)=|B_\mu|+R$ with $R=O\left(\mu^1\right)$ in 1834 . The more accurate estimate for the remainder have been studied from then on. We list some results obtained recently:

$$R=O_\varepsilon\left(\mu^{\frac{7}{11}+\varepsilon}\right),\quad \text{Iweniec}\ -\ \text{Mozzochi (1988)}$$

$$R=O_\varepsilon\left(\mu^{0.6298\ldots+\varepsilon}\right),\quad \text{Huxley (2002)}$$

$$R=O_\varepsilon\left(\mu^{0.6289\ldots+\varepsilon}\right),\quad Li-\text{Yang(2023)}$$

For a lower bound of the remainder, Hardy obtained $R=\Omega\left(\mu^{\frac{1}{2}}(\log\mu)^{\frac{1}{4}}\right)$. It is thus conjectured that $R=O_\varepsilon\left(\mu^{\frac{1}{2}+\varepsilon}\right)$ which remains open until now.

In this section, we shall discuss this problem in $\mathbb{R}^n$ and first prove Gauss's bound via a simple geometry observation. Then we shall use the estimate for $\widehat{d\sigma}$ established a better estimate for the remainder.

Denote by $Q = \left[-\frac{1}{2}, \frac{1}{2}\right]^n$, every lattice point of $\mathbb{Z}^n$ is corresponding to a unique $\mathbb{Z}^n-$ translation of Q . Notice that if some $\mathbb{Z}^n-$ translation of Q intersects B_μ, then it is contained in $B_{\mu+\frac{\sqrt{n}}{2}}$. Thus

$$\#\left(\mathbb{Z}^n \cap B_\mu\right) \leq \left|B_{\mu+\frac{\sqrt{n}}{2}}\right| = |B_\mu| + O\left(\mu^{n-1}\right)$$

As to the standard estimates of the Gauss problem for $\mathbb{R}^n$, for some $0<\epsilon<1$, one can easily obtain the following inequalities,

$$\sum_{k\in\mathbb{Z}^d} \mathbb{1}_{B_{t-\epsilon}(0)} * \rho_\epsilon(k) \leq \#\left(B_t(0) \cap \mathbb{Z}^n\right) = \sum_{k\in\mathbb{Z}^d} \mathbb{1}_{B_t(0)}(k) \leq \sum_{k\in\mathbb{Z}^d} \mathbb{1}_{B_{t+\epsilon}(0)} * \rho_\epsilon(k). \tag{4.4.2}$$

LHS of inequalities 4.4.2 can be obtained by the following observation

$$\operatorname{supp}\left(\mathbb{1}_{B_{t-\epsilon}(0)} * \rho_\epsilon\right) \subset B_{t-\epsilon}(0) + B_\epsilon(0) \subset B_t(0) \quad \textit{and} \quad 0 \leq \mathbb{1}_{B_{t-\epsilon}(0)} * \rho_\epsilon \leq 1,$$

and RHS of inequalities 4.4.2 can be obtained by the following observation

$$\mathbb{1}_{B_{t+\epsilon}(0)} * \rho_\epsilon|_{B_t(0)} \equiv 1.$$

By the Poisson summation formula,

$$\sum_{k\in\mathbb{Z}^d} \left(\mathbb{1}_{B_{t-\epsilon}(0)} * \rho_\epsilon\right)^\wedge(k) \leq \#\left(B_t(0) \cap \mathbb{Z}^n\right) \leq \sum_{k\in\mathbb{Z}^d} \left(\mathbb{1}_{B_{t+\epsilon}(0)} * \rho_\epsilon\right)^\wedge(k).$$

Set $T = t \pm \epsilon$,

$$\begin{aligned}\sum_{k\in\mathbb{Z}^d} \left(\mathbb{1}_{TB_1(0)} * \rho_\epsilon\right)^\wedge(k) &= T^n \sum_{k\in\mathbb{Z}^d} \widehat{\mathbb{1}_{B_1(0)}}(Tk)\,\widehat{\rho}(\epsilon k) \\ &= T^n\left|B_1(0)\right| + T^n \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \widehat{\mathbb{1}_{B_1(0)}}(Tk)\,\widehat{\rho}(\epsilon k),\end{aligned}$$

denote

$$\widetilde{R}(T) := T^n \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \widehat{\mathbb{1}_{B_1(0)}}(Tk)\,\widehat{\rho}(\epsilon k),$$

then inequalities 4.4.2 becomes

$$\left|B_1(0)\right|(t-\epsilon)^n + \widetilde{R}(t-\epsilon) \leq \sum_{k\in\mathbb{Z}^d} \mathbb{1}_{B_t(0)}(k) \leq \left|B_1(0)\right|(t+\epsilon)^n + \widetilde{R}(t+\epsilon),$$

thus we get

$$O\left(t^{n-1}\epsilon\right) + \widetilde{R}(t-\epsilon) \leq R(t) \leq O\left(t^{n-1}\epsilon\right) + \widetilde{R}(t+\epsilon).$$

Moreover, we have

$$|R(t)| \lesssim \left|\widetilde{R}(t-\epsilon)\right| + \left|\widetilde{R}(t-\epsilon)\right| + t^{n-1}\epsilon.$$

Estimate $\widetilde{R}(T)$, here we use a proposition in Hörmander Vol.1. 3.1.5, that proposition shows that the following equality

$$2\pi i\xi_j\widehat{\mathbb{1}_{B_1(0)}}(\xi) = -\widehat{n_j \mathrm{d}\sigma}(\xi),$$

where σ be the surface measure in $B_1(0)$ and n_j is the j^{th} component of the Gauss map $\overrightarrow{n} : \partial B_1(0) \to S^{n-1}$, i.e. $\forall x \in \partial B_1(0)$, $\overrightarrow{n}(x)$ be the unit exterior normal of x. By the result of $\widehat{\sigma}$, we have

$$\widehat{\mathbb{1}_{B_1(0)}} = O\left(|\xi|^{-(n-1)/2}|\xi|^{-1}\right) = O\left(|\xi|^{-(n+1)/2}\right),$$

therefore, notice that Schwartz function decay fast, $\forall N \geq 0$, we have

$$\begin{aligned}
\left|\widetilde{R}(T)\right| &\lesssim T^n\left(\sum_{1\leq|k|\leq\epsilon^{-1}}\left|\widehat{\mathbb{1}_{B_1(0)}}(Tk)\right| + \sum_{\epsilon^{-1}<|k|}\left|\widehat{\mathbb{1}_{B_1(0)}}(Tk)\right|(1+|\epsilon k|)^{-N}\right)\\
&\lesssim T^n\sum_{1\leq|k|\leq\epsilon^{-1}}|Tk|^{-(n+1)/2} + T^n\sum_{\epsilon^{-1}<|k|}|Tk|^{-(n+1)/2}(1+|\epsilon k|)^{-N}\\
&\lesssim t^{(n-1)/2}\epsilon^{-(n-1)/2},
\end{aligned}$$

Indeed, one can easily verify the following inequalities

$$T^{(n-1)/2} \asymp t^{(n-1)/2}, \quad \sum_{1\leq|k|\leq\epsilon^{-1}}|k|^{-(n+1)/2} \asymp \epsilon^{-(n-1)/2}$$

and

$$\sum_{\epsilon^{-1}<|k|}|k|^{-(n+1)/2}(1+|\epsilon k|)^{-N} \asymp \epsilon^{-(n-1)/2}$$

if we fix some N. Hence we get

$$|R(t)| \lesssim t^{(n-1)/2}\epsilon^{-(n-1)/2} + t^{n-1}\epsilon.$$

Set $t^{(n-1)/2}\epsilon^{-(n-1)/2} = t^{n-1}\epsilon$, then we have $\epsilon = t^{-(n-1)/(n+1)}$. Therefore,

$$|R(t)| \lesssim t^{n-2+2/(n-1)}.$$

Remark. We have proved the following results:

1. (Trivial) $R(t) = O(t^{n-1})$;

2. (Standard) $R(t) = O\left(t^{n-2+2/(n-1)}\right)$.

In fact, for large n, we have some sharp results like the following theorem:

Theorem (sharp). For $n \geq 5$, we have

$$\#\left(\partial B_t(0) \cap \mathbb{Z}^n\right) = O\left(t^{n-2}\right) \quad \textit{and} \quad \Omega\left(t^{n-2}\right).$$

the proof of the above theorem mainly based on the circle method. Actually, it's suffice to estimate the number of the integral roots of the following equation

$$x_1^2 + \cdots + x_n^2 = t^2 \in \mathbb{N}.$$

For $n = 2$, the Gauss circle problem reamins open, and for $n = 3$, the Gauss sphere problem also remains open. For $n = 4$, we have the following theorem:

Theorem .For $n = 4$, we have

$$R(t) = \#\left(B_t(0) \cap \mathbb{Z}^4\right) - |B_1(0)|\, t^4 = O\left(t^{2+\epsilon}\right).$$

$\square$

4.5 Oscillatory integrals (1-dim case)

This part is based on Elias M. Stein's book Harmonic Analysis chapter VIII. Here we consider the following oscillatory integrals

$$I(\lambda) = \int_a^b e^{i\lambda\phi(x)}\psi(x)\, \mathrm{d}x,$$

where $\lambda > 1$, ϕ smooth, real-valued and ψ smooth, complex-valued. Similarily, we have method of nonstationary phase and method of stationary phase . Moreover, we have some uniform estimates for oscillatory integrals.

Proposition 4.5.1. Let ϕ and ψ be smooth functions so that ψ has compact support in (a, b) and $\phi'(x) \neq 0$ for any $x \in [a, b]$. Then

$$I(\lambda) = O\left(\lambda^{-N}\right) \quad \forall N \geq 0.$$

Proof. Let D denote the differential operator

$$Df(x) = \left(i\lambda\phi'(x)\right)^{-1} \frac{\mathrm{d}f}{\mathrm{d}x}$$

and let ${}^tD$ denote its transpose

$$^tDf(x) = -\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f}{i\lambda\phi'(x)}\right).$$

Then,

$$D^N\left(e^{i\lambda\phi(x)}\right) = e^{i\lambda\phi(x)}$$

and integration by parts shows that

$$\int_a^b e^{i\lambda\phi(x)}\psi(x)\,\mathrm{d}x = \int_a^b D^N\left(e^{i\lambda\phi(x)}\right)\psi(x)\,\mathrm{d}x = \int_a^b e^{i\lambda\phi(x)}\left({}^tD\right)^N(\psi)(x)\,\mathrm{d}x,$$

thus we have

$$|I(\lambda)| \lesssim_N \lambda^{-N}.$$

□

Remark. If we set $N = 1$, then

$$|I(\lambda)| \leq \int_a^b \left|{}^tD\psi(x)\right|\,\mathrm{d}x \leq \frac{1}{\lambda}\int_a^b \frac{|\psi'\phi' - \phi''\psi|}{(\phi')^2}\,\mathrm{d}x.$$

□

Remark. By partition of unity, one can naturally obtain the following principle: (Localization principle) The asymptotic behaviour of $I(\lambda)$ is determined by those points in (a, b) with $\phi'(x) = 0$ together with the contribution from the endpoints of the interval. □

The following proposition we called the van der Corput type estimate.

Proposition 4.5.2 (van der Corput)**.** Suppose ϕ is real-valued and smooth in (a, b) and $\left|\phi^{(k)}(x)\right| \geq 1$ for all $x \in (a, b)$. Then

$$\left|\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x\right| \leq C_k\lambda^{-1/k}$$

holds where

(a). $k = 1$ and ϕ' is monotonic;

or

(b). $k \geq 2$.

The constant C_k is independent of ϕ and λ.

Proof. First, assume (a) holds. Notice that

$$\begin{aligned}\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x &= \int_a^b \frac{1}{i\lambda\phi'(x)}\,\mathrm{d}\left(e^{i\lambda\phi(x)}\right)\\ &= \frac{1}{i\lambda\phi'(x)}e^{i\lambda\phi(x)}\Big|_a^b - \int_a^b e^{i\lambda\phi(x)}\left(\frac{1}{i\lambda\phi'(x)}\right)'\,\mathrm{d}x,\end{aligned}$$

hence by the assumption in (a), we have

$$\left|\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x\right| \le 3\lambda.$$

Second, note that when $k=2$, by $|\phi''|\ge 1$ one can get ϕ' is monotonic, hence one can prove by induction on k. Suppose the case k is known, assume that $\phi^{(k+1)}(x)\ge 1$ for all $x\in[a,b]$. Let $x=c$ be the unique point in $[a,b]$ where $\left|\phi^{(k)}(x)\right|$ assume its minimum value. If $\left|\phi^{(k)}(c)\right|=0$, then outside the interval $(c-\delta,c+\delta)$, we have $\left|\phi^{(k)}(x)\right|\ge\delta$, hence

$$\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x = \left(\int_a^{c-\delta} + \int_{c-\delta}^{c+\delta} + \int_{c+\delta}^b\right) e^{i\lambda\phi(x)}\,\mathrm{d}x := I + II + III.$$

For I and III we use the induction hypothesis, then

$$|I| + |III| \le 2C_k(\lambda\delta)^{-1/k},$$

for II we use the trivial estimate,

$$|II| \le 2\delta,$$

hence we get

$$\left|\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x\right| \le 2\delta + 2C_k(\lambda\delta)^{-1/k}.$$

If $\left|\phi^{(k)}(c)\right|\ne 0$, then c is one of the endpoint of $[a,b]$, a similiar argument shows that

$$\left|\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x\right| \le \delta + C_k(\lambda\delta)^{-1/k}.$$

To conclude,

$$\left|\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x\right| \le 2\delta + 2C_k(\lambda\delta)^{-1/k},$$

Set $\delta=(\lambda\delta)^{-1/k}$, then $\delta=\lambda^{-1/(k+1)}$ and

$$\left|\int_a^b e^{i\lambda\phi(x)}\,\mathrm{d}x\right| \le (2C_k+2)\,\lambda^{-1/(k+1)}.$$

Take $C_{k+1}=2C_k+2$ and $C_1=3$, one can easily get $C_k=5\cdot 2^{k-1}-2$, hence we get the proof. $\square$

Remark. For the following oscillatory integrals

$$\int_a^b e^{i\lambda\phi(x)}\psi(x)\,\mathrm{d}x,$$

if we denote $F(x) = \int_a^x e^{i\lambda\phi(t)}\,\mathrm{d}t$, then we have

$$\begin{aligned}\int_a^b e^{i\lambda\phi(x)}\psi(x)\,\mathrm{d}x &= \int_a^b F'(x)\psi(x)\,\mathrm{d}x \\ &= F(b)\psi(b) - F(a)\psi(a) - \int_a^b F(x)\psi'(x)\,\mathrm{d}x.\end{aligned}$$

□

The next proposition is called the method of stationary phase.

Proposition 4.5.3. Suppose $k \geq 2$ and

$$\phi(x_0) = \phi'(x_0) = \cdots = \phi^{(k-1)}(x_0) = 0 \quad \textit{while} \quad \phi^{(k)}(x_0) \neq 0.$$

If ψ is supported in a sufficiently small neighborhood of x_0. For

$$I(\lambda) = \int_a^b e^{i\lambda\phi(x)}\psi(x)\,\mathrm{d}x,$$

we have

$$I(\lambda) \sim \lambda^{-1/k}\sum_{j=0}^{\infty} a_j\lambda^{-j/k}$$

in the sense that for all nonnegative integers N and r,

$$\frac{\mathrm{d}^r}{\mathrm{d}\lambda^r}\left(I(\lambda) - \lambda^{-1/k}\sum_{j=0}^{N} a_j\lambda^{-j/k}\right) = O\left(\lambda^{-r-(N+2)/k}\right) \quad \textit{as } \lambda \to \infty.$$

Remark. 1. Each constant a_j depends on finitely many derivatives of ϕ and ψ at x_0. If $k = 2$,

$$a_0 = \left(\frac{2\pi}{-i\phi''(x_0)}\right)^{1/2}\psi(x_0).$$

The implicit constant depend on upper bounds of finitely many derivaties of ϕ and ψ in the support of ψ, the size of the support of ψ and a lower bound for $\left|\phi^{(k)}(x_0)\right|$.
2. If k is even, then $a_j = 0$ for all odd j. For example, $k = 2$,

$$I(\lambda) = \lambda^{-1/2}\left(a_0 + a_2\lambda^{-2/k} + a_4\lambda^{-4/k} + \cdots + a_N\lambda^{-N/k} + O\left(\lambda^{-(N+2)/k}\right)\right).$$

□

Proof of proposition 4.5.3. As to the simple case, here we first consider the case when $k = 2$, $\phi(x) = x^2$ and $\psi \in \mathscr{S}(\mathbb{R})$. Thus

$$I(\lambda) = \int_{\mathbb{R}} e^{i\lambda x^2} \psi(x) \, \mathrm{d}x = \int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} e^{x^2} \psi(x) \, \mathrm{d}x.$$

By Taylor theorem,

$$e^{x^2} \psi(x) = \sum_{j=0}^{N} b_j x^j + x^{N+1} R_N(x),$$

where

$$R_N(x) = \int_0^1 \frac{(1-t)^N}{N!} \left(e^{x^2} \psi(x)\right)^{(N+1)} (tx) \, \mathrm{d}t \quad \textit{and} \quad b_j = (j!)^{-1} \left(e^{x^2} \psi(x)\right)^{(j)} (0).$$

Then

$$I(\lambda) = \sum_{j=0}^{N} b_j \int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^j \, \mathrm{d}x + \int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x) \, \mathrm{d}x := I + II.$$

For I, denote $y = (1 - i\lambda)^{1/2} x$, here we select an analytic branch of $\sqrt{z}$ for $\lambda > 0$. By Cauchy integration formula, we have

$$\begin{aligned}
&\sum_{j=0}^{N} b_j \int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^j \, \mathrm{d}x \\
=& \sum_{j=0}^{N} b_j \int_{\mathbb{R}} e^{(1-i\lambda)x^2} x^j \, \mathrm{d}x \\
=& \sum_{j=0}^{N} b_j \int_{(1-i\lambda)^{1/2}\mathbb{R}} e^{-y^2} y^j (1 - i\lambda)^{-(j+1)/2} \, \mathrm{d}y \\
=& \sum_{j=0}^{N} b_j \int_{\mathbb{R}} e^{-y^2} y^j (1 - i\lambda)^{-(j+1)/2} \, \mathrm{d}y \\
=& \sum_{j=0}^{N} b_j \left(\int_{\mathbb{R}} e^{-y^2} y^j \, \mathrm{d}x \right) \lambda^{-(j+1)/2} \left(\sum_{l=0}^{L} c_l^{(j)} \lambda^{-l} + O\left(\lambda^{-L-1}\right) \right).
\end{aligned}$$

It remains to consider II,

$$\begin{aligned}
\int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x) \, \mathrm{d}x = &\int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x) \, \alpha\left(\frac{x}{\epsilon}\right) \mathrm{d}x + \\
&\int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x) \left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \mathrm{d}x,
\end{aligned}$$

where $\alpha \in \mathscr{C}_c^\infty(\mathbb{R})$ be a standard bump function which satisfies $\alpha|_{[-1,1]} \equiv 1$, $\operatorname{supp} \alpha \subset [-2, 2]$ and $\alpha \geq 0$. Hence

$$\left| \int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x) \, \alpha\left(\frac{x}{\epsilon}\right) \mathrm{d}x \right| \lesssim_{N,\psi} \epsilon^{N+1} \cdot 1 \cdot \epsilon = \epsilon^{N+2},$$

indeed, this inequality holds by the following trivial estimates:

$$\forall\, x \in \operatorname{supp} \alpha \left(\frac{\cdot}{\epsilon}\right) \subset [-2\epsilon, 2\epsilon]\,, \ \left|x^{N+1}\right| \lesssim \epsilon^{N+1} \ \textit{and} \ \left|e^{-x^2} R_N(x)\right| \lesssim_{N,\psi} 1.$$

Denote $\eta(x) = e^{-x^2} R_N(x)$, $Df(x) = (2i\lambda x)^{-1} \cdot \mathrm{d}f/\mathrm{d}x$ and ${}^tDf(x) = (i\lambda)^{-1} \cdot (\mathrm{d}/\mathrm{d}x)\,[f/2x]$, thus for positive integer M, we have

$$\begin{aligned}
&\int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x) \left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \mathrm{d}x \\
&= \int_{\mathbb{R}} e^{i\lambda x^2} x^{N+1} \eta(x) \left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \mathrm{d}x \\
&= \int_{\mathbb{R}} D^M \left(e^{i\lambda x^2}\right) x^{N+1} \eta(x) \left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \mathrm{d}x.
\end{aligned}$$

Note that

$$\begin{aligned}
x^{N+1}\eta(x)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) &= e^{-x^2} x^{N+1} R_N(x) \left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \\
&= e^{-x^2}\left(e^{x^2}\psi(x) - \sum_{j=0}^{N} b_j x^j\right)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \\
&= \left(\psi(x) - e^{-x^2}\sum_{j=0}^{N} b_j x^j\right)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right)
\end{aligned}$$

and $\psi(x) - e^{-x^2}\sum_{j=0}^{N} b_j x^j \in \mathscr{S}(\mathbb{R})$, therefore, integration by parts, we have

$$\begin{aligned}
&\left|\int_{\mathbb{R}} D^M\left(e^{i\lambda x^2}\right) x^{N+1}\eta(x)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \mathrm{d}x\right| \\
&= \left|\int_{\mathbb{R}} e^{i\lambda x^2} \left({}^tD\right)^M \left(x^{N+1}\eta(x)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right)\right) \mathrm{d}x\right| \\
&\leq \int_{|x|>\epsilon} \left|\left({}^tD\right)^M \left(x^{N+1}\eta(x)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right)\right)\right| \mathrm{d}x.
\end{aligned}$$

By induction,

$$\left|\left({}^tD\right)^M \left(x^{N+1}\eta(x)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right)\right)\right| \lesssim_{N,\psi,M} \lambda^{-M} |x|^{N+1-2M}.$$

Thus

$$\left|\int_{\mathbb{R}} e^{i\lambda x^2} x^{N+1}\eta(x)\left(1 - \alpha\left(\frac{x}{\epsilon}\right)\right) \mathrm{d}x\right| \lesssim \lambda^{-M} \int_{|x|>\epsilon} |x|^{N+1-2M}\, \mathrm{d}x \lesssim \lambda^{-M}\epsilon^{N+2-2M}$$

if $N + 2 - 2M < 0$. To sum up,

$$\left|\int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2} x^{N+1} R_N(x)\, \mathrm{d}x\right| \lesssim \epsilon^{N+2} + \lambda^{-M}\epsilon^{N+2-2M} \lesssim \lambda^{-(N+2)/2},$$

here we set $\epsilon^{N+2}=\lambda^{-M}\epsilon^{N+2-2M}$, i.e. $\epsilon=\lambda^{-1/2}$. Thus

$$\begin{aligned}
I(\lambda) &= \sum_{j=0}^{N} b_j \left(\int_{\mathbb{R}} e^{-y^2} y^j \, \mathrm{d}x\right) \lambda^{-(j+1)/2} \left(\sum_{l=0}^{L} c_l^{(j)} \lambda^{-l} + O\left(\lambda^{-L-1}\right)\right) + O\left(\lambda^{-(N+2)/2}\right). \\
&= \lambda^{-1/2} \sum_{j=0}^{N} a_j \lambda^{-j/2} + O\left(\lambda^{-(N+2)/2}\right).
\end{aligned}$$

For $k=2$ and general ϕ which satisties $\phi(x_0)=\phi'(x_0)=0$, $\phi''(x_0)\neq 0$ and ψ is supported in a sufficiently small neighborhood of x_0. Take ψ's Taylor expansion near x_0, we have

$$\phi(x) = \frac{\phi''(x_0)}{2}(x-x_0)^2 + O\left(|x-x_0|^2\right) = \frac{\phi''(x_0)}{2}(x-x_0)^2(1+\tilde{\epsilon}(x))$$

where $\tilde{\epsilon}(x)=O(|x-x_0|)$. Let $y=(x-x_0)(1+\tilde{\epsilon}(x))^{1/2}$, thus $\phi(x)=\phi''(x_0)\,y^2/2$ and

$$I(\lambda) = \int e^{i\lambda\phi''(x_0)y^2/2}\psi(x(y))\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| \mathrm{d}y,$$

by the special case above, we have

$$\begin{aligned}
I(\lambda) &= \left(\frac{\phi''(x_0)}{2}\lambda\right)^{-1/2} \sum_{j=0}^{N} a_j \left(\lambda\frac{\phi''(x_0)}{2}\right)^{-j/2} + O\left(\lambda^{-(N+2)/2}\right) \\
&= \lambda^{-1/2} \sum_{j=0}^{N} \left[a_j \left(\frac{\phi''(x_0)}{2}\right)^{-(j+1)/2}\right] \lambda^{-j/2} + O\left(\lambda^{-(N+2)/2}\right).
\end{aligned}$$

Notice that

$$\frac{\mathrm{d}^r}{\mathrm{d}\lambda^r}(I(\lambda)) = \int e^{i\lambda\phi(x)}(i\phi(x))^r \psi(x)\,\mathrm{d}x,$$

apply the asymptotics in the case above.

At last, for general $k\geq 3$, similarly we get the proof. □

4.6 An application for Bessel functions

Recall the definition about the Bessel function $J_m(r)$,

$$J_m(r) = \frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta} e^{-im\theta}\,\mathrm{d}\theta$$

for some $m\in\mathbb{Z}$. Denote $\phi(\theta)=\sin\theta$, then $\phi'(\theta)=\cos\theta$ and there are two roots $\theta=\pi/2,\ \theta=3\pi/2$ when $\phi'(\theta)=0$ for $\theta\in[0,2\pi]$. Note that $\phi''(\pi/2)=-1$ and $\phi''(3\pi/2)=1$, hence ϕ has two nondegenerate critial points in $[0,2\pi]$. Choose two bump functions ψ_1 which supports in a sufficiently small neighborhood of $\pi/2$ and equals 1 near $\pi/2$ and ψ_2 which supports in a sufficiently small neighborhood of $3\pi/2$ and equals 1 near $3\pi/2$. Set $\psi_3 = 1 - \psi_1 - \psi_2$, thus we get

$$\begin{aligned}
J_m(r) &= \frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta}e^{-im\theta}\psi_1(\theta)\,\mathrm{d}\theta + \frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta}e^{-im\theta}\psi_2(\theta)\,\mathrm{d}\theta \\
&+ \frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta}e^{-im\theta}\psi_3(\theta)\,\mathrm{d}\theta \\
&:= I + II + III.
\end{aligned}$$

For III, integration by parts twice, we have

$$\begin{aligned}
&\frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta}e^{-im\theta}\psi_3(\theta)\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{ir}\left(e^{ir\sin\theta}\right)' \frac{e^{-im\theta}\psi_3(\theta)}{\cos\theta}\,\mathrm{d}\theta \\
&= \frac{1}{2\pi ir}\left[e^{ir\sin\theta}\frac{e^{-im\theta}\psi_3(\theta)}{\cos\theta}\Big|_0^{2\pi} - \int_0^{2\pi} e^{ir\sin\theta}\left(\frac{e^{-im\theta}\psi_3(\theta)}{\cos\theta}\right)'\,\mathrm{d}\theta\right] \\
&= -\frac{1}{2\pi ir}\int_0^{2\pi} e^{ir\sin\theta}\left(\frac{e^{-im\theta}\psi_3(\theta)}{\cos\theta}\right)'\,\mathrm{d}\theta \\
&= -\left(2\pi r^2\right)^{-1}\int_0^{2\pi} e^{ir\sin\theta}\left(\frac{1}{\cos\theta}\left(\frac{e^{-im\theta}\psi_3(\theta)}{\cos\theta}\right)'\right)'\,\mathrm{d}\theta \\
&= O\left(r^{-2}\right).
\end{aligned}$$

For I, by the proposition 4.5.3, one can easily get the following estimate:

$$\begin{aligned}
&\frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta}e^{-im\theta}\psi_1(\theta)\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}e^{ir}\int_0^{2\pi} e^{ir(\sin\theta-\sin(\pi/2))}e^{-im\theta}\psi_1(\theta)\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}e^{ir}e^{-im\pi/2}\psi_1\left(\frac{\pi}{2}\right)\left(\frac{2\pi}{i\sin(\pi/2)}\right)^{1/2} r^{-1/2} + O\left(r^{-3/2}\right) \\
&= e^{i(r-m\pi/2-\pi/4)}\left(2\pi\right)^{-1/2} r^{-1/2} + O\left(r^{-3/2}\right).
\end{aligned}$$

For II, similarly, we have

$$\frac{1}{2\pi}\int_0^{2\pi} e^{ir\sin\theta}e^{-im\theta}\psi_2(\theta)\,\mathrm{d}\theta = e^{i(r-m\pi/2-\pi/4)}\left(2\pi\right)^{-1/2} r^{-1/2} + O\left(r^{-3/2}\right).$$

Thus,

$$J_m(r) = \sqrt{\frac{2}{\pi}}\cos\left(r - m\frac{\pi}{2} - \frac{\pi}{4}\right) r^{-1/2} + O\left(r^{-3/2}\right), \quad as\ r \to \infty.$$

Remark. Actually, one can use the Bessel function $J_m(r)$ to get the same asymptotic expansion about the $\widehat{\sigma}$, where σ be the measure on $S^{n-1} \subset \mathbb{R}^n$. By the following identity

$$\widehat{\sigma}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} d\sigma(x) = 2\pi |\xi|^{-(n-2)/2} J_{(n-2)/2}(2\pi|\xi|) \quad for\ \xi \in \mathbb{R}^n.$$

and note that we have proved

$$J_m(r) = \sqrt{\frac{2}{\pi}} \cos\left(r - m\frac{\pi}{2} - \frac{\pi}{4}\right) r^{-1/2} + O\left(r^{-3/2}\right), \quad as\ r \to \infty,$$

let $m = (n-2)/2$ and $r = 2\pi|\xi|$, we get

$$\widehat{\sigma}(\xi) = 2|\xi|^{-(n-1)/2} \cos(2\pi(|\xi| - (n-1)/8)) + O\left(|\xi|^{-(n+1)/2}\right).$$

Moreover, one can use the Frostman condition $|\sigma(B_r(x))| \lesssim_{\sigma,n} r^{n-1}$ and the decay behavior $|\widehat{\sigma}(\xi)| \lesssim_{\sigma,n} (1+|\xi|)^{-(n-1)/2}$ to get the Tomas's argument, which is a classical theorem in Fourier restriction theory. □

4.7 Application in a problem in spectral geometry

Here we consider the classical Laplace eigenvalue problem:

Suppose $D \subset \mathbb{R}^d$ be a open, bounded, connected domain, which (piecewise) smooth boundary ∂D. Consider the following Laplacian equation with Dirichlet boundary condition or Neumann boundary condition,

$$\begin{cases} -\Delta u(x) = \lambda^2 u(x), & x \in D; \\ u(x) = 0\,(Dirichlet)\ or\ \frac{\partial u}{\partial n}(x) = 0\,(Neumann), & x \in \partial D. \end{cases}$$

Spectrum: $0 \le \lambda_1^2 < \lambda_2^2 \le \cdots \le \lambda_n^2 \le \cdots \nearrow \infty$(Counting eigenvalues with multiplicity).

Define the following egienvalue counting function

$$\mathcal{N}_D(\mu) := \#\left\{n \in \mathbb{N} : \lambda_n^2 \le \mu^2\right\}.$$

In 1911, Weyl proved the result:

$$\mathcal{N}_D(\mu) = \frac{\omega_d}{(2\pi)^d} \mathrm{Vol}(D)\, \mu^d + o\left(\mu^d\right) \quad as\ \mu \to \infty,$$

where ω_k is the volume of the unit ball in $\mathbb{R}^k$. Nowadays, we call the similar estimates of $\mathcal{N}_D(\mu)$ be the Weyl's law. In 1913, Weyl guessed the following estimate of $\mathcal{N}_D(\mu)$ (the Weyl's conjecture)

$$\mathcal{N}_D(\mu) = \frac{\omega_d}{(2\pi)^d} \mathrm{Vol}(D)\,\mu^d \mp \frac{\omega_{d-1}}{4(2\pi)^{d-1}} \mathrm{Vol}(\partial D)\,\mu^{d-1} + o\left(\mu^{d-1}\right) \quad as\ \mu \to \infty.$$

where "$-$"("$+$") refers to the Dirichlet(Neumann) boundary condition. Belows are some results of Weyl's conjecture:

1. Courant(1920) $O\left(\mu^{d-1}\log\mu\right)$.
2. Corleman(1934, 1936) arbitrary spacial dimension.
3. Levitan(1952), Avakvmovic(1956) $O\left(\mu^{d-1}\right)$ for some class of manifolds.
4. Hörmander(1968, 1969) Fourier integral operators.
5. Duistermant-Guillemin(1975) $o\left(\mu^{d-1}\right)$ under certain assumption.(related to methods in dynamical system)
6. Ivrii(1980), Melrose(1980) $o\left(\mu^{d-1}\right)$ under certain non-periodicity assumption.

In particular, Weyl's conjecture holds for bounded, real-analytic, convex domains, e.g. disks, ellipses, balls, etc. Also, one can consider the remainder $\mathcal{R}_D(\mu)$ in Weyl's law:

$$\mathcal{N}_D(\mu) = c_1\mu^d \mp c_2\mu^{d-1} + \mathcal{R}_D(\mu).$$

Goal: refine the bound of $\mathcal{R}_D(\mu)$ from $o\left(\mu^{d-1}\right)$ to $O\left(\mu^k\right)$ with $k < d-1$ as small as possible.

Known:

1. Lazutkin-Terman(1981) ($\mathbb{R}^2$-case) For each $k < 1$, construct a domain D, s.t. $\mathcal{R}_D(\mu) \neq O\left(\mu^k\right)$.
2. For some special domains in $\mathbb{R}^2$, e.g. squares, disks, ellipses, k can be smaller than 1.

3. As to the lower bound of the remainder Eswarathasan–Polterovich–Toth (2016)

   $\mathbb{R}^2$. $\lambda^{-1}\int_\lambda^{2\lambda}|\mathcal{R}_D(\mu)|\,\mathrm{d}\mu \geq c\lambda^{1/2}$ holds for ellipses.

   $\mathbb{R}^d\,(d\geq 2)$, $D=$ ball. $\lambda^{-1}\int_\lambda^{2\lambda}|\mathcal{R}_D(\mu)|\,\mathrm{d}\mu \geq c\lambda^{d-2+1/2}$.

Here we focus on some classical cases in $\mathbb{R}^2$ with Dirichlet boundary conditions:

Case I. Square $D=[0,1]\times[0,1]$.

By separation of variables, one can easily obtain the eigenvalues of $-\Delta$ have the following representations

$$\lambda^2=\pi^2\left(m^2+n^2\right)$$

for $m,n\in\mathbb{N}$. Hence, we have

$$\mathcal{N}_D(\mu)=\#\{(m,n)\in\mathbb{N}^2:\pi^2\left(m^2+n^2\right)\leq\mu^2\},$$

which is the number of lattice points in

$$\left\{(x,y)\in\mathbb{R}^2: x\geq 0,\ y\geq 0,\ \sqrt{x^2+y^2}\leq\mu/\pi\right\}.$$

Recall the Gauss circle problem, we have the following identity with the remainder

$$\#\left(\mathbb{Z}^2\cap B_t(0)\right)=|B_t(0)|+\textit{remainder}.$$

Thus $\mathcal{R}_D(\mu)$ corresponds to the remainder in the Gauss circle problem. Belows are some results related to the $\mathcal{R}_{square}(\mu)$:

1. Huxley(2003) $\mathcal{R}_{square}(\mu)=O_\epsilon\left(\mu^{131/208+\epsilon}\right)$ where $131/208=0.6298\cdots$.

2. Li-Yang(2023) $\mathcal{R}_{square}(\mu)=O_\epsilon\left(\mu^{\theta+\epsilon}\right)$ where $\theta=0.6289\cdots$. Actually, in Li-Yang's paper, θ is defined in such way that $-\theta/2$ is unique solution to the equation

$$-\frac{8}{25}x-\frac{1}{200}\left(\sqrt{2\left(1-14x\right)}-5\sqrt{-1-8x}\right)^2+\frac{51}{200}=-x$$

   on the interval $[-0.35,-0.3]$.

Case II. Disk $D=B_1(0)\subset\mathbb{R}^2$.

Known:

1. Kuznecov–Fedosov (1965), Colin de Verdière (2011), $\mathcal{R}_{disk}(\mu)=O\left(\mu^{2/3}\right)$.

2. (2019) $\mathcal{R}_{disk}(\mu) = O\left(\mu^{2/3-1/295}\right)$.

3. (2021) $\mathcal{R}_{disk}(\mu) = O_\epsilon\left(\mu^{131/208+\epsilon}\right)$.

Here we shall introduce Yves Colin de Verdière's strategy briefly.

- approximations of eigenvalues. (1-dim stationary phase)
- correspondence betweeen eigenvalues and lattice points.
- eigenvalues counting $\to$ lattice points counting.

We shall only focus on the first step. As it is well known and can be checked by separation of variables, the eigenvalues of disk are the squares of positive zeros $j_{\nu,k}$ of the Bessel function $J_\nu(x)$ where $\nu = n + d/2 - 1$, $n \in \mathbb{Z}^+$ and $k \in \mathbb{N}$. Hence $\nu = n$ for some $n \in \mathbb{Z}^+$ if we set $d = 2$. To approximate zeros of $J_\nu(x)$, we expend $J_\nu(x)$ by stationary phase and Olver's expension

$$J_\nu(x) = \left(\frac{2}{\pi}\right)^{1/2}\left(x^2-\nu^2\right)^{-1/4}\left[\cos\left(\pi x g\left(\frac{\nu}{x}\right) - \frac{\pi}{4}\right) + O_c\left(x^{-1}\right)\right]$$

if $x \geq (1+c)\nu$ for some $c > 0$ with $g(t) = \pi^{-1}\left(\sqrt{1-t^2} - t\arccos t\right)$. Use zeros of the main term $\cos\left(\pi x g\left(\frac{\nu}{x}\right) - \frac{\pi}{4}\right)$ to approximate zeros of $J_\nu(x)$. Hence we get

- $j_{\nu,k} \sim (\nu, k-1/4)$, $\nu \in \mathbb{Z}^+$ and $k \in \mathbb{N}$.
- $j_{\nu,k}^2 \leq \mu^2 \sim (\nu, k-1/4) \in \mu\Omega$ where $\Omega := \{(x,y) \in \mathbb{R}^2 : x, y \geq 0,\ y \leq g(x)\}$.
- lattice counting $\#\left(\mu\Omega \cap \left(\mathbb{Z}^2 - (0,1/4)\right)\right)$.

Remark. There are some results in annulus, balls, etc. $\square$

4.8 Fourier restriction operator

If g is continous on $\mathbb{R}^n (n \geq 2)$, its restriction to a hypersurface $S \subset \mathbb{R}^n$ ($\operatorname{codim}(S) = 1$)(e.g. $S = \mathbb{S}^{n-1}$) is a well-defined function $g\big|_S$. If $f \in L^1(\mathbb{R}^n)$, $\widehat{f}$ is continuous, hence its restriction $\widehat{f}\big|_S$ on S is well-defined. If $f \in L^2(\mathbb{R}^n)$, $\widehat{f} \in L^2(\mathbb{R}^n)$, only defined almost everywhere, is completely arbitrary on a set of measure zero. $\widehat{f}\big|_{S^{n-1}}$ is not well-defined. Recall that if $f \in L^p(\mathbb{R}^n)$ $1 < p < 2$, $\widehat{f} := \widehat{f_1} + \widehat{f_2}$. A natural question is that is $\widehat{f}\big|_{S^{n-1}}$ well-defined? At first sight, $\widehat{f}$ seems to be defined a.e., $\widehat{f}\big|_{S^{n-1}}$ does not make sense.

Remarkable fact: In fact, we have the following theorem which called the Stein-Tomas theorem.

Theorem 4.8.1 (Stein-Tomas)**.** Suppose $n \geq 2$ and $1 \leq p \leq \frac{2(n+1)}{n+3} < 2$,

$$\left\|\widehat{f}\right\|_{L^2(S^{n-1})} \leq C \|f\|_{L^p(\mathbb{R}^n)} \quad \forall f \in \mathscr{S}(\mathbb{R}^n)$$

where $C = C(p)$.

Remark. If $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \frac{2(n+1)}{n+3}$, choose $\{f_n\}_{n=1}^{\infty} \subset \mathscr{S}(\mathbb{R}^n)$ s.t. $f_n \to f \ in\ L^p(\mathbb{R}^n)$ as $n \to \infty$, then $\{f_n\}_{n=1}^{\infty}$ Cauchy in $L^p(\mathbb{R}^n)$, hence by Stein-Tomas theorem, $\{f_n\}_{n=1}^{\infty}$ Cauchy in $L^2(S^{n-1})$, $\exists h \in L^2(S^{n-1})$ s.t. $h = \lim_{n\to\infty} \widehat{f_n}$ in $L^2(S^{n-1})$. Define $h = \widehat{f}\Big|_{S^{n-1}}$, one can check $\widehat{f}\Big|_{S^{n-1}}$ is well-defined easily, also, we have the following priori estimate

$$\left\|\widehat{f}\right\|_{L^2(S^{n-1})} \leq C \|f\|_{L^p(\mathbb{R}^n)} \quad \forall f \in L^p(\mathbb{R}^n), \quad 1 \leq p \leq \frac{2(n+1)}{n+3}.$$

Hence the restriction of Fourier transform on S^{n-1} is well-defined if $1 \leq p \leq 2(n+1)/(n+3)$. □

Formulate the problem.

Suppose S is a compact hypersurface in $\mathbb{R}^n$, for all $f \in L^1(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$(or just $f \in \mathscr{S}(\mathbb{R}^n)$), denote $\mathcal{R}_{p\to q}(S)$ if the following inequality holds

$$\left\|\widehat{f}\right\|_{L^q(S)} \leq C \|f\|_{L^p(\mathbb{R}^n)},$$

where $C = C(p, q, n, S)$. Trivially, by the definition of Fourier transform, we have

$$\left\|\widehat{f}\right\|_{L^\infty(S)} \leq C \|f\|_{L^1(\mathbb{R}^n)}.$$

Below we mainly consider $\mathcal{R}_{p\to 2}(S^{n-1})$.

4.9 Convolution of a Schwartz function with a measure

Let $\mu \in \mathcal{M}(\mathbb{R}^n)$, where $\mathcal{M}(\mathbb{R}^n)$ be the space of finite complex-valued Radon measures on $\mathbb{R}^n$ with the norm $\|\mu\| = |\mu|(\mathbb{R}^n)$, here $|\mu|$ is the total variation measure of μ.(Those definitions can be found in T. Wolff's book) Assume μ has compact support for simplicity. Define

$$\varphi * \mu(x) := \int_{\mathbb{R}^n} \phi(x-y)\, \mathrm{d}\mu(y), \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^n).$$

By D.C.T., one can get $\varphi * \mu \in \mathscr{C}^{\infty}(\mathbb{R}^n)$. Also, denote $\check{\mu}(\cdot) := \widehat{\mu}(-\cdot)$ where $\widehat{\mu}(\xi) := \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi}\, \mathrm{d}\mu(x)$. Below are some basic facts and one can check them by Fubini-Tonelli theorem.

Proposition 4.9.1. Suppose $\mu \in \mathcal{M}(\mathbb{R}^n)$ with compact support, then for any $\varphi \in \mathscr{S}(\mathbb{R}^n)$,

$$\widehat{\varphi\check{\mu}} = \widehat{\varphi} * \mu \quad \textit{and} \quad \widehat{\varphi\mu} = \widehat{\varphi} * \widehat{\mu}.$$

Lemma 4.9.2. Let μ be a finite positive measure(e.g. $\mu = \mathcal{H}^{n-1}\big|_{\mathbb{S}^{n-1}}$ on $\mathbb{S}^{n-1}$). The following are equivalent for any q and any C:

1. $\left\|\widehat{f\mu}\right\|_{L^q(\mathbb{R}^n)} \leq C\|f\|_{L^2(\mu)},\ \forall f \in L^2(\mu);$
2. $\|\widehat{g}\|_{L^2(\mu)} \leq C\|g\|_{L^{q'}(\mathbb{R}^n)},\ \forall g \in \mathscr{S}(\mathbb{R}^n);$
3. $\|\widehat{\mu} * f\|_{L^q(\mathbb{R}^n)} \leq C^2\|f\|_{L^{q'}(\mathbb{R}^n)},\ \forall f \in \mathscr{S}(\mathbb{R}^n).$

Proof. Let $g \in \mathscr{S}(\mathbb{R}^n)$ and $f \in L^2(\mu)$. By Fubini-Tonelli theorem, we have

$$\int_{\mathbb{R}^n} \widehat{g} f\, \mathrm{d}\mu = \int_{\mathbb{R}^n} g\widehat{f\mu}\, \mathrm{d}x. \tag{4.9.1}$$

$1 \Rightarrow 2$. By identity 4.9.1 and Hölder's inequality,

$$\begin{aligned}
\|\widehat{g}\|_{L^2(\mu)} &= \sup_{\|f\|_{L^2(\mu)}\leq 1}\left|\int_{\mathbb{R}^n} \widehat{g} f\, \mathrm{d}\mu\right| = \sup_{\|f\|_{L^2(\mu)}\leq 1}\left|\int_{\mathbb{R}^n} g\widehat{f\mu}\, \mathrm{d}x\right| \\
&\leq \sup_{\|f\|_{L^2(\mu)}\leq 1}\|g\|_{L^{q'}(\mathbb{R}^n)}\left\|\widehat{f\mu}\right\|_{L^q(\mathbb{R}^n)} \leq C\sup_{\|f\|_{L^2(\mu)}\leq 1}\|g\|_{L^{q'}(\mathbb{R}^n)}\|f\|_{L^2(\mu)} \\
&= C\|g\|_{L^{q'}(\mathbb{R}^n)}.
\end{aligned}$$

$2 \Rightarrow 1$. By identity 4.9.1, one can obtain it similarly like above.

$3 \Rightarrow 2$. Notice that

$$\|\widehat{g}\|_{L^2(\mu)} = \int_{\mathbb{R}^n} \widehat{g}\overline{\widehat{g}}\, \mathrm{d}\mu = \int_{\mathbb{R}^n} (\widehat{\mu} * \overline{g})\, g\, \mathrm{d}x \leq \|\widehat{\mu} * \overline{g}\|_{L^q(\mathbb{R}^n)}\|g\|_{L^{q'}(\mathbb{R}^n)} \leq C^2\|g\|^2_{L^{q'}(\mathbb{R}^n)}.$$

$2 \Rightarrow 3$. Notice that

$$\begin{aligned}
\|\widehat{\mu} * f\|_{L^q(\mathbb{R}^n)} = \left\|\widehat{\check{f}\mu}\right\|_{L^q(\mathbb{R}^n)} &= \sup_{g\in\mathscr{S}(\mathbb{R}^n),\ \|g\|_{L^{q'}(\mathbb{R}^n)}\leq 1}\left|\int_{\mathbb{R}^n} \widehat{\check{f}\mu} g\right| \\
&= \sup_{g\in\mathscr{S}(\mathbb{R}^n),\ \|g\|_{L^{q'}(\mathbb{R}^n)}\leq 1}\left|\int_{\mathbb{R}^n} \widehat{g}\check{f}\, \mathrm{d}\mu\right| \\
&\leq \sup_{g\in\mathscr{S}(\mathbb{R}^n),\ \|g\|_{L^{q'}(\mathbb{R}^n)}\leq 1}\|\widehat{g}\|_{L^2(\mu)}\left\|\check{f}\right\|_{L^2(\mu)} \leq C^2\|f\|_{L^{q'}(\mathbb{R}^n)}.
\end{aligned}$$

□

With the lemma above, we can prove Stein-Tomas theorem. Notice that Stein-Tomas throem 4.8.1 is equivalent to the following theorem:

Theorem 4.9.3. For $n \geq 2$ and $1 \leq q' \leq 2(n+1)/(n+3)$ (which is equivalent to $q \geq (2n+2)/(n-1)$),

$$\|\widehat{\sigma} * f\|_{L^q(\mathbb{R}^n)} \leq C_q \|f\|_{L^{q'}(\mathbb{R}^n)}, \ \forall f \in \mathscr{S}(\mathbb{R}^n),$$

or

$$\left\|\widehat{f\sigma}\right\|_{L^q(\mathbb{R}^n)} \leq C \|f\|_{L^2(\mu)}, \ \forall f \in L^2(\sigma).$$

Remark. By Young's inequality, we have

$$\|\widehat{\sigma} * f\|_{L^q(\mathbb{R}^n)} \leq \|\widehat{\sigma}\|_{L^{q/2}(\mathbb{R}^n)} \|f\|_{L^{q'}(\mathbb{R}^n)}.$$

Note that $\widehat{\sigma}(\xi) = O\left((1+|\xi|)^{-(n-1)/2}\right)$, hence if $q/2 \cdot (n-1)/2 > n$, we have $\widehat{\sigma} \in L^{q/2}(\mathbb{R}^n)$. So Stein-Tomas theorem holds for $q > 4n/(n-1)$. Also, one will consider whether the range in Stein-Tomas theorem is sharp or not, the answer is that the range in Stein-Tomas theorem is the best possible! i.e. if $\|\widehat{\sigma} * f\|_{L^q(\mathbb{R}^n)} \leq C_q \|f\|_{L^{q'}(\mathbb{R}^n)}$ holds for any $f \in \mathscr{S}(\mathbb{R}^n)$, then $q \geq (2n+2)/(n-1)$. Here we introduce a classical example which can explain why Stein-Tomas theorem does not hold for some $q < (2n+2)/(n-1)$, it's called the Knapp example: Let $C_\delta = \{x \in S^{n-1} : 1 - x_n \cdot e_n \leq \delta^2\}$, here $e_n = (0, \cdots, 0, 1)$. We often call C_δ be the cap with height δ^2. It's easy to verify

$$C_\delta = S^{n-1} \cap B_{\sqrt{2}\delta}(e_n).$$

Let $f = \mathbb{1}_{C_\delta}$, plug this function in the priori estimate $\left\|\widehat{f\sigma}\right\|_{L^q(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\sigma)}$, notice that

$$\|f\|_{L^2(\sigma)} = \left(\int_{\mathbb{R}^n} \mathbb{1}_{C_\delta}^2 \, \mathrm{d}\sigma\right)^{1/2} \asymp \left(\delta^{n-1}\right)^{1/2}$$

and $f\sigma$ is supported in the rectangle centered at e_n with side length δ^2 in the x_n direction, length about δ in the orthogonal directions. Consider $\widehat{f\sigma}$ on the dual rectangle in particular centered at 0, $|\xi_n| \leq C_1^{-1}\delta^{-2}$, $|\xi_j| \leq C_1^{-1}\delta^{-1}$ for $j = 1, \ldots, n-1$, where C_1 is a large constant. When $\xi \in$ dual rectangle, $x \in C_\delta$, we have

$$|(x - e_n) \cdot \xi| \lesssim C_1^{-1},$$

hence

$$\left|\widehat{f\sigma}(\xi)\right| = \left|\int_{C_\delta} e^{-2\pi i x\cdot\xi}\,\mathrm{d}\sigma(x)\right| = \left|\int_{C_\delta} e^{-2\pi i (x-e_n)\cdot\xi}\,\mathrm{d}\sigma(x)\right|$$
$$\geq \left|\int_{C_\delta} \cos\left(2\pi\left(x-e_n\right)\cdot\xi\right)\,\mathrm{d}\sigma(x)\right|.$$

If C_1 is large enough, we have $|2\pi(x-e_n)\cdot\xi| < \pi/3$ for all $x \in C_\delta$ and all $\xi \in$ dual rectangle. Thus $\left|\widehat{f\sigma}(\xi)\right| \geq \sigma(C_\delta)/2 \asymp \delta^{n-1}$ on dual rectangle. If the priori estimate holds, we have

$$\delta^{n-1}\cdot\delta^{-(n+1)/q} \lesssim \left\|\widehat{f\sigma}\right\|_{L^q(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\sigma)} \lesssim \delta^{(n-1)/2},$$

in particular, we have

$$\delta^{(n-1)/2-(n+1)/q} \lesssim 1 \quad \textit{holds for all sufficiently small } \delta > 0,$$

then taking $\delta \downarrow 0$ we get $(n-1)/2-(n+1)/q \geq 0$, i.e.

$$q \geq 2(n+1)/(n-1).$$

□

4.10 Fourier restriction conjecture

Conjecture 4.10.1 (Stein's conjecture)**.** $\left\|\widehat{f\sigma}\right\|_{L^q(\mathbb{R}^n)} \leq C_q\|f\|_{L^\infty(\mathbb{S}^{n-1})}$ holds for all $q > \frac{2n}{n-1}$. More generally,

$$\left\|\widehat{f\sigma}\right\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|f\|_{L^{q'}(\mathbb{S}^{n-1})} \textit{ whenever } p' > \frac{2n}{n-1},\ 1 \leq q \leq \frac{n-1}{n+1}p',$$

or

$$\left\|\widehat{f}\right\|_{L^q(\mathbb{S}^{n-1})} \lesssim \|f\|_{L^p(\mathbb{R}^n)} \textit{ whenever } p' > \frac{2n}{n-1},\ 1 \leq q \leq \frac{n-1}{n+1}p'.$$

By Stein-Tomas theorem and Hölder's inequality,

$$\left\|\widehat{f\sigma}\right\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\mathbb{S}^{n-1})} \lesssim \|f\|_{L^\infty(\mathbb{S}^{n-1})} \textit{ whenever } q \geq 2(n+1)/(n-1).$$

Results on Stein's conjecture:

1. $\mathbb{R}^2$ Solved;

2. $\mathbb{R}^3$ Conjecture remains open for $q > 3$.

   Stein-Tomas: $q > 4$;

J. Bourgain(1991): $q > 4 - 2/15$;

T. Wolff(2001): Cone $q > 10/3$;

T. Tao(2003): Paraboloid $q > 10/3$;

Bourgain-Guth(2011): $q > 36/17 = 3.29\cdots$;

L. Guth(2016)(polynomial partitioning): $q > 3.25\cdots$;

Hong Wang(2018)(polynomial partitioning): $q > 3 + 3/13$;

Hong Wang ,Shukun Wu(2022 arXiv): $q > 3 + 3/14$.

3. $\mathbb{R}^n (n \geq 4)$

Tomas(1975);

T. Tao(2023);

Hickman ,Rogers(2019);

$\cdots$

Proof of Stein-Tomas Theorem 4.8.1. Only prove the case $q > 2\,(n+1)\,/\,(n-1)$. For the proof of the endpoint one can see Stein's book Harmonic Analysis and Schlag's book. By theorem 4.9.3, it suffices to show that if $q > 2\,(n+1)\,/\,(n-1)$, then

$$\|\widehat{\sigma} * f\|_{L^q(\mathbb{R}^n)} \leq C_q \|f\|_{L^{q'}(\mathbb{R}^n)}, \quad \forall f \in \mathscr{S}(\mathbb{R}^n).$$

Let $\phi \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ be a bump function such that supp$\phi \subset \{1/4 \leq |x| \leq 1\}$ and $\sum_{j\geq 0} \phi(2^{-j}x) = 1$ if $|x| \geq 1$. Then

$$\widehat{\sigma}(\xi) = \left(1 - \sum_{j\geq 0} \phi\left(2^{-j}\xi\right)\right) \widehat{\sigma}(\xi) + \sum_{j\geq 0} \phi\left(2^{-j}\right) \widehat{\sigma}(\xi) = K_{-\infty}(\xi) + \sum_{j=0}^{\infty} K_j(\xi)$$

where $K_{-\infty}(\xi) = \left(1 - \sum_{j\geq 0} \phi(2^{-j}\xi)\right) \widehat{\sigma}(\xi)$ and $K_j(\xi) = \phi(2^{-j})\,\widehat{\sigma}(\xi)$ for $j \geq 0$. We estimate $K_{-\infty} * f$ and $K_j * f$ $\forall j \geq 0$.
Case $K_{-\infty} * f$.

Note that $K_{-\infty} \in \mathscr{C}_c^\infty(\mathbb{R}^n)$, then by Young's inequality, we obtain

$$\|K_{-\infty} * f\|_{L^q(\mathbb{R}^n)} \leq \|K_{-\infty}\|_{L^{q/2}(\mathbb{R}^n)} \|f\|_{L^{q'}(\mathbb{R}^n)}.$$

Case $K_j * f$.

By the decay estimation of $\widehat{\sigma}$ and trivial estimations, one can easily obtain that

$$\|K_j * f\|_{L^\infty(\mathbb{R}^n)} \leq \|K_j\|_{L^\infty(\mathbb{R}^n)} \|f\|_{L^1(\mathbb{R}^n)} \lesssim 2^{-j\cdot(n-1)/2} \|f\|_{L^1(\mathbb{R}^n)}$$

and

$$\|K_j * f\|_{L^2(\mathbb{R}^n)} = \left\|\widehat{K_j}\widehat{f}\right\|_{L^2(\mathbb{R}^n)} \leq \left\|\widehat{K_j}\right\|_{L^\infty(\mathbb{R}^n)} \|f\|_{L^2(\mathbb{R}^n)}. \tag{4.10.1}$$

Hence we need to estimate $\left\|\widehat{K_j}\right\|_{L^\infty(\mathbb{R}^n)}$. Note that σ is radial, hence we have

$$\widehat{K_j}(\xi) = \widehat{\phi(2^{-j}\cdot)} * \sigma(\xi) = \int_{S^{n-1}} 2^{jn}\widehat{\phi}\left(2^j(\xi-\eta)\right) \mathrm{d}\sigma(\eta).$$

By the fast decay condition of the Schwartz function $\widehat{\phi}$, for some $N \geq 0$, we have

$$\begin{aligned}
&\left|\widehat{K_j}(\xi)\right| \\
&\lesssim 2^j \int_{S^{n-1}} \left(1+\left(2^j|\xi-\eta|\right)\right)^{-N} \mathrm{d}\sigma(\eta) \\
&= 2^{jn} \int_{S^{n-1}\cap B_{2^{-j}}(\xi)} \left(1+\left(2^j|\xi-\eta|\right)\right)^{-N} \mathrm{d}\sigma(\eta) + \\
&\quad 2^{jn} \sum_{k=0}^{\infty} \int_{S^{n-1}\cap\left\{\eta:2^k\leq 2^j|\xi-\eta|<2^{k+1}\right\}} \left(1+\left(2^j|\xi-\eta|\right)\right)^{-N} \mathrm{d}\sigma(\eta) \\
&\lesssim 2^{jn}\sigma\left(S^{n-1}\cap B_{2^{-j}}(\xi)\right) + \sum_{k=0}^{\infty} 2^{-kN}\sigma\left(S^{n-1}\cap\left\{\eta : 2^k \leq 2^j|\xi-\eta| < 2^{k+1}\right\}\right),
\end{aligned}$$

It's easy to verify

$$\sigma\left(S^{n-1}\cap B_{2^{-j}}(\xi)\right) \lesssim 2^{-j(n-1)}$$

and

$$\sigma\left(S^{n-1}\cap\left\{\eta : 2^k \leq 2^j|\xi-\eta| < 2^{k+1}\right\}\right) \lesssim 2^{(k-j+1)(n-1)} \lesssim 2^{(k-j)(n-1)}.$$

Thus

$$\left|\widehat{K_j}(\xi)\right| \lesssim 2^{jn}\left(2^{-j(n-1)} + \sum_{k=0}^{\infty} 2^{-kN}2^{(k-j)(n-1)}\right) = 2^j\left(1+\sum_{k=0}^{\infty} 2^{-k(N-(n-1))}\right) \lesssim 2^j.$$

Plugging the above inequality into the inequality 4.10.1, we have

$$\|K_j * f\|_{L^2(\mathbb{R}^n)} \lesssim 2^j \|f\|_{L^2(\mathbb{R}^n)}.$$

By the Riesz-Thorin interpolation theorem, we have

$$\|K_j * f\|_{L^q(\mathbb{R}^n)} \lesssim \left(2^{-j\cdot(n-1)/2}\right)^{1-t} \left(2^j\right)^t \|f\|_{L^p(\mathbb{R}^n)}$$

with

$$\frac{1}{p}=\frac{1-t}{1}+\frac{t}{2},\quad \frac{1}{q}=\frac{1-t}{\infty}+\frac{t}{2}.$$

Note that $p=q',\ t=2/q$. Thus

$$\|K_j*f\|_{L^q(\mathbb{R}^n)}\lesssim\left(2^j\right)^{(n+1)/q-(n-1)/2}\|f\|_{L^{q'}(\mathbb{R}^n)}$$

for any $q\geq 2$. If $q>2\,(n+1)\,/\,(n-1)$, then $(n+1)\,/q-(n-1)\,/2<0$ which leads to

$$\sum_{j\geq 0}\|K_j*f\|_{L^q(\mathbb{R}^n)}\lesssim\|f\|_{L^{q'}(\mathbb{R}^n)}\,.$$

Finally, by Fatou's lemma, we obtain

$$\begin{aligned}
\|\widehat{\sigma}*f\|_{L^q(\mathbb{R}^n)}&=\left\|\left(K_{-\infty}(\xi)+\sum_{j=0}^{\infty}K_j(\xi)\right)*f\right\|_{L^q(\mathbb{R}^n)}\\
&\leq\|K_{-\infty}*f\|_{L^q(\mathbb{R}^n)}+\liminf_{N\to\infty}\left\|\left(\sum_{j=0}^{N}K_j\right)*f\right\|_{L^q(\mathbb{R}^n)}\\
&\lesssim\|f\|_{L^{q'}(\mathbb{R}^n)}\,.
\end{aligned}$$

□

4.11 Hausdorff measures

Fix $\alpha\geq 0$ and let $E\subset\mathbb{R}^n$. For $\epsilon>0$, one defines

$$H_\alpha^\epsilon(E)=\inf\left\{\sum_{j\in\mathbb{N}}r_j^\alpha\right\}$$

where the infimum is taken over all countable coverings of E by open balls $B_{r_j}(x_j)$ with $r_j<\epsilon$. It's easy to check $H_\alpha^\epsilon(E)\nearrow$ *as* $\epsilon\searrow 0$ and we define

$$H_\alpha(E)=\lim_{\epsilon\downarrow 0}H_\alpha^\epsilon(E)\,.$$

$H_\alpha(E)$ is called the (outer) Hausdorff measure of E.

Belows are some basic proposition.

- H_α is a metric outer measure, i.e. for any $E,F\subset\mathbb{R}^n$ satisfy $\operatorname{dist}(E,F):=\inf_{x\in E,y\in F}\operatorname{dist}(x,y)>0$, then $H_\alpha(E\cup F)=H_\alpha(E)+H_\alpha(F)$;

- $H_\alpha(E)$ is a nonincreasing function of α.

- $H_\alpha(E) = 0$ for all E if $\alpha > n$.

Lemma 4.11.1. There is a unique number α_0, called the Hausdorff dimension of E, s.t.

$$H_\alpha(E) = \infty \; if \; \alpha < \alpha_0 \; and \; H_\alpha(E) = 0 \; if \; \alpha > \alpha_0.$$

The proof of the lemma 4.11.1 is obvious. Actually, one can calculate some examples, belows are some basic results.

1. $1/3-$Cantor set $\mathcal{C}$. Hausdorff dimension of $\mathcal{C} = \log 2/\log 3$;

2. $\mathbb{R}^3$. $dim\,(a\ point) = 0$, $dim\,(an\ interval) = 1$, $dim\,(a\ square) = 2$.

4.12 Besicovitch sets, Kakeya conjecture

A Besicovitch set (or a Kakeya set) is a compact set $E \subset \mathbb{R}^n$ which contains a unit line segment in every direction, i.e. for any $e \in S^{n-1}$, $\exists x \in E$, s.t. $\{x + te : -1/2 \le t \le 1/2\} \subset E$.
$\mathbb{R}^2$. Kakeya needle problem : What is the smallest area you need in order to turn a unit needle around by 180°?(line segment can be continuously moved i.e. translated and rotated)
Question: determine the minimum area of a set in $\mathbb{R}^2$ that contains unit line segments in all directions.

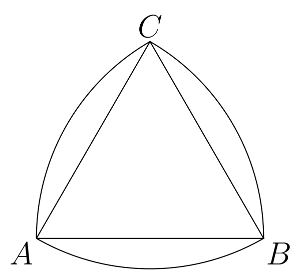
Theorem 4.12.1 (Besicovitch 1920)**.** If $n \ge 2$, then there are Kakeya sets in $\mathbb{R}^n$ which Lebesgue measure 0.

The Kakeya conjecture :

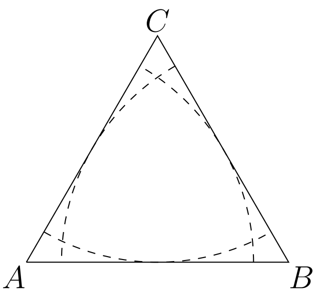
If $E \subset \mathbb{R}^n$ is a Kakeya set. E has Hausdorff dimension n.

Remark. Restriction conjecture $\Rightarrow$ Kakeya conjecture. □

Example. In $\mathbb{R}^2$, then the ball $B_{1/2}(0)$ is a Kakeya set. Area $= \pi/4 = 0.7853\cdots$;

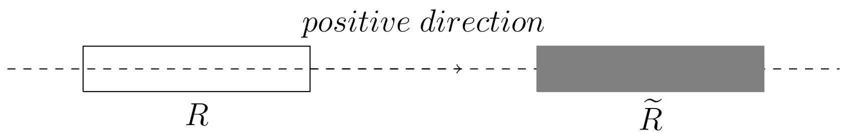


For the above figure, $\triangle ABC$ is an equilateral triangle with length 1. $\overset{\frown}{AB}$ is a part of a circle of radius 1 which has the center C, $\overset{\frown}{BC}$ is a part of a circle of radius 1 which has the center A and $\overset{\frown}{AC}$ is a part of a circle of radius 1 which has the center B. This set is a Besicovitch set. Area $= 3\left(\pi/6 - \sqrt{3}/4\right) = 0.7047\cdots$;



For the above figure, $\triangle ABC$ is an equilateral triangle with length $2/\sqrt{3}$. Area $= 1/\sqrt{3} = 0.5773\cdots$. This set is the optimal among convex shapes(Pal 1921).

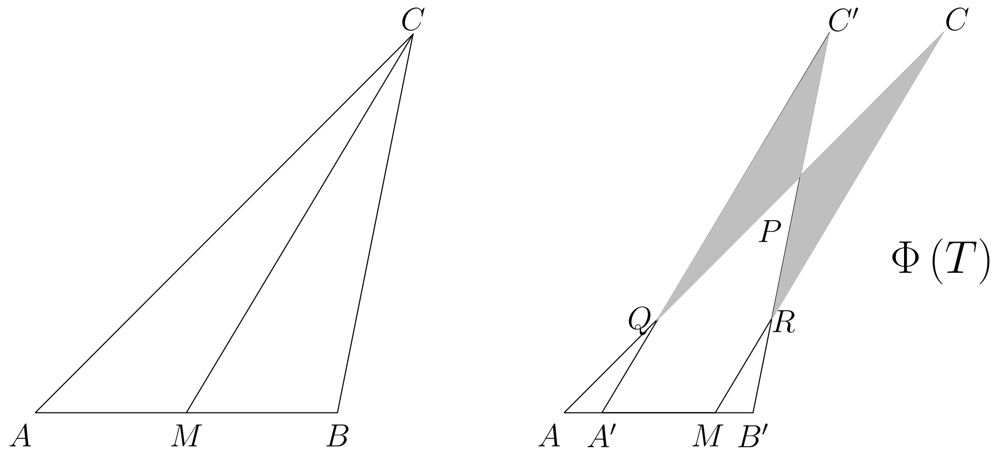
Definition 4.12.2 (The reach of a rectangle of length 1 and 2^{-N})**.** For the rectangle R of length 1 and 2^{-N}, one defines its reach $\widetilde{R}$ be the rectangle which generated by translation of R by 2 units in the positive direction along the longer side of R.



Theorem 4.12.3. Given any $\epsilon > 0$, $\exists$ an integer $N = N_\epsilon$ and 2^N rectangles $R_1, \ldots, R_{2^N}$ each having side lengths 1 and 2^{-N}. So that

1. $\left|\bigcup_{j=1}^{2^N} R_j\right| < \epsilon$;
2. The $\widetilde{R}_j$ are munutally disjoint $j = 1, \ldots, 2^N$ and $\left|\bigcup_{j=1}^{2^N} \widetilde{R}_j\right| = 1$.

Construction: Fix a constant of proportionality $\alpha \in (1/2, 1)$, we translate the "right" triangle leftwards to obtain the overlapping figure $\Phi(T)$.



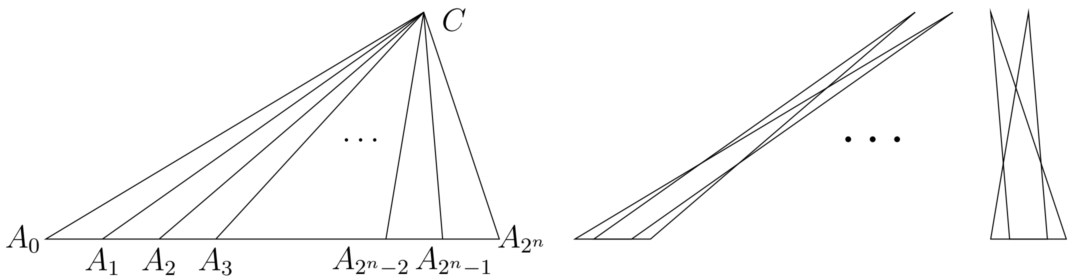
Denote $\Phi_h(T) = \triangle AB'P$ and $\Phi_a(T) =$ two small shaded triangles $\triangle C'QP \cup \triangle PRC$. Set $\alpha = |AB'|/|AB|$. $\Phi_h(T)$ similar to $\triangle ABC(:= T)$. It's easy to obtain that

$$|\Phi_h(T)| = \alpha^2 |T| \quad \textit{and} \quad |\Phi_a(T)| = 2(1-\alpha)^2 |T|.$$

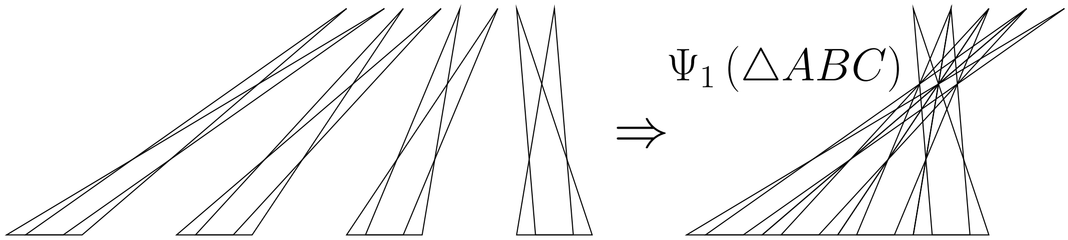
Therefore

$$|\Phi(T)| = |\Phi_h(T)| + |\Phi_a(T)| = \left(\alpha^2 + 2(1-\alpha)^2\right) |T|.$$

An $n-$fold iteration of this construction. Subdivide AB into 2^n equal subintervals, hence $\triangle ABC$ is divided into 2^{n-1} disjoint smaller triangles $\triangle A_{2j}A_{2j+2}C$ for $0 \leq j \leq 2^{n-1}$.



Fix α, construct the figure $\Phi(\triangle A_{2j}A_{2j+2}C)$. In so doing, we obtain 2^{n-1} "hearts" and 2^{n-1} pairs of "arms". The right side of $\Phi_h(\triangle A_{2j}A_{2j+2}C)$ is parallel to the line CA_{2j+2} as is the left side of $\Phi_h(\triangle A_{2j+2}A_{2j+4}C)$ for $0 \leq j \leq 2^{n-1}-1$. Thus the triangle $\triangle A_{2j+2}A_{2j+4}C$ can be moved leftwords so that the left side of $\Phi_h(\triangle A_{2j+2}A_{2j+4}C)$ coincides with the right side of $\Phi_h(\triangle A_{2j}A_{2j+2}C)$. We can incorporate each of 2^{n-1} hearts into one composite heart which is similar to the original triangle $\triangle ABC$.



To summarize, we have translated the 2^n subtriangles of $\triangle ABC$ forming a figure that we call $\Psi_1(\triangle ABC)$. It contains a "heart" namely the disjoint union of the translates of $\Phi_h(\triangle A_{2j}A_{2j+2}C)$. The rest of $\Psi_1(\triangle ABC)$ consists of the union of the translated $\Phi_a(\triangle A_{2j}A_{2j+2}C)$ which we refer to as the "arms" of $\Psi_1(\triangle ABC)$. Hence

$$\begin{aligned}
|heart\ of\ \Psi_1(\triangle ABC)| &= \sum_{j=0}^{2^{n-1}-1} |\Phi_h(\triangle A_{2j}A_{2j+2}C)| \\
&= \sum_{j=0}^{2^{n-1}-1} \alpha^2 |\triangle A_{2j}A_{2j+2}C| = \alpha^2 |T|,
\end{aligned}$$

and

$$\begin{aligned}
|arms\ of\ \Psi_1(\triangle ABC)| &\leq \sum_{j=0}^{2^{n-1}-1} |\Phi_a(\triangle A_{2j}A_{2j+2}C)| \\
&\leq \sum_{j=0}^{2^{n-1}-1} 2(1-\alpha)^2 |T|/2^{n-1} = 2(1-\alpha)^2 |T|.
\end{aligned}$$

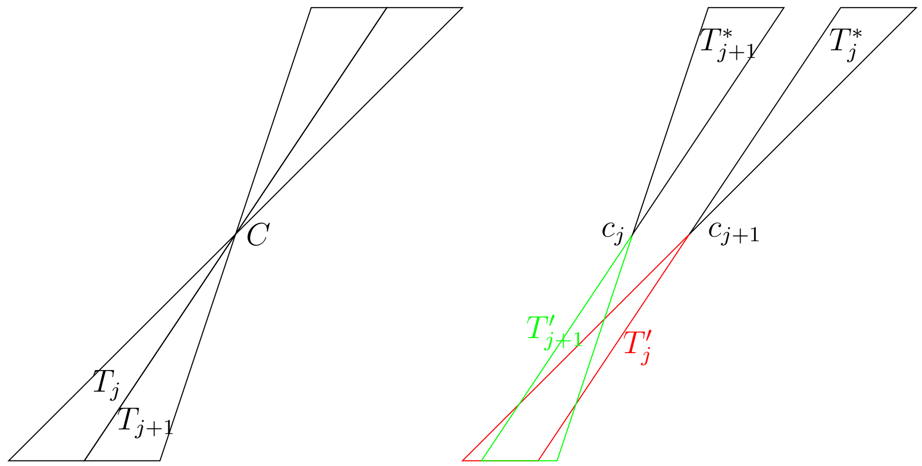
Thus

$$|\Psi_1(\triangle ABC)| \leq \left(\alpha^2 + 2(1-\alpha)^2\right)|T|.$$

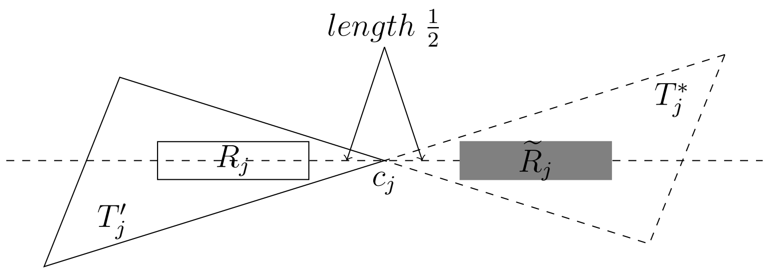
(Iteration)The heart of $\Psi_1(\triangle ABC)$ is the union of 2^{n-1} triangles and so we carry out the above process on the heart of $\Psi_1(\triangle ABC)$ with n replaced by $n-1$. The area of its heart is $\alpha^2(\alpha^2|T|)$, the area of the additional arms generated at this stage will not exceed $2(1-\alpha)^2\alpha^2|T|$. We continue in this way, finally obtain $\Psi_n(\triangle ABC)$ and

$$\begin{aligned}
|\Psi_1(\triangle ABC)| &\leq \left(\left(\alpha^2\right)^n + 2(1-\alpha)^2 + 2(1-\alpha)^2\alpha^2 + \cdots + 2(1-\alpha)^2\alpha^{2n-2}\right)|T| \\
&= \left(\alpha^{2n} + 2(1-\alpha)^2\frac{1-\alpha^{2n}}{1-\alpha^2}\right)|T| \\
&\leq \left(\alpha^{2n} + 2(1-\alpha)\right)|T|,
\end{aligned}$$

where $\alpha \in (1/2, 1)$. If $\alpha \uparrow 1$ and n large, then $\alpha^{2n} + 2(1-\alpha)$ can be made as small as we wish. For $T_j = \triangle A_jA_{j+1}C$, let T_j' denote the corresponding translated triangles comprising $\Psi_n(\triangle ABC)$ and c_j be the vertices of the T'_j. Denote by T^*_j the $\triangle$ obtained by reflecting T'_j through c_j. T'_j overlap a lot, T^*_j are mutally disjoint.



Fix $T = \triangle ABC$ is an equilateral triangle with height 2. Let R_j denote the rectangle whose major axios is P_1P_2 whose side length are 1 and 2^{-N}. $N = n + c_1$, c_1 is a fixed large integer(independent of n). Let $\widetilde{R}_j \subset T^*_j$ be the reflection of R_j through c_j. We have 2^n rectangles R_j of size 1×2^{-N} so that reaches $\widehat{R}_j$ are disjoint.



Take 2^{c_1} disjoint translates of such sets, finally gives us 2^N rectangles. With size 1×2^{-N} which we contained in a set of measure $\le 2^{c_1}\left[\alpha^{2n} + 2(1-\alpha)\right]|T|$ with disjoint reaches. Let $\alpha \uparrow 1$, n large, $N = n + c_1$, we can make the measure abitrarily small.

Remark. Theorem(Fefferman 1971). T_B is unbounded in $L^p(\mathbb{R}^2)$ if $p \neq 2$. Here

$$(T_B f)^\wedge = \mathbb{1}_B \widehat{f} \quad and \quad B = \left\{x \in \mathbb{R}^2 : |x| \le 1\right\}.$$

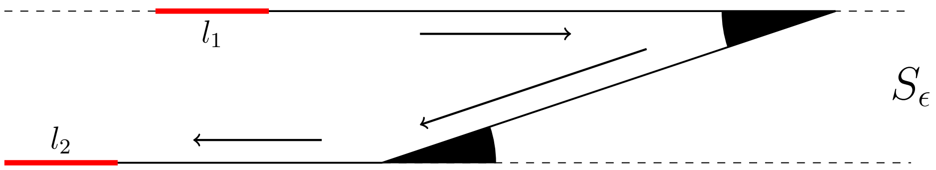
One can see Stein Harmonic Analysis page 450 §2.5 "Unboundness of disk multiplier". $\square$

Results:

- For each $\epsilon > 0$, $\exists$ a set K with $|K| < \epsilon$, that contains unit line segments in all possible orientations.

- For each $\epsilon > 0$, there is a set $\widetilde{K}$, with $\left|\widetilde{K}\right| < \epsilon$ and inside of which a unit line segment can be moved continuously, changing its orientation by 180°.(Kakeya needle problem)

  Observation: Suppose l_1, l_2 are two parallel unit line segment in the plane. Then for any $\epsilon > 0$, there is a set S_ϵ $|S_\epsilon| < \epsilon$, so that, with S_ϵ, l_1 can be continuously moved into the position l_2. Below is a construction which called the Pal joint.



- There is a compact set K, with $|K| = 0$ that contains unit line segment in all direction.

- Suppose π is a parallelogram. Then given any $\epsilon > 0$, we can find parallelograms $\pi_i \subset \pi$. $\left|\cup_{i=1}^{N}\pi_i\right| < \epsilon$ and so that any line segment in π that joints the line $y = 0$ and $y = 1$ has a translate that is contained in one of the π_i. One can see some details in Stein's book Harmonic Analysis chapter X §3. Futher results.

References

[1] Duoandikoetxea, J., Fourier analysis. Translated and revised from the 1995 Spanish original by David Cruz-Uribe Grad. Stud. Math., 29 American Mathematical Society, Providence, RI, 2001. xviii+222 pp.

[2] Folland, G. B., Real analysis. Modern techniques and their applications. Second edition Pure Appl. Math. (N. Y.) Wiley-Intersci. Publ. John Wiley & Sons, Inc., New York, 1999. xvi+386 pp.

[3] Grafakos, L., Classical Fourier analysis. Third edition Grad. Texts in Math., 249 Springer, New York, 2014. xviii+638 pp.

[4] Grafakos, L., Modern Fourier analysis. Third edition Grad. Texts in Math., 250 Springer, New York, 2014. xvi+624 pp.

[5] Muscalu, C., Schlag, W., Classical and multilinear harmonic analysis. Vol. I. Cambridge Stud. Adv. Math., 137 Cambridge University Press, Cambridge, 2013. xviii+370 pp.

[6] Muscalu, C., Schlag, W., Classical and multilinear harmonic analysis. Vol. II. Cambridge Stud. Adv. Math., 138 Cambridge University Press, Cambridge, 2013. xvi+324 pp.

[7] Stein, E. M., Singular integrals and differentiability properties of functions. Princeton Math. Ser., No. 30 Princeton University Press, Princeton, NJ, 1970. xiv+290 pp.

[8] Stein, E. M., Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.(English summary) With the assistance of Timothy S. Murphy Princeton Math. Ser., 43 Monogr. Harmon. Anal., III Princeton University Press, Princeton, NJ, 1993. xiv+695 pp.

[9] Stein, E. M., Weiss, G., Introduction to Fourier analysis on Euclidean spaces. Princeton Math. Ser., No. 32 Princeton University Press, Princeton, NJ, 1971. x+297 pp.