NOTES - 8630

Paul Simanjuntak

12/12/2019

Remark: This is my (slightly more complete, but with far less foresight) notes for Harmonic Analysis I, as taught by Steve Hoffman in University of Missouri.

1 Hardy-Littlewood Maximal Operator

We begin with a definition:

Definition 1.1. Let f be a measurable function. The Hardy-Littlewood maximal function is given by

$$Mf := \sup_{Q \ni x} \fint_Q |f(y)| \, dy$$

where Q is a cube (not necessarily centered) containing x.

From this definition, we state the following theorem, which proof we will delay.

Theorem 1.2 (Hardy-Littlewood Maximal). *The maximal operator $M : f \mapsto Mf$ is of weak type $(1,1)$, i.e. for all $\lambda > 0$,*

$$|\{x \in \mathbb{R}^n \mid Mf(x) > \lambda\}| \leq \frac{C_n}{\lambda} \|f\|_1$$

for some constant C_n depending only on n.

Note. The weak type $(1,1)$ (or weak L^1) is weaker than L^1. In general, $Mf \notin L^1$ even when $f \in L^1$ itself. For an example of this, consider the characteristic of the unit ball χ_B. It can be computed that

$$M\chi_B(x) \equiv |x|^{-n}$$

outside some sufficiently large ball centered at the origin. Clearly such function is not integrable over $\mathbb{R}^n$, because writing in polar coordinate,

$$\int_{\mathbb{R}^n} M\chi_B(x) \, dx \equiv C \int_0^\infty \frac{dr}{r}$$

While we save the proof of Theorem 1.2 for later, we'll apply the theorem immediately to prove the following important fact:

Theorem 1.3 (Lebesgue Differentiation). *Let $f \in L^1_{loc}(\mathbb{R}^n)$ for $n \geq 1$. Then,*

$$\lim_{Q \to x} \fint_Q f(y) \, dy = f(x) \qquad \textit{a.e. on } \mathbb{R}^n$$

where Q is a cube containing x.

Proof. Let $f \in L^1(\mathbb{R}^n)$ and define

$$\Lambda f(x) = \overline{\lim_{r\to 0}} \fint_{Q_r(x)} f - \underline{\lim_{r\to 0}} \fint_{Q_r(x)} f$$

where $Q_r(x)$ is a cube centered at x with side lenght r. We will show first that the wanted limit exists a.e. and show that the limit is indeed $f(x)$. The first result is given by the following claim:

Claim. *$\Lambda f(x) = 0$ a.e. on $\mathbb{R}^n$.*

Proof. Observe that if $f \in C_c(\mathbb{R}^n)$, then $\Lambda f(x) = 0$. Indeed for $x \in \operatorname{supp} f$, let M_r and m_r be supremum and infimum of f on $Q_r(x)$ (for fixed $r > 0$). Then,

$$\Lambda f(x) \le M_r - m_r \qquad (\Lambda f \ge 0)$$

Because f is continuous, $M_r - m_r \to 0$ as $r \to 0$, which proves the claim for the class C_c.
Next, by density of this class, we may choose a sequence $\{f_k\} \subseteq C_c(\mathbb{R}^n)$ so that $f_k \to f$ in $L^1(\mathbb{R}^n)$.
Because Λf is sub-linear,

$$0 \le \Lambda f \le \Lambda(f - f_k) + \Lambda f_k = \Lambda(f - f_k)$$

Now we see that $\Lambda f \le 2Mf$ (because $f \le Mf$). Fix $\varepsilon > 0$. Then

$$\begin{aligned} |\{\Lambda f > \varepsilon\}| &\le |\{\Lambda(f - f_k) > \varepsilon\}| \\ &\le \left|\left\{ M(f - f_k) > \frac{\varepsilon}{2} \right\}\right| \\ &\le \frac{2C_n}{\varepsilon} \|f - f_k\|_1 \end{aligned}$$

where we use the weak estimate in the last line. The quantity $\|f - f_k\|_1$ vanishes as $k \to \infty$, so we conclude that $|\{\Lambda f > \varepsilon\}| = 0$ for all $\varepsilon > 0$, i.e. $\Lambda f = 0$ a.e.

□

Now that we've established that the limit exists a.e., it remains to find its value. Consider again the sequence $\{f_k\}$ as in proof of claim above. Consider then

$$\begin{aligned} \left|\lim_{r\to 0} \fint_{Q_r(x)} f(y)\,dy - f(x)\right| &= \left|\lim_{r\to 0} \fint_{Q_r(x)} (f(y) - f_k(y))\,dy + (f_k(x) - f(x))\right| \\ &\le M(f - f_k)(x) + |f_k(x) - f(x)| \end{aligned}$$

where we uses the fact that for the f_k's, the averaging yields the value of the function itself at x. Thus,

$$\left|\left\{ \left|\lim_{r\to 0} \fint_{Q_r(x)} f(y)\,dy - f(x)\right| > \varepsilon \right\}\right| \le |\{M(f - f_k) > \varepsilon/2\}| + |\{|f_k - f| > \varepsilon/2\}|$$

If we now apply the weak estimate and Markov inequality to respective term above, we get

$$\left|\left\{\left|\lim_{r\to 0} \fint_{Q_r(x)} f(y)\,dy - f(x)\right| > \varepsilon\right\}\right| \leq \frac{2}{\varepsilon}(C_n+1)\left\|f_k - f\right\|_1$$

which vanishes as $k \to 0$, hence the limit is indeed $f(x)$ a.e.

□

It is worth noting here two important ingredients that will be used repeatedly: (i) convergence in some dense class and (ii) bound of the maximal operator.
In this proof we used Theorem 1.2 twice eventhough we haven't given proof to the fact. The remaining discussion in this section will focus on proving the Hardy-Littlewood Maximal theorem. A standard proof of the theorem uses the following result:

Lemma 1.4 (Vitali Covering). *Let $E \subseteq \mathbb{R}^n$ and $E \subseteq \bigcup_\alpha Q_\alpha$ where the Q_α's are cubes with uniformly bounded diameters. Then there exists a countable sub-collection $\{Q_{\alpha_j}\}$ so that the Q_{α_j}'s are pairwise disjoint and there exists an uniform constant C_n so that*

$$|E| \leq C_n \sum_j |Q_{\alpha_j}|$$

Proof of Lemma 1.4. The proof will be constructive: assume such collection $\{Q_\alpha\}$. We may choose Q_1 from this collection so that

$$\operatorname{diam} Q_1 \geq \frac{1}{2}\sup_\alpha \operatorname{diam} Q_\alpha$$

because the diameter of the cubes are uniformly bounded. Remove then from the collection Q_1 and all the cubes meeting it (e.g. if any side is touching) to get a sub-collection $\{Q_\beta\}$. We can then again find Q_2 in this sub-collection so that

$$\operatorname{diam} Q_2 \geq \frac{1}{2}\sup_\beta \operatorname{diam} Q_\beta$$

Repeating this procedure will give us a (possibly finite) sequence of disjoint cubes $\{Q_k\}$. There are then two cases to consider. If $\sum_k |Q_k| = \infty$, then we're done (because we just say the $|E|$ is at most infinite, always a true statement).
Suppose now that $\sum_k |Q_k|$ is finite. Observe that for any k, because $\operatorname{diam} Q_k$ is at least twice of the remaining Q_α's in the collection, $Q_\alpha \subseteq 5Q_k$. Then

$$E \subseteq \bigcup_\alpha Q_\alpha \subseteq \bigcup_k 5Q_k \qquad \Rightarrow \qquad |E| \leq \left|\bigcup_k 5Q_k\right| \leq \sum_k |5Q_k| = 5^n \sum_k |Q_k| \tag{1}$$

Alternatively we may get that $\{Q_k\}$ is an infinite sequence, so $\operatorname{diam} Q_k \to 0$ as $k \to \infty$. It then remains, as in the finite case, that every Q_α is contained in some Q_k. There are two cases: trivially,

Q_α might be one of the Q_k's. Otherwise, because $\operatorname{diam} Q_k \to 0$, given Q_α we may find the smallest index k_0 so that

$$\operatorname{diam} Q_{k_0+1} \leq \frac{1}{2} \operatorname{diam} Q_\alpha$$

This implies Q_α meets with at least one of $Q_1, \cdots, Q_{k_0}$. Call the largest of these k_0 cubes to be Q_{k_1}. By same argument as above, $Q_\alpha \subseteq 5Q_{k_1}$. The estimate is then given by Equation (1).

□

Now that we've collected the necessary tools, we can give a short proof of the Hardy-Littlewood Maximal inequality.

Proof of Theorem 1.2. Fix $\lambda > 0$ and set $E_\lambda = \{Mf > \lambda\}$. By definition of the maximal operator, given $x \in E_\lambda$ there exists a cube Q_x containing x so that $\fint_{Q_x} |f| > \lambda$.

Moreover,

$$\lambda |Q_x| \leq \int_{Q_x} |f| \leq \|f\|_1 \tag{2}$$

so all such Q_x has uniformly bounded diameter. If we consider the collection of such Q_x, easily

$$E_\lambda \subseteq \bigcup_{x \in E_\lambda} Q_x$$

This sets us up precisely for application of the Covering lemma: we can find a sequence of disjoint cubes $\{Q_{x_k}\}$ so that

$$|E_\lambda| \leq 5^n \sum_k |Q_{x_k}|$$

Combined with estimate 2 and using the fact the Q_{x_k}'s are disjoint,

$$|E_\lambda| \leq \frac{5^n}{\lambda} \sum_k \int_{Q_{x_k}} |f| = \frac{5^n}{\lambda} \int_{\bigcup_k Q_{x_k}} |f| \leq \frac{5^n}{\lambda} \|f\|_1$$

□

As a generalization, we may state the following corollary

Corollary 1.5. *For all $1 < p \leq \infty$, $\|Mf\|_p \leq C \|f\|_p$.*

This corollary is a special case of a more general theorem, which we will proof (in one of its more restricted version):

Theorem 1.6 (Marcinkiewicz Interpolation)**.** *Suppose T is a sub-linear operator and of weak type $(1,1)$ and (r,r) (for $r > 1$), i.e.*

$$|\{|Tf| > \lambda\}| \leq \frac{C_1}{\lambda} \|f\|_1 \qquad \text{and} \qquad |\{|Tf| > \lambda\}| \leq \left(\frac{C_r}{r} \|f\|_r\right)^r \qquad \text{for } 1 < r < \infty$$

$$\|Tf\|_\infty \leq C_\infty \|f\|_\infty \qquad \text{for } r = \infty$$

Then $T : L^p \to L^p$ for all $1 < p < r$, i.e. there is a C_p so that $\|Tf\|_p \leq C_p \|f\|_p$ for any $f \in L^p$.

Note. The $r = \infty$ bound is given by the fact that the 'average' of f is at most its highest value $\|f\|_\infty$, while the lower bound $r = 1$ is given by Theorem 1.2. These two facts, with the theorem, will give the corollary.

Proof. Suppose first that $1 < r < \infty$ and let $1 < p < r$. Take any $f \in L^p$ and fix $\lambda > 0$. We can write $f = f_1 + f_2$, where

$$f_1 = \begin{cases} f & \text{if } |f| > \lambda \\ 0 & \text{o.w.} \end{cases}$$

and f_2 is the complement. Observe that not only $f_2 \in L^\infty$, but because $f \in L^p$,

$$\int |f_2|^r \leq \int |f_2|^p \cdot \lambda^{r-p} \leq \lambda^{r-p} \|f\|_p$$

and hence $f_2 \in L^r$. Similarly,

$$\int |f_1| \leq \int |f_1| \left(\frac{|f_1|}{\lambda}\right)^{p-1} \leq \frac{1}{\lambda^{p-1}} \|f\|_p^p$$

and thus $f_1 \in L^1$.
Since

$$\{|Tf| > \lambda\} \subseteq \left\{|Tf_1| > \frac{\lambda}{2}\right\} \cup \left\{|Tf_2| > \frac{\lambda}{2}\right\} \tag{3}$$

we can then use the distribution definition of integral to write

$$\begin{aligned} \int |Tf|^p &= p \int_0^\infty \lambda^{p-1} |\{|Tf| > \lambda\}| \, d\lambda \\ &\leq \underbrace{p \int_0^\infty \lambda^{p-1} \left|\left\{|Tf_1| > \frac{\lambda}{2}\right\}\right|}_{I} + \underbrace{p \int_0^\infty \lambda^{p-1} \left|\left\{|Tf_2| > \frac{\lambda}{2}\right\}\right|}_{II} \end{aligned}$$

We will then give bounds of I and II separately, so that the bound is proportional to $\|f\|_p$, as wanted. For the first we'll use the w(1, 1)-estimate and the observation that $f_1 \in L^1$:

$$\begin{aligned} p \int_0^\infty \lambda^{p-1} \left|\left\{|Tf_1| > \frac{\lambda}{2}\right\}\right| &\leq C_1 p \int_0^\infty \lambda^{p-1} \frac{2}{\lambda} \|f_1\|_1 \, d\lambda \\ &= 2C_1 p \int_0^\infty \lambda^{p-2} \int_{|f|>\lambda} |f(x)| \, dx \, d\lambda \\ &= 2C_1 p \int |f(x)| \int_0^{|f|} \lambda^{p-2} \, d\lambda \, dx \\ &= \frac{2C_1 p}{p-1} \|f\|_p^p \end{aligned} \tag{4}$$

Similarly for II we'll use the w(r, r) bound:

$$p \int_0^\infty \lambda^{p-1} \left|\left\{|Tf_2| > \frac{\lambda}{2}\right\}\right| \leq p \int_0^\infty \lambda^{p-1} \left(\frac{2C_r}{\lambda^r}\right)^r \int |f_2|^r \, dx \, d\lambda$$

$$= pC_r^r \int_0^\infty |f|^r \int_{|f|}^\infty \lambda^{p-1-r}\, d\lambda\, dx$$
$$= \frac{pC_r^r}{r-p} \|f\|_p^p \tag{5}$$

Combining (4) and (5) will give the wanted bound for some universal constant C that depends only on p and r.
It remains to prove the case $r = \infty$. We'll apply similar strategy: let

$$f_1 = \begin{cases} f & \text{if } |f| > \frac{\lambda}{2C_\infty} \\ 0 & \text{o.w.} \end{cases}$$

and let $f_2 = f - f_1$. Observe then

$$\|Tf_2\|_\infty \le C_\infty \|f_2\|_\infty = \frac{\lambda}{2}$$

(we choose the cut-off for f_1 precisely for this purpose). In particular, this implies $\left|\left\{ |Tf_2| > \frac{\lambda}{2} \right\}\right| = 0$. From (3) we deduce then

$$|\{|Tf| > \lambda\}| \le \left|\left\{ |Tf_1| > \frac{\lambda}{2} \right\}\right|$$

Again from the distribution definition of p-th moment and the observation above,

$$\begin{aligned}
\int |Tf|^p &= p \int_0^\infty \lambda^{p-1} |\{|Tf| > \lambda\}|\, d\lambda \\
&\le p \int_0^\infty \lambda^{p-1} \left|\left\{ |Tf_1| > \frac{\lambda}{2} \right\}\right| d\lambda \\
&\le p \int_0^\infty \lambda^{p-1} \frac{2C_1}{\lambda} \int |f_1|\, dx\, d\lambda
\end{aligned}$$

where we use the w$(1,1)$ bound above. From Fubini, we may continue

$$\begin{aligned}
\int |Tf|^p &\le 2C_1 p \int_0^\infty \lambda^{p-2} \int_{|f| > \lambda/2C_\infty} |f|\, dx\, d\lambda \\
&= 2C_1 p \int |f(x)| \int_0^{2C_1|f(x)|} \lambda^{p-2}\, d\lambda\, dx \\
&= 2^p C_1 C_\infty^{p-1} \frac{p}{p-1} \int |f|^p
\end{aligned}$$

In other words, we've found an universal constant depending only on p so that $\|Tf\|_p$ is bounded by $C\|f\|_p$. This completes the proof.

□

2 Approximate Identities

Consider a function $\varphi \in L^1(\mathbb{R}^n)$ so that $\int \varphi = 1$. For $\varepsilon > 0$, we'll set

$$\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)$$

so $\int \varphi_\varepsilon = 1$. Here we take the 'good kernel' φ_ε to be quite general. In its application, we'll often consider when φ is non-negative, smooth, and with compact support (i.e. of class C_c^∞).
Recall that given two (L^1) functions f and g, their convolution is defined as

$$(g * f)(x) = \int g(x-y)f(y)\,dy = \int f(x-y)g(y)\,dy$$

Then the operation $f \mapsto P_\varepsilon f =: \varphi_\varepsilon * f$ is called the approximation to the identity. We can imagine that φ_ε converges to the Dirac delta function as $\delta \to 0$ and thus $(\varphi_\varepsilon * f)(x)$ converges to $f(x)$ a.e. We will show some version of this result later, but for now we'll begin with a weaker version of it:

Proposition 2.1. *Let $1 \le p < \infty$ and denote $P_\varepsilon f := \varphi_\varepsilon * f$. Then $P_\varepsilon f \to f$ as $\varepsilon \to 0$ in L^p-norm for any $f \in L^p$.*

Proof. Let $f \in L^p$ and $\varepsilon > 0$. Recall that translation of L^p functions is a continuous operation, i.e.

$$\|f(\cdot + h) - f(\cdot)\|_p \to 0 \qquad h \to 0$$

Since we can write

$$P_\varepsilon f(x) - f(x) = \int \varphi_\varepsilon(x-y)[f(y) - f(x)]\,dy$$

Then

$$\begin{aligned}\left\|P_\varepsilon f - f\right\|_p &= \left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}[f(x-y) - f(x)]\,\varphi_\varepsilon(y)\,dy\right|^p dx\right)^{1/p}\\ &= \left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}[f(x-\varepsilon y) - f(x)]\,\varphi(y)\,dy\right|^p dx\right)^{1/p}\end{aligned}$$

Applying Fubini's and then Minkowski's Integral inequality,

$$\begin{aligned}\left\|P_\varepsilon f - f\right\|_p &\le \int_{\mathbb{R}^n}\varphi(y)\left(\int_{\mathbb{R}^n}|f(x-\varepsilon y) - f(x)|^p\,dx\right)^{1/p} dy\\ &= \int_{\mathbb{R}^n}\varphi(y)\,\|f(\cdot - \varepsilon y) - f(\cdot)\|_p\,dy\end{aligned}$$

As remarked, the norm in the integral vanishes, so in particular it is bounded. Then we may apply DCT (also because φ integrates to 1) and the continuity of translation to conclude that the upper bound vanishes as $\varepsilon \to 0$, i.e $P_\varepsilon f \to f$ in L^p-norm.

□

Note. If we take $\varphi(x) = \dfrac{1}{|B_1(0)|}\chi_{B_1(0)}$ or $\varphi(x) = \chi_{Q_1(0)}$ the proposition just reduces to the Lebesgue Differentiation theorem (Theorem 1.3)

The next theorem we're going to prove is significantly stronger, as it happens that the convergence also holds a.e. (instead of weakly a.e.). For that purpose, we'll use the following definition

Definition 2.2. Given $\varphi \in L^1$, the least decreasing majorant (LDM) of φ is defined as

$$\psi(x) = \sup_{|y| \geq |x|} |\varphi(y)|$$

Now we can state the main result of this section

Theorem 2.3. *Let $\varphi \in L^1$ and ψ be its LDM. Suppose that $\psi \in L^1$ with $\|\psi\|_1 = A$ for some $A > 0$. Then:*

(i) Pointwise bound: $\sup_{\varepsilon>0} |(f * \varphi_\varepsilon)(x)| \leq C \cdot A \cdot Mf(x)$

(ii) Pointwise convergence: $\lim_{\varepsilon \to 0}(f * \varphi_\varepsilon)(x) = f(x)$ *a.e. for* $f \in L^p$, $1 \leq p < \infty$.

Proof. We begin by showing (i). Fix $x \in \mathbb{R}^n$ and $\varepsilon > 0$. Then

$$\begin{aligned} |(f * \varphi_\varepsilon)(x)| &\leq \int_{\mathbb{R}^n} |\varphi_\varepsilon(x-y)|\,|f(y)|\,dy \\ &\leq \int_{\mathbb{R}^n} |\psi_\varepsilon(x-y)|\,|f(y)|\,dy \\ &= \varepsilon^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \psi\left(\frac{|x-y|}{\varepsilon}\right) |f(y)|\,dy \end{aligned}$$

We can split the space into annuli so that the last integral turns into sum, e.g.

$$|(f * \varphi_\varepsilon)(x)| \leq \sum_{j=-\infty}^{\infty} \underbrace{\varepsilon^{-n} \int_{2^j\varepsilon < |x-y| \leq 2^{j+1}\varepsilon} \psi\left(\frac{|x-y|}{\varepsilon}\right) |f(y)|\,dy}_{I_j}$$

Observe here that since ψ is decreasing, on each annulus we have $\psi\left(\dfrac{|x-y|}{\varepsilon}\right) \leq \psi(2^j)$. Then we can estimate I_j as

$$\begin{aligned} I_j &\leq \varepsilon^{-n} \int_{2^j\varepsilon < |x-y| \leq 2^{j+1}\varepsilon} \psi(2^j)\,|f(y)|\,dy \\ &\leq (2^j\varepsilon)^{-n} \int_{|x-y| \leq 2^{j+1}\varepsilon} |f(y)|\,dy \cdot 2^{jn}\psi(2^j) \end{aligned}$$

But observe that $|B(x, 2^j\varepsilon)| = C(n)(2^j\varepsilon)^n$, so we recognize the integral in front as the average of f. This average is always dominated by Mf, so

$$I_j \leq C(n) \cdot Mf(x) \cdot 2^{jn}\psi(2^j)$$

and thus

$$|(f * \varphi_\varepsilon)(x)| \leq C(n) \cdot Mf(x) \sum_{j=-\infty}^{\infty} 2^{jn}\psi(2^j)$$

The first two terms are already in the wanted form. It then remains that the last infinite sum can be bounded by A. To see this, note that because ψ is radial and decreasing, $\psi(2^j) \leq \psi(y)$ whenever $2^{j-1} \leq |y| < 2^j$. Moreover, since the size of the annulus is $C2^{jn}$ for some appropriate constant, we can bound

$$2^{jn}\,\psi(2^j) = C\int_{2^{j-1}\leq|y|<2^j} \psi(2^j)\,dy \leq C\int_{2^{j-1}\leq|y|<2^j} \psi(y)\,dy$$

Summing for all j gives us

$$\sum_{j=-\infty}^{\infty} 2^{jn}\psi(2^j) \leq C\sum_{j=-\infty}^{\infty}\int_{2^{j-1}\leq|y|<2^j} \psi(y)\,dy = C\int_{\mathbb{R}^n}\psi(y)\,dy \leq C \cdot A$$

which completes the proof for (i).
For (ii), the proof is the same as in proof of Theorem 1.3 except that the bound of Theorem 1.2 is replaced with (i)

□

Some examples of the LDM and how it's applied:

Example 2.4. The Poisson kernel in $\mathbb{R}^n \times \mathbb{R}_+$ is given by

$$p_t(x) = \frac{C_n}{(t^2 + |x|^2)^{(n+1)/2}} \equiv t^{-n} \cdot \varphi\left(\frac{x}{t}\right)$$

where

$$\varphi(x) = \frac{C_n}{(1 + |x|^2)^{(n+1)/2}}$$

and C_n is chosen so that $\|\varphi\|_1 = 1$. Note that φ is radial, decreasing, and integrable. Let $u(x,t) = (p_t * f)(x)$, then $u(x,t)$ solves the L^p-Dirichlet problem for Laplace equation in $\mathbb{R}^n \times \mathbb{R}_+$, i.e.

$$(\mathrm{D})\begin{cases} \Delta u = 0 & \text{in } \mathbb{R}^n \times \mathbb{R}_+ \\ \lim_{t\to 0} u(x,t) = f(x) \in L^p & \text{a.e. in } \mathbb{R}^n \\ \sup_{t>0} |u(x,t)| \leq C \cdot Mf & \text{in } L^p \text{ for } p > 1 \end{cases}$$

Example 2.5. The Gaussian or heat kernel is given by

$$w_t(x) = C_n t^{-n/2} \exp\left(-\frac{|x|^2}{4t}\right) \cong \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)$$

where

$$\varphi(x) = C_n \exp\left(-\frac{|x|^2}{4}\right)$$

Again, w_t is radial, decreasing, and integrable with the constant C_n is chosen so that $\|w_t\|_1 = 1$. Then, $u(x,t) = (w_t * f)(x)$ solves the initial value problem (IVP) for the heat equation, i.e.

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta_x u & \text{in } \mathbb{R}^n \times \mathbb{R}_+ \\ \lim_{t\to 0^+} u(x,t) = f(x) & \text{a.e. in } \mathbb{R}^n \\ \sup_{t>0} |u(x,t)| \le C \cdot Mf & \end{cases}$$

3 Fourier Transform

Definition 3.1. The Fourier transform of f is defined as

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi \cdot x}\, dx$$

Analogously, the inverse Fourier transform of g is defined as

$$\check{g}(x) = \int_{\mathbb{R}^n} g(\xi) e^{2\pi i x \cdot \xi}\, d\xi$$

We will start with some easy proposition:

Proposition 3.2. *Let $f \in L^1$. Then*

(i) $\left\|\widehat{f}\right\|_\infty \le \|f\|_1$

(ii) $\widehat{f}$ *is uniformly continuous in* $\mathbb{R}^n$.

Proof. The first one is easy: given $\xi \in \mathbb{R}^n$,

$$|\widehat{f}(\xi)| \le \int_{\mathbb{R}^n} |f(x)|\, |e^{-2\pi i \xi \cdot x}|\, dx = \int_{\mathbb{R}^n} |f(x)|\, dx = \|f\|_1$$

because $|e^{-2\pi i \xi \cdot x}| = 1$ for all x, ξ. The uniform bound follows immediately.
For (ii), let $\xi_1, \xi_2 \in \mathbb{R}^n$. Then

$$\begin{aligned} |\widehat{f}(\xi_1) - \widehat{f}(\xi_2)| &= \left| \int_{\mathbb{R}^n} f(x)(e^{-2\pi i \xi_1 \cdot x} - e^{-2\pi i \xi_2 \cdot x})\, dx \right| \\ &\le \int_{\mathbb{R}^n} |f(x)|\, \left|\exp[-2\pi i(\xi_1 - \xi_2)\cdot x] - 1\right|\, dx \end{aligned}$$

which gives (1) $|\widehat{f}(\xi_1) - \widehat{f}(\xi_2)| \le 2\|f\|_1$ and (2) from (1) and by DCT, $|\widehat{f}(\xi_1) - \widehat{f}(\xi_2)|$ tends to 0 as $|\xi_1 - \xi_2| \to 0$. This shows $\widehat{f}$ is uniformly continuous.

□

Note. In this section, everything that holds for $\widehat{f}$ almost always also holds for $\check{f}$.

The following theorem guarantees that the Fourier transform is well-defined.

Theorem 3.3 (Riemann-Lebesgue)**.** *If $f \in L^1$, then $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$.*

Proof. Recall that finite linear combination of characteristic functions of rectangles are dense in $L^1(\mathbb{R}^n)$. Using density argument, we can reduce the consideration to when $f = \chi_R$ with $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Indeed,

$$\begin{aligned}\widehat{\chi_R}(\xi) &= \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} e^{-2\pi i \xi_1 x_1} \cdots e^{-2\pi i \xi_n x_n}\, dx_1 \cdots dx_n \\ &= \prod_{k=1}^{n} \int_{a_k}^{b_k} e^{-2\pi i \xi_k x_k}\, dx_k\end{aligned}$$

We can then reduce the consideration further to the 1-D case, because the result follows if at least one of the factors in the last product vanishes. Then,

$$\begin{aligned}\left|\int_a^b e^{-2\pi i \xi x}\, dx\right| &= \left|\frac{1}{-2\pi i \xi} e^{-2\pi i \xi x}\Big|_a^b\right| \\ &\le \frac{2}{2\pi|\xi|} \overset{|\xi|\to\infty}{\to} 0\end{aligned}$$

□

Note. This theorem allows us to give another proof of the fact $\widehat{f}$ is uniformly continuous: since $\widehat{f}$ is continuous (easy to check) and vanishes at ∞, it's a simple fact that $\widehat{f}$ is then continuous.

The following proposition show that the convolution is well-defined if f and g are sufficiently integrable:

Proposition 3.4. *If $1 \le p < \infty$,*

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

Proof.

$$\begin{aligned}\|f * g\|_p &= \left(\int_{\mathbb{R}^n} \left|\int_{\mathbb{R}^n} f(x-y) g(y)\, dy\right|^p dx\right)^{1/p} \\ &\le \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\, dy\right)^p dx\right)^{1/p} \\ &\le \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^p\, dy\right)^{1/p} dx \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)|^p\, dx\right)^{1/p} |g(y)|\, dy \\ &= \|f\|_p \, \|g\|_1\end{aligned}$$

where we used Fubini and Minkowski's integral inequality. □

A couple more basic propositions:

Proposition 3.5. *For $f, g \in L^1$, $\widehat{f * g} = \widehat{f}\widehat{g}$*

Proof. The proof is by writing out the definition:

$$\begin{aligned} (\widehat{f * g})(\xi) &= \int_{\mathbb{R}^n} e^{-2\pi i \xi \cdot x} \int_{\mathbb{R}^n} f(x-y)\, g(y)\, dy\, dx \\ &= \int_{\mathbb{R}^n} g(y) e^{-2\pi i \xi \cdot y} \int_{\mathbb{R}^n} f(x-y) e^{-2\pi i \xi \cdot (x-y)}\, dx\, dy \end{aligned}$$

For the latter integral in the last line, use change of variable $x \mapsto x + y$. In this case,

$$(\widehat{f * g})(\xi) = \int_{\mathbb{R}^n} g(y) e^{-2\pi i \xi \cdot y} \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \xi \cdot x}\, dx = \widehat{g}\, \widehat{f}$$

□

Proposition 3.6 (Invariance properties)**.** *Let $f \in L^1$.*

(i) $(f(\cdot - h))^\wedge(\xi) = e^{-2\pi i \xi \cdot h} \widehat{f}(\xi)$

(ii) $(e^{2\pi i x \cdot h} f(x))^\wedge(\xi) = \widehat{f}(\xi - h)$

(iii) $\left(t^{-n} f\left(\frac{x}{t}\right)\right)^\wedge(\xi) = \widehat{f}(t\xi)$ *for* $t > 0$

(iv) Let φ be an orthogonal transformation of $\mathbb{R}^n$, i.e. $\rho^* = \rho^{-1}$. In particular, if $|\det \rho| = 1$, ρ is a rotation. Then $\widehat{f(\rho x)}(\xi) = \widehat{f}(\rho\xi)$.

(v) If f is radial, so is $\widehat{f}$.

Proof. For (i),

$$\begin{aligned} (f(\cdot - h))^\wedge(\xi) &= \int_{\mathbb{R}^n} f(x-h) e^{-2\pi i \xi \cdot x}\, dx \\ &= e^{-2\pi i \xi \cdot h} \int_{\mathbb{R}^n} f(x-h) e^{-2\pi i \xi \cdot (x-h)}\, dx \\ &= e^{-2\pi i \xi \cdot h} \widehat{f}(\xi) \end{aligned}$$

For (ii),

$$\begin{aligned} (e^{2\pi x \cdot h} f(x))^\wedge(\xi) &= \int_{\mathbb{R}^n} e^{2\pi i x \cdot h} f(x) e^{-2\pi i x \cdot \xi}\, dx \\ &= \int_{\mathbb{R}^n} f(x) e^{2\pi i x \cdot (\xi - h)}\, dx \\ &= \widehat{f}(\xi - h) \end{aligned}$$

For (iii), given $t > 0$,

$$\left(t^{-n} f\left(\frac{x}{t}\right)\right)^{\wedge}(\xi) \;=\; \int_{\mathbb{R}^n} t^{-n} f\left(\frac{x}{t}\right) e^{-2\pi i \xi\cdot x}\, dx$$

Change the variable $x \mapsto tx$. Then

$$\begin{aligned}\left(t^{-n} f\left(\frac{x}{t}\right)\right)^{\wedge}(\xi) &= \int_{\mathbb{R}^n} f(x) e^{-2\pi i(t\xi)\cdot x}\, dx \\ &= \widehat{f}(t\xi)\end{aligned}$$

For (iv), we'll again apply change of variable $x \mapsto \rho^{-1}x$. Then, if $|\det \rho| = 1$,

$$\begin{aligned}\widehat{f(\rho x)}(\xi) &= \int_{\mathbb{R}^n} f(\rho x) e^{-2\pi i \xi\cdot x}\, dx \\ &= \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi\cdot \rho^{-1} x} |\rho|\, dx \\ &= \int_{\mathbb{R}^n} f(x) e^{-2\pi i(\rho\xi)\cdot x}\, dx \\ &= \widehat{f}(\rho\xi)\end{aligned}$$

Last for (v), suppose ρ is a rotation. Because f is radial, $f(\rho x) = f(x)$ for all x. Moreover,

$$\widehat{f}(\rho\xi) = (f(\rho^{-1}x))^{\wedge}(\xi) = (f(x))^{\wedge}(\xi) = \widehat{f}(\xi)$$

and thus $\widehat{f}$ is also radial.

□

Theorem 3.7. *Suppose $f \in L^1$. Then*

(i) $\dfrac{\partial \widehat{f}}{\partial \xi_k}(\xi) = (-2\pi x_k f(x))^{\wedge}(\xi)$ *if* $x_k f(x) \in L^1$.

(ii) More generally, if P is a polynomial of degree $d \geq 1$, then $P(D)\widehat{f}(\xi) = (P(-2\pi i x) f(x))^{\wedge}(\xi)$.

(iii) $(P(D)f)^{\wedge}(\xi) = P(2\pi i\xi)\widehat{f}(\xi)$ *provided that for all multi-indices α with $|\alpha| \leq d$, $D^{\alpha} f \in L^1$ and also $f, D^{\alpha} f \to 0$ at infinity.*

Proof. For (i) is from definition:

$$\begin{aligned}\frac{\partial \widehat{f}}{\partial \xi_k}(\xi) &= \frac{\partial}{\partial \xi_k}\int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi\cdot x}\, dx \\ &= \int_{\mathbb{R}^n} (-2\pi i x_k) f(x)\, e^{-2\pi i\xi\cdot x}\, dx \\ &= (-2\pi i x_k\, f(x))^{\wedge}(\xi)\end{aligned}$$

Part (ii) follows from induction, so for any $m \in \mathbb{N}$,

$$\frac{\partial^m \widehat{f}}{\partial x_k^m}(\xi) = (-2\pi i x_k)^m \widehat{f}(\xi)$$

and thus

$$D^m \widehat{f}(\xi) = (-2\pi i x)^k \widehat{f}(\xi)$$

This expression is linear (and also derivative operator is linear), so the result extends to polynomial. Last, the proof of (iii) also relies on induction on degree of P. We'll just do the base case, as the recursion step is similar to (ii). Observe that

$$\left(\frac{\partial}{\partial \xi_k} f(x)\right)^{\wedge}(\xi) = \int e^{-2\pi i \xi \cdot x} \frac{\partial f}{\partial x_k}(x)\, dx$$

Let η_R be the smooth cut-off function of a ball of radius $2R$, i.e.

$$0 \leq \eta_R \leq 1, \qquad \eta_R \cong 1 \text{ on } B_R(0), \qquad \nabla \eta_R(x) < \frac{1}{R} \text{ for } R \leq |x| \leq 2R$$

Thus with integration by parts, since the function vanish at infinity (the boundary value),

$$\begin{aligned}\left(\frac{\partial}{\partial \xi_k} f(x)\right)^{\wedge}(\xi) &= \lim_{R\to\infty} -\int_{\mathbb{R}^n} \frac{\partial}{\partial x_k}(e^{-2\pi i \xi \cdot x} \eta_R(x))\, f(x)\, dx \\ &= \lim_{R\to\infty} \left[(2\pi i \xi_k) \int_{\mathbb{R}^n} e^{-2\pi i \xi \cdot x} f(x)\, \eta_R(x)\, dx + O\left(\frac{1}{R}\right) \int_{R \leq |x| \leq 2R} |f(x)|\, dx\right]\end{aligned}$$

The second integral in last line will vanish as $R \to \infty$, provided f decays at fast enough rate. The first integral converges, by DCT, to $2\pi i \xi_k \widehat{f}(\xi)$, so we conclude that

$$\nabla_\xi \widehat{f}(\xi) = (2\pi i \xi) \widehat{f}(\xi)$$

□

After we're done with the basic properties of Fourier transform, the latter part of this section will concern itself with proving that the inverse Fourier transform is basically the inverse of the Fourier transform. For that purpose, we'll introduce a couple tools that should be reminiscent of the previous section.

Definition 3.8. Let $\varepsilon > 0$. The ε-Gauss mean of g is defined as

$$G_\varepsilon(g) = \int_{\mathbb{R}^n} g(\xi)\, \exp(-4\pi^2 \xi^2 |\xi|^2)\, d\xi$$

If, for some $l \in \mathbb{R}$, $\lim_{\varepsilon\to 0} G_\varepsilon(g) = l$, we'll call g to be Gauss summable.

Note. If $g \in L^1$, then DCT implies that $G_\varepsilon(g) \to \int g$.

Now some basic properties of the Gauss means:

Proposition 3.9.

$$\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} e^{-4\pi^2|\xi|^2}\, d\xi = (4\pi)^{-n/2} e^{-|x|^2/4}$$

Proof. We may split the exponent on integrand on left expression by coordinate (easy to do because it's linear). By Fubini, we can then write iterated integral as

$$\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} e^{-4\pi^2|\xi|^2}\, d\xi = \prod_{j=1}^{n} \int_{-\infty}^{\infty} \exp(-4\pi^2\xi_j^2 - 2\pi i x_j \xi_j)\, d\xi_j \tag{6}$$

Because we have product of the same integral, it suffices to compute one of them. By completing the square, we may write

$$\int_{-\infty}^{\infty} \exp(-4\pi^2\xi^2 - 2\pi i x\xi)\, d\xi = e^{-x^2/4} \int_{\mathbb{R}} \exp[-(2\pi\xi + ix/2)^2]\, d\xi$$

If we change the variable $2\pi\xi + ix/2 \mapsto \xi$, the expression above turns to

$$\int_{-\infty}^{\infty} \exp(-4\pi^2\xi^2 - 2\pi i x\xi)\, d\xi = e^{-x^2/4} \cdot \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\xi^2}\, d\xi = \frac{e^{-x^2/4}}{\sqrt{4\pi}}$$

Applying the last computation to (6) gives precisely the wanted expression.

□

There are two easy but practical consequences of this last proposition.

Corollary 3.10. *If $a > 0$, then*

$$\int_{\mathbb{R}^n} e^{-\pi a|\xi|^2} e^{-2\pi i \xi\cdot x}\, d\xi = a^{-n/2} e^{-\pi|x|^2/a}$$

Proof. Apply change of variable $\xi \mapsto \sqrt{\frac{4\pi}{a}}\xi$. □

Corollary 3.11.

$$(e^{-\pi|x|^2})^\wedge(\xi) = e^{-\pi|\xi|^2}$$

Note. The last result says that the transform of Gaussian is also Gaussian.

The next proposition is a weaker version of a more general result, but for now the following is sufficient (and we have enough tools to prove it).

Proposition 3.12. *Given $f, g \in L^1$, $\int_{\mathbb{R}^n} \widehat{f} g = \int_{\mathbb{R}^n} f \widehat{g}$*

Proof. The proof is by unfolding the definition:

$$\begin{aligned}\int_{\mathbb{R}^n} \hat{f} g &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x)\, e^{-2\pi i x\cdot\xi}\, dx \right) g(\xi)\, d\xi \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} g(\xi)\, e^{-2\pi i x\cdot\xi}\, d\xi \right) f(x)\, dx \\ &= \int_{\mathbb{R}^n} \hat{g} f\end{aligned}$$

where we used Fubini to change the order of integration (OK because we have L^1 functions). □

We can generalize the Gauss mean with the following notion

Definition 3.13. The Φ means of g are integrals of the form

$$\Phi_\varepsilon(g) = \int_{\mathbb{R}^n} g(y)\Phi(\varepsilon y)\, dy$$

If Φ is continuous, bounded, and $\Phi(0) = 1$, then $\Phi_\varepsilon(g) \to \int g$. In particular, we may take Φ to be a Gaussian function.

While we noted above that Φ mean converges to the integral of the function itself for nice enough Φ, the following theorem gives a stronger result: the same holds for the inverse Fourier transform.

Theorem 3.14. *Let $f, \varphi, \Phi \in L^1$ so that $\hat{\varphi} = \Phi$, $\check{\Phi} = \varphi$, and $\int \varphi = 1$. Then the Φ means of $\hat{f}(\xi)e^{2\pi i\xi\cdot x}$ converges to f in L^1. Moreover, the Gauss means also converges to f a.e. (for each value of x), i.e.*

$$\lim_{\varepsilon\to 0} G_\varepsilon(\hat{f}(\xi)\, e^{2\pi i\xi\cdot x}) = [\hat{f}(\xi)\Phi(\varepsilon\xi)]^\vee(x) \text{ a.e.}$$

Note. Observe that

$$\Phi_\varepsilon(\hat{f}(\xi)e^{2\pi i\xi\cdot x}) = (\hat{f}(\xi)\Phi(\varepsilon\xi))^\vee(x)$$

and since $\hat{\varphi} = \Phi$,

$$\Phi_\varepsilon(\hat{f}(\xi)e^{2\pi i\xi\cdot x}) = \left(\hat{f} \left(\varepsilon^{-n}\varphi\left(\frac{\xi}{\varepsilon}\right)\right)^\wedge \right)^\vee (x) = ((f * \varphi_\varepsilon)^\wedge)^\vee(x)$$

The proof of Theorem 3.14 will use the following lemma:

Lemma 3.15. *Let $f, \Phi, \varphi \in L^1$ where $\varphi = \check{\Phi}$. Then*

$$\Phi_\varepsilon(\hat{f} e^{2\pi i\xi\cdot x}) = \int_{\mathbb{R}^n} \varphi(x-y) f(y)\, dy = (\varphi_\varepsilon * f)(x)$$

Note. The cumbersome hypothesis is apparently satisfied simply by letting φ or Φ to be Gaussian.

Proof. For fixed x, set $g(\xi) = e^{2\pi i \xi\cdot x}\Phi(\varepsilon\xi)$. Using Proposition 3.12 (possible because $\Phi \in L^1$) we may write $\Phi_\varepsilon(\widehat{f}(\xi)e^{2\pi i\xi\cdot x}) = \int \widehat{f} g = \int f\widehat{g}$, so it's sufficient to show that $\widehat{g}(y) = \varphi_\varepsilon(x-y)$.
From the definition of Fourier transform,

$$\begin{aligned}
\widehat{g}(y) &= \int_{\mathbb{R}^n} e^{2\pi i\xi\cdot x}\Phi(\varepsilon\xi)\, e^{-2\pi i\xi\cdot y}\, d\xi \\
&= \int_{\mathbb{R}^n} e^{2\pi i\xi\cdot(x-y)}\,\Phi(\varepsilon\xi)\, d\xi \\
&= \check{\Phi}(\varepsilon\xi)(x-y)
\end{aligned}$$

and we recognize the last expression as $\varphi_\varepsilon(x-y)$. This completes the proof.

□

Proof of Theorem 3.14. From the last lemma, the proof is simple: we may write

$$\Phi_\varepsilon(\widehat{f}e^{2\pi i\xi\cdot x}) = (\varphi_\varepsilon * f)(x)$$

Note that after appropriate normalization we may take φ to be an approximation to the identity. Then Proposition 2.1 gives the necessary convergence in L^1.
In the case Φ is the Gaussian, we note that the Gaussian function is its own LDM (see Definition 2.2), so Theorem 2.3 gives the pointwise convergence.

□

Now we're ready to state the proof the main result of this section: that the inverse Fourier transform is 'almost' the inverse of the Fourier transform.

Theorem 3.16 (Fourier Inversion in L^1)**.** *If both f and $\widehat{f}$ are in L^1, then $f(x) = (\widehat{f})^\vee(x)$ a.e. on $\mathbb{R}^n$.*

Proof. Let $f, \widehat{f} \in L^1$ and set $\Phi(\varepsilon\xi) = \exp(-4\pi^2\xi^2|\xi|^2)$. Then by the last theorem,

$$\int_{\mathbb{R}^n} \widehat{f}(\xi)e^{2\pi i\xi\cdot x}e^{-4\pi^2\varepsilon^2|\xi|^2}\, d\xi = G_\varepsilon(\widehat{f}(\xi)e^{2\pi i\xi\cdot x}) \overset{\varepsilon\to 0}{\to} f(x)$$

in L^1 an, in particular, a.e. On the other hand, we may write the integral above as the inverse Fourier transform $(\widehat{f}(\xi)\Phi(\varepsilon\xi))^\vee(x)$. Since $\widehat{f}$ is integrable and $\Phi(\varepsilon\xi)$ is uniformly bounded (by 1, no less), we may apply DCT to find that

$$(\widehat{f}(\xi)\Phi(\varepsilon\xi))^\vee(x) \overset{\varepsilon\to 0}{\to} \int_{\mathbb{R}^n} \widehat{f}(\xi)e^{2\pi i\xi\cdot x}\, d\xi = (\widehat{f})^\vee(x)$$

By uniqueness of limit, we conclude that $f(x) = (\widehat{f})^\vee(x)$ a.e.

□

4 Extending Fourier Transform to L^2

In the previous section we've introduced the basics of Fourier transform and gives several important results that come up when f and $\widehat{f}$ are merely integrable. In this section, we will try to extend these results using interpolation technique and density of certain class of function in $L^1 \cap L^2$. To motivate this section, we will start with the following theorem:

Theorem 4.1 (Plancherel for L^1)**.** *Suppose that $f, \widehat{f}, h, \widehat{h} \in L^1$, then*

$$\int_{\mathbb{R}^n} \widehat{f}\overline{\widehat{f}} = \int_{\mathbb{R}^n} f\overline{h} \tag{7}$$

where $\overline{h}$ is the complex conjugate of h. In particular, if $f = h$, $\|f\|_2 = \left\|\widehat{f}\right\|_2$.

Proof. Set $g = \overline{\widehat{f}}$. Then by Proposition 3.12

$$\int_{\mathbb{R}^n} \widehat{f}\overline{\widehat{f}} = \int_{\mathbb{R}^n} \widehat{f} g = \int_{\mathbb{R}^n} f\widehat{g}$$

It then suffices to show $\widehat{g} = \overline{h}$. This is just unpacking the definition:

$$\begin{aligned} \widehat{g}(x) &= \int_{\mathbb{R}^n} g(\xi)\, e^{-2\pi i \xi\cdot x}\, d\xi \\ &= \int_{\mathbb{R}^n} g(-\xi)\, e^{2\pi i \xi\cdot x}\, d\xi \\ &= (g(-\xi))^\vee(x) = (\overline{\widehat{h}}(-\xi))^\vee(x) \end{aligned}$$

However note that since conjugation commutes with complex conjugate,

$$\overline{\widehat{h}}(-\xi) = \int_{\mathbb{R}^n} \overline{h}(-x)\, \exp\overline{(-2\pi i(-\xi)\cdot x)}\, dx = \widehat{\overline{h}(x)}$$

and thus

$$\widehat{g}(x) = (\overline{\widehat{h}}(-\xi))^\vee(x) = (\widehat{\overline{h}})^\vee(x)$$

The Fourier inversion theorem says the last expression is equal to $\overline{h}$ a.e., so $\widehat{g} = \overline{h}$ and hence the proof is complete.

□

Notice however that this result is quite restricted and unnatural. For one, the theorem is useful only when $f \in L^2$. Moreover, it is natural to understand (7) as an inner product, where the natural domain is on L^2 functions (as is the case for Fourier transform). Thus we may extend the result to the space L^2. The key argument will be density in a special class, which we will now describe.

Definition 4.2. The Schwartz class $\mathcal{S}$ is a collection of smooth function which eventually derivative decreases faster than any polynomial power. To be precise

$$\mathcal{S} = \left\{ \varphi \in C^\infty \,\middle|\, \forall N \in \mathbb{N}, \forall \beta \text{ multi-index with } |\beta| \geq 0,\ \sup_{x \in \mathbb{R}^n} |x|^N |D^\beta \varphi(x)| \leq C_{N,\beta,\varphi} < \infty \right\}$$

There are several ways we can characterize this space. For one, we can see that if $\varphi \in \mathcal{S}$, $|\varphi(x)| \leq C_{N,\varphi}(1+|x|)^{-N}$ for all N, because φ decays faster than any polynomial power. In particular, this gives that $\varphi \in L^p$ for $1 \leq p < \infty$. Last, we'll note that $\mathcal{S}$ is dense in L^p: the C_c belongs to $\mathcal{S}$.
The following proposition ensures that Fourier transform doesn't change $\mathcal{S}$:

Proposition 4.3. *If $f \in \mathcal{S}$ then $\widehat{f} \in \mathcal{S}$. Hence if $f \in \mathcal{S}$, then f and $\widehat{f}$ belongs to $L^1 \cap L^2$.*

Proof. Fix $m \geq 0$ and a multi-index β so that $|\beta| \geq 0$. Using the differentiation rule for Fourier transform (Theorem 3.7), we may write

$$\xi^m D^\beta \widehat{f}(\xi) = C_{m,\beta} \xi^m (\widehat{x^\beta f(x)})(\xi) = C_{m,\beta} [D^m (x^\beta f(x))]^\wedge(\xi)$$

Since $f \in \mathcal{S}$,

$$|\xi^m D^\beta \widehat{f}(\xi)| \leq C_{m,\beta} \left\| [D^m (x^\beta f(x))]^\wedge \right\|_\infty < \infty$$

Since m and β are arbitrary, $\widehat{f}$ belongs to $\mathcal{S}$. $\square$

Using the last proposition, we will now describe the construction of $\widehat{f}$ for $f \in L^2$. Let $f \in L^2(\mathbb{R}^n)$. By density of $\mathcal{S}$, there is a sequence $\{f_k\} \subseteq \mathcal{S}$ so that $f_k \overset{2}{\to} f$.
Because this sequence is in $L^1 \cap L^2$, (i) Theorem 4.1 implies $\|f_k\|_2 = \|f_k\|_2$ for all k. Moreover, since this sequence is also Cauchy, $\left\| f_k - f_j \right\|_2 = \left\| \widehat{f}_k - \widehat{f}_j \right\|_2$ for all k, j. In particular, this last quantity vanishes as $k, j \to \infty$. The completeness of L^2 then gives an element $g \in L^2$ so that $\widehat{f}_k \overset{2}{\to} g$. Define then $\widehat{f}$ to be g. This gives a way to extend the Fourier transform to L^2 (and allows us to do the same for most of the previous results).
It remains to check that this definition is really well-defined. Suppose $\{\widetilde{f}_k\} \subseteq \mathcal{S}$ is another sequence that converges to f in L^2. By the same argument, $\{\widehat{\widetilde{f}}_k\}$ is Cauchy in L^2, so it converges to an element $\widetilde{g}$. We'll then show that $g = \widetilde{g}$.
To see this, consider the sequence $f_1, \widetilde{f}_1, f_2, \widetilde{f}_2, \cdots$. Because $\{f\}$ and $\{\widetilde{f}\}$ are Cauchy sequence converging to the same limit, this new sequence, call it $\{f_k^\sharp\}$, also converges to f. Moreover, by the same argument, the transform's sequence also converges, say to $g^\sharp$. Now note the following: $\{\widehat{f_k^\sharp}\}$ is convergent, so all its subsequences converges to the same limit. In particular, this means $g = g^\sharp = \widetilde{g}$ (from the odd and even limit, respectively). This proves the uniqueness of the limit, so $\widehat{f}$ is well-defined, regardless of construction.
Rewriting the symbols appropriately, we have just essentially given proof to the following statement:

Proposition 4.4. *Suppose $T : X \to Y$ is a linear or a non-negative sub-linear operator (i.e. $T(-f) = T(f)$) between Banach spaces and $Y \subseteq \mathbb{R}$. Assume also that T satisfies $\|Tf\|_Y \leq C\|f\|_X$ for all $f \in X' \subseteq X$ a dense subspace. Then T extends to a bounded operator from X to Y.*

Now we can state Plancherel's formula in its natural domain.

Theorem 4.5 (Plancherel)**.** *Let $f \in L^2$. Then $\widehat{f}$ exists and belongs to L^2. Moreover, $\|f\|_2 = \left\|\widehat{f}\right\|_2$.*

Proof. Recall that Proposition 4.3 gives that (i) the Schwartz space $\mathcal{S}$ is dense and (ii) given $f \in \mathcal{S}$, $\widehat{f} \in \mathcal{S}$. Then let $f \in L^2$ and $\{f_k\} \subseteq \mathcal{S}$ so that $f_k \overset{2}{\to} f$ (and consequently $\widehat{f}_k \overset{2}{\to} \widehat{f}$. Then, we can approximate as follows:

$$\begin{aligned}\left\|\widehat{f}\right\|_2 &\leq \left\|\widehat{f} - \widehat{f}_k\right\|_2 + \left\|\widehat{f}_k\right\|_2 \\ &= \left\|\widehat{f} - \widehat{f}_k\right\|_2 + \|f_k\|_2 \\ &\leq \left\|\widehat{f} - \widehat{f}_k\right\|_2 + \|f_k - f\|_2 + \|f\|_2\end{aligned}$$

where above we used $\left\|\widehat{f}_k\right\|_2 = \|f_k\|_2$ for all k because $f_k \in L^1$ and Theorem 4.1. Moreover, by the construction outlined previously, we know that f_k and its transform converges, in L^2 norm, to f and $\widehat{f}$, respectively. Letting $k \to \infty$, we then get $\left\|\widehat{f}\right\|_2 \leq \|f\|_2$. Arguing similarly will give the opposite inequality, so we obtain the wanted equality for $f \in L^2$.

□

Recall that in the previous section, we've observed that $\left\|\widehat{f}\right\|_2 \leq \|f\|_1$, i.e. the FT is a bounded operator from L^1 to L^∞. Similarly, the proof of Plancherel shows that the FT is also bounded operator from L^2 to L^2. Based on these end cases, the following theorem interpolates to the intemediate cases.

Theorem 4.6 (Hausdorff-Young)**.** *For all $1 \leq p \leq 2$, $\left\|\widehat{f}\right\|_{p'} \leq \|f\|_p$.*

This theorem is a special case of the following interpolation theorem.

Theorem 4.7 (Riesz-Thorin)**.** *Let T be a linear operator so that*

$$\begin{aligned}&\|Tf\|_{q_0} \leq k_0 \|f\|_{p_0} &\qquad& 1 \leq p_0, q_0 \leq \infty \\ &\|Tf\|_{q_1} \leq k_1 \|f\|_{p_1} && 1 \leq p_1, q_1 \leq \infty\end{aligned}$$

(p and q are not necessarily dual). Then

$$\|Tf\|_{q_\theta} \leq k_0^{1-\theta} k_1^\theta \|f\|_{p_\theta}$$

for all $0 \leq \theta \leq 1$ with

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \qquad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$$

The proof of the Hausdorff-Young theorem is then simple: in the setting of Riesz-Thorin, take $p_0 = 1$, $q_0 = \infty$, and $p_1 = q_1 = 2$. Note here that we can define the space L^p for $1 < p < 2$ as follows: a function f belongs to L^p iff it can be written as $f = f_1 + f_2$, where $f \in L^1$ and $f \in L^2$ (e.g. by writing $f_2 = f \cdot \chi_{\{|f| \leq N\}}$ for some large N). The function f is independent of this decomposition.

Before we prove the Riesz-Thorin theorem, we'll extend to L^2 some facts we've shown in previous function for L^1 functions:

Proposition 4.8. *(i) Let $f, g \in L^2$. Then* $\int \widehat{f} g = \int f \widehat{g}$.

(ii) Given $f \in L^1$ and $g \in L^p$ for $1 \le p \le 2$, $\widehat{f * g} = \widehat{f}\widehat{g}$ *a.e.*

Proof. Recall that we've shown the equality for L^1 functions in Proposition 3.12. Fix $f, g \in L^2$. Let $\{f_k\}, \{g_k\} \subseteq \mathcal{S}$ so that $f_k \overset{2}{\to} f$ and $g_k \overset{2}{\to} g$. Then the wanted statement holds already for f_k and g_k. We can then approximate the integral as follows:

$$\begin{aligned} \int \widehat{f} g &= \int (\widehat{f} - \widehat{f}_k) g + \int \widehat{f}_k g \\ &= \int (\widehat{f} - \widehat{f}_k) g + \int (g - g_k) \widehat{f}_k + \int \widehat{f}_k g_k \end{aligned}$$

Similary we can write

$$\int f \widehat{g} = \int (\widehat{g} - \widehat{g}_k) f + \int (f - f_k) \widehat{g}_k + \int f_k \widehat{g}_k$$

Substracting the two expression, we get

$$\begin{aligned} \left| \int \widehat{f} g - \int f \widehat{g} \right| \le & \left| \int (\widehat{f} - \widehat{f}_k) g \right| + \left| \int (g - g_k) \widehat{f}_k \right| + \left| \int (\widehat{g} - \widehat{g}_k) f \right| \\ & + \left| \int (f - f_k) \widehat{g}_k \right| + \left| \int \widehat{f}_k g_k - \int f_k \widehat{g}_k \right| \end{aligned}$$

Observe that by application of Cauchy-Schwarz, the first four terms in the right side will vanish in the limit. Moreover, since $f_k, g_k \in L^1$, Proposition 3.12 gives that the last term should vanish for all k. Therefore in the limit we should have

$$\left| \int \widehat{f} g - \int f \widehat{g} \right| = 0$$

which implies the wanted expression.
For (ii) we note that Hausdorff-Young and Proposition 3.4,

$$\left\| \widehat{f * g} \right\|_{p'} \le \| f * g \|_p \le \| f \|_1 \, \| g \|_p \qquad \left\| \widehat{g} \right\|_{p'} \le \| g \|_p$$

Let $\{g_k\} \subseteq \mathcal{S}$ so that $g_k \to g$ in L^p. By the observation above,

$$\left\| \widehat{g}_k - \widehat{g} \right\|_{p'} \le \left\| g_k - g \right\|_p \overset{k \to \infty}{\to} 0$$

and thus $g_k \to g$ also in $L^{p'}$. Similarly, the observation above also gives

$$\widehat{f * g_k} \overset{p'}{\to} \widehat{f * g}$$

From Proposition 3.5, since $g_k \in L^1$, we know $\widehat{f * g_k} = \widehat{f}\widehat{g_k}$. Then

$$\left\|\widehat{f * g} - \widehat{f}\widehat{g}\right\|_{p'} \leq \left\|\widehat{f * g} - \widehat{f * g_k}\right\|_{p'} + \left\|\widehat{f * g_k} - \widehat{f}\widehat{g_k}\right\|_{p'} + \left\|\widehat{f}\widehat{g_k} - \widehat{f}\widehat{g}\right\|_{p'}$$

The last observation says that the second term on the right side vanishes, while the two remarks right before that says that the first and last expression vanishes in the limit. Thus

$$\left\|\widehat{f * g} - \widehat{f}\widehat{g}\right\|_{p'} = 0 \qquad \Rightarrow \qquad \widehat{f * g} = \widehat{f}\widehat{g} \text{ a.e. on } \mathbb{R}^n$$

□

To finish this section it remains to prove the Riesz-Thorin theorem. The proof is complex analytic and will rely on this neat little complex analysis lemma.

Lemma 4.9 (Hadamard's Three Lines)**.** *Let f bounded in a closed vertical strip*

$$S = \{z = x + iy \in \mathbb{C} \mid 0 \leq x \leq 1\}$$

and analytic in interior of S but continuous up to the boundary. Suppose $|F(iy)| \leq M_0$ and $|F(1 + iy)| \leq M_1$. Then for all $z \in S$, $|F(x + iy)| \leq M_0^{1-x} M_1^x$.

Proof. Look at complex analysis books. □

Proof of Theorem 4.7. Long stuff □

Just like in the previous section, we'll close this section by presenting another version of Fourier Inversion theorem. Before that, we'll give two auxiliary results (that are important in its own right)

Lemma 4.10. *Let $T \in \mathcal{L}(X)$, where X is any Banach space, so that $\|x\| \leq c_0 \|Tx\|$ for all $x \in X$. Then T is injective with closed range.*

Proof. The coercive condition implies injectivity (as $Tx = 0$ can only implies $x = 0$). Now we'll show that $R(T)$ is closed: let $\{Tx_k\} = \{y_k\}$ be a Cauchy sequence in X. Because X is complete, there is $y \in X$ so that $y_k \to y$.
At the same time, the coercive condition gives

$$\left\|x_k - x_j\right\| \leq c_0 \left\|Tx_k - Tx_j\right\| = c_0 \left\|y_k - y_j\right\|$$

which vanishes as $j, k \to \infty$, so $\{x_k\}$ is also Cauchy. We can then find its limit $x \in X$. Therefore,

$$\left\|Tx_k - Tx\right\| \leq \|T\| \left\|x_k - x\right\|$$

and hence $Tx_k \to Tx$. By uniqueness of limit, $y = Tx$, i.e. the range of T is closed.

□

We'll use this lemma to prove the following key fact about Fourier transform.

Theorem 4.11. *The Fourier transform $\mathscr{F}$ is unitary on $L^2(\mathbb{R}^n)$, i.e. the operator is surjective and an isometry.*

Proof. Isometry is already given by Plancherel's formula (Theorem 4.5). The previous lemma gives that $\mathscr{F}$ is coercive (because of Plancherel's formula), so its range is closed. To get surjectivity, it then suffices to show that $R(\mathscr{F})$ is dense.
Suppose it is not dense. Then there is a non-zero $g \in L^2$ so that $g \perp R(\mathscr{F})$, i.e. for any $f \in L^2$, because $\widehat{f} \in R(\mathscr{F})$,

$$\int f\widehat{\overline{g}} = \int \widehat{f}\overline{g} = 0$$

But the first integral is zero for all functions iff $\widehat{\overline{g}} = 0$. In particular, if we take $f = \widehat{g}$, this means $\|\widehat{g}\|_2 = \|g\|_2 = 0$. This gives the wanted contradiction.

□

Finally we'll prove the important theorem of this section.

Theorem 4.12 (Fourier Inversion)**.** *The Fourier transform is invertible on L^2 and $(\widehat{f})^\vee = f$ a.e. for all $f \in L^2$.*

Proof. We've shown that Fourier transform is injective (in the lemma above) and surjective (the previous theorem), so we've shown $\mathscr{F}$ is invertible. It remains to show that it gives the right inverse. To show the formula, it suffices to show that it holds in a dense class, as we can then pass to the general case. Recall that an unitary operator's inverse is its adjoint (i.e. $\mathscr{F}^{-1} = \mathscr{F}^*$). Let $f, g \in \mathcal{S}$. Then by Fubini (OK because $f, g \in L^1$)

$$\begin{aligned}
\langle f, \widehat{g}\rangle &= \int f(x) \int \overline{g}(\xi)e^{2\pi i x\cdot\xi}\, d\xi\, dx \\
&= \int\int f(x)e^{2\pi i x\cdot\xi}\, d\xi\, f(x)\, dx \\
&= \langle \check{f}, g\rangle \equiv \langle \mathscr{F}^* f, g\rangle
\end{aligned}$$

and thus $\check{f} - \mathscr{F}^* f$ is orthogonal to g. Letting g roam over $\mathcal{S}$, we conclude that $\check{f} = \mathscr{F}^* f = \mathscr{F}^{-1} f$. This proves the inversion formula.

□

5 Singular Integral Operator

In the following section, we will talk about integral operators (mostly on L^2, but more generally on class of test functions C_c^∞) that contains singularity in its definition. Before we give the definition, we'll start with some major examples:

Example 5.1. The Hilbert transform of a function $f : \mathbb{R} \to \mathbb{R}$ is defined as

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

The limit is called the principal value of the integral.

Example 5.2. Let $1 \le j \le n$ for fixed n. The j-th Riesz transform of f in $\mathbb{R}^n$ is defined as

$$R_j f(x) = C_n \lim_{\varepsilon\to 0} \int_{|x-y|>\varepsilon} \frac{x_j - y_j}{|x-y|^{n+1}} f(y)\, dy$$

More generally, if we let $\Omega \in L^1(S^{n-1})$, we can extend this function to $\mathbb{R}^n \setminus \{0\}$ by setting $\Omega(x) = \Omega\left(\frac{x}{|x|}\right)$, so the function Ω is homogenous along rays. Assume also for normalization that $\int_{S^{n-1}} \Omega(\theta)\, d\theta = 1$. Then as generalization, we can consider the operator

$$Tf(x) = \text{p.v.} \int_{\mathbb{R}^n} \frac{\Omega(x-y)}{|x-y|^n} f(y)\, dy$$

To get back the j-th Riesz transform, set $\Omega(x) = \dfrac{x_j}{|x|}$.

Observe that all the examples above are of the convolution type. That is, formally, $Tf = K * f$ where $K(x) = \dfrac{\Omega(x)}{|x|^n}$. We know that FT plays well with convolution, so we should expect that, formally, $\widehat{Tf} = m\hat{f}$, where $m = \hat{K}$. In fact, this statement can be interpreted as

$$m(z) = \lim_{\varepsilon\to 0} \hat{K}_{\varepsilon,R}(z)$$

where $K_{\varepsilon,R}(x) = K(x)\chi_{\{\varepsilon<|x|<R\}}$. Later in the section, we will show that for the Hilbert transform H

$$m(z) = i\,\text{sgn}(z)$$

and for the Riesz transform R_j

$$m_j(z) = i\frac{z_j}{|z|}$$

Assuming the last for now, we can see that

$$\left(-\partial_j \partial_k f\right)^\wedge = 4\pi^2 z_j z_k \hat{f}(z) = -\frac{z_j}{|z|}\frac{z_k}{|z|}\left(-4\pi^2|z|^2\hat{f}(z)\right)$$

which we recognize as $(R_j R_k \Delta f)^\wedge$. This observation then proves the following proposition:

Proposition 5.3. *Consider the Poisson problem $-\Delta u = f \in L^2(\mathbb{R}^n)$. Then we have the $W_{2,2}$ estimate*

$$\left\| \frac{\partial^2 u}{\partial x_k \partial x_j} \right\|_2 \le \|f\|_2$$

Proof. By the discussion above,

$$\left\| \frac{\partial^2 u}{\partial x_k \partial x_j} \right\|_2 = \left\| R_j R_k \Delta u \right\|_2 = \left\| R_j R_k f \right\|_2 = \left\| \frac{z_j}{|z|} \frac{z_k}{|z|} \hat{f} \right\|_2 \leq \left\| \hat{f} \right\|_2 = \|f\|_2$$

□

After we're done with the motivation part, we'll try to build the major theorems needed. In general, SIO theory is split between L^2 theory and L^p theory for $1 < p < \infty$. Now we'll give general definition of SIO. The set-up is as follows: let $\mathcal{D} = C_c^\infty(\mathbb{R}^n)$ and $\mathcal{D}'$ be the dual of $\mathcal{D}$, i.e. the space of distributions.

Definition 5.4. A generalized SIO T is a mapping from $\mathcal{D}$ to $\mathcal{D}'$ and is associated with a Calderon-Zygmund kernel $K(x, y)$ in the sense that for $f, g \in \mathcal{D}$ with disjoint supports,

$$\langle Tf, g \rangle = \int_{\mathbb{R}^n} g(x) \int_{\mathbb{R}^n} K(x, y) f(y)\, dy\, dx$$

Note that $\langle \cdot, \cdot \rangle$ is not an inner product.

Definition 5.5. A Calderon-Zygmund kernel satisfies the following conditions:

(i) $|K(x, y)| \leq \dfrac{C}{|x - y|^n}$

(ii) $|K(x, y) - K(x, y + h)| + |K(x, y) - K(x + h, y)| \leq C \dfrac{|h|^\alpha}{|x - y|^{n+\alpha}}$

for some $\alpha \in (0, 2]$ and $|x - y| \geq 2|h|$.

Note. From the definition, given a SIO with C-Z kernel $K(x, y)$, we may write

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y)\, dy$$

for $x \notin \operatorname{supp} f$. This integral is absolutely convergent.

The first examples of this section has been of the convolution types. Here we'll give an example of a SIO that cannot be written as convolution.

Example 5.6. The first Calderon commutator on $\mathbb{R}$ is defined as

$$C_1 f(x) = \lim_{\varepsilon \to 0} \frac{1}{2\pi} \int_{|x-y|>\varepsilon} \frac{A(x) - A(y)}{(x - y)^2} f(y)\, dy$$

where $A : \mathbb{R} \to \mathbb{C}$ is Lipschitz. Here the associated kernel is $K(x, y) = \dfrac{1}{2\pi} \dfrac{A(x) - A(y)}{(x - y)^2}$.

Note. The examples of kernel so far has been anti-symmetric: $K(x, y) = -K(y, x)$.

In fact, from the observation above, we can say a little bit more of the SIO:

Proposition 5.7. *For an anti-symmetric kernel $K(x, y)$, the associated p.v. SIO always makes sense in terms of distribution, i.e. as mapping from $\mathcal{D}$ to $\mathcal{D}'$.*

Proof. Let $f, g \in C_c^\infty$. Define

$$T_\varepsilon f(x) = \int_{|x-y|>\varepsilon} K(x, y)\, f(y)\, dy$$

Then,

$$\begin{aligned}\langle T_\varepsilon f, g\rangle &= \iint_{|x-y|>\varepsilon} K(x, y)\, f(y)\, g(x)\, dy\, dx \\ &= -\iint_{|x-y|>\varepsilon} K(y, x) f(y)\, g(x)\, dy\, dx \\ &= -\iint_{|x-y|>\varepsilon} K(x, y)\, f(x)\, g(y)\, dx\, dy\end{aligned}$$

and hence

$$\langle T_\varepsilon f, g\rangle = \frac{1}{2}\iint_{|x-y|>\varepsilon} K(x, y)\, [f(y)g(x) - f(x)g(y)]\; dy\, dx \tag{8}$$

For the proof we'll need the following claim:

Claim.

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |K(x, y)||f(y)g(x) - f(x)g(y)|\, dy\, dx \le C\, R^{n+1}(\|f\|_\infty \|\nabla g\|_\infty + \|\nabla f\|_\infty \|g\|_\infty)$$

where f and g are supported inside $B_R(x_0)$ for some $R > 0$.

Proof of Claim. Notice that by Mean Value Theorem,

$$\begin{aligned}|f(x)g(y) - f(y)g(x)| &\le |f(y) - f(x)|\, |g(x)| + |f(x)|\, |g(y) - g(x)| \\ &\le |x - y| \left(\|\nabla f\|_\infty \|g\|_\infty + \|f\|_\infty \|\nabla g\|_\infty\right)\end{aligned}$$

Moreover, since the support of f and g are contained in $B_R(x_0)$, $|x - x_0|, |y - x_0| < R$ for all x and y in the ball. In particular, $|x - y| < 2R$. Thus from the estimate above and the C-Z kernel condition,

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |K(x, y)||f(y)g(x) - f(x)g(y)|\, dy\, dx &\le C\int_{|x-x_0|>R}\int_{|x-y|<2R} \frac{1}{|x-y|^{n-1}}\, dy\, dx \cdot \\ &\quad \left(\|\nabla f\|_\infty \|g\|_\infty + \|f\|_\infty \|\nabla g\|_\infty\right) \\ &\le C\left(\|\nabla f\|_\infty \|g\|_\infty + \|f\|_\infty \|\nabla g\|_\infty\right) \cdot R \int_{|x-x_0|<R} dx \\ &= C \cdot R^{n+1}\left(\|\nabla f\|_\infty \|g\|_\infty + \|f\|_\infty \|\nabla g\|_\infty\right)\end{aligned}$$

□

Since we know the expression $\langle T_\varepsilon f, g\rangle$ is bounded, we may apply DCT to (8) and take limit as $\varepsilon \to 0$ to properly define the operator.

□

In fact, the proposition above is an example of a more general phenomenon.

Definition 5.8. A SIO T satisfies the weak boundedness property (WBP) if for all $f, g \in C_c^\infty$ and for all $B = B_R(x_0)$, $x_0 \in \mathbb{R}^n$ and $0 < R < \infty$ such that supp f, supp $g \subseteq B_R(x_0)$,

$$|\langle Tf, g\rangle| \leq CR^n(\|f\|_\infty + R\|\nabla f\|_\infty)(\|g\|_\infty + R\|\nabla g\|_\infty)$$

Note that the proof of the proposition above also proves the following statement.

Corollary 5.9. *If T is a p.v. SIO associated with an anti-symmetric kernel, then T satisfies WBP.*

Note. T is L^2-bounded if for all $f, g \in C_c^\infty$, $|\langle Tf, g\rangle| \leq C\|f\|_2\|g\|_2$. Moreover in $B_R(x_0)$ (chosen large enough to contain the support of both f and g), the L^2-norm is controlled by L^infty-norm. By the pairing bound above, we can then extend by density to all of L^2. Hence WBP is a necessary condition for WBP.

The following theorem shows the usefulness of SIOs: they can be interpolated to the whole L^p for $1 < p < \infty$.

Theorem 5.10 (Calderon-Zygmund)**.** *Let T be an SIO associated to a standard C-Z kernel. Suppose that $T : L^2 \to L^2$. Then*

(i) T is of $w(1,1)$-type.

(ii) $T : L^p \to L^p$, i.e. $\|Tf\|_p \leq C_p\|f\|_p$ for $1 < p < \infty$ for some constant C_p.

Proof. We begin by showing that (ii) follows from (i): by hypothesis, there is a constant $C_2 > 0$ so that $\|Tf\|_2 \leq C_2\|f\|_2$. Observe that

$$|\{|Tf| > \lambda\}| = \left|\{|Tf|^2 > \lambda^2\}\right| \leq \frac{1}{\lambda^2}\int |Tf|^2 \leq \frac{C_2^2}{\lambda^2}\|f\|_2^2$$

In particular, the L^2 bound for T implies T is of w$(2,2)$-type. Assuming we've shown (i), the Marcinkiewicz Interpolation theorem (Theorem 1.6) gives that $T : L^p \to L^p$ for all $1 < p < 2$.
To extend this for $2 < p < \infty$, we begin by the following observation: since L^2 is self-dual, $T : L^2 \to L^2$ iff $T^* : L^2 \to L^2$. Moreover, the kernel associated with T^* is $K^*(x, y) = K(y, x)$, which is also a C-Z kernel (because the definition is symmetric). Thus by the same reasoning, $T^* : L^p \to L^p$ for $1 < p < 2$.
Now let $2 < p < \infty$. For any $f, g \in \mathcal{S}$,

$$\begin{aligned} |\langle Tf, g\rangle| &\leq |\langle f, T^*g\rangle| \\ &\leq \|f\|_p\|Tg\|_{p'} && \text{with } 1 < p' < 2 \\ &\leq C_{p'}\|f\|_p\|g\|_{p'} \end{aligned}$$

Taking supremum over all g with $\|g\|_{p'} = 1$, we conclude that $\|T\|_{p\to p} \leq C_{p'}$, i.e. $T : L^p \to L^p$ for $2 < p < \infty$. This covers all the cases for (ii).
Before we prove (i), we'll show this highly versatile (and surprising) lemma:

Lemma 5.11 (Calderon-Zygmund decomposition)**.** *Let $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$. Then there exists a family $\{Q_k\}_{k=1}^{\infty}$ of non-overlapping (i.e. at most touching boundaries) dyadic cubes and a decomposition of $f = b + g$ so that*

(i) $\|g\|_\infty \leq 2^n \lambda$ *and* $\|g\|_1 \leq \|f\|_1$.

(ii) $\displaystyle\sum_{k=1}^{\infty} |Q_k| \leq \frac{\|f\|_1}{\lambda}$

(iii) We can decompose b further, e.g. $b = \sum_{k=1}^{\infty} b_k$ *where* $\operatorname{supp} b_k \subseteq Q_k$ *and* $\int b_k = 0$ *for each k. Moreover,* $\|b\|_1 \leq 2\,\|f\|_1$.

(iv) $\lambda \leq \displaystyle⨍_{Q_k} |f| \leq 2^n \lambda$ *for all k.*

Proof of Lemma. The proof is a version of a typical argument called the 'Stopping Time Argument'. The set-up is as follows: give any cube Q,

$$⨍_Q |f| \leq \frac{1}{|Q|}\,\|f\|_1$$

which is bounded by λ whenever $|Q| > \frac{1}{\lambda}\,\|f\|_1$, i.e. $\ell(Q) > \left(\frac{\|f\|_1}{\lambda}\right)^{1/n}$.

Choose m large enough so that $2^m > \left(\frac{\|f\|_1}{\lambda}\right)^{1/n}$. Now let $\mathbb{D}(m)$ be the grid of dyadic cubes with side length 2^m. By our choice of m, for any $Q \in \mathbb{D}(m)$, $⨍_Q |f| < \lambda$.

The stopping time argument goes as follows: let $Q \in \mathbb{D}(m)$. We'll divide Q into 2^n equal parts. Then for any child Q' of Q, either

$$⨍_{Q'} |f| < \lambda \qquad \text{or} \qquad ⨍_{Q'} |f| \geq \lambda$$

If the second case happens, then we'll stop and include that Q' into the family we want. Otherwise, we'll continue the dyadic subdivision. Applying this argument to each $Q \in \mathbb{D}(m)$ and enumerating the resulting collection, we obtain our family $\{Q_k\}$. Note that this family is (i) non-overlapping and (ii) maximal, i.e. if Q_k^* is the (unique) parent of Q_k, then $⨍_{Q_k^*} |f| < \lambda$.

From this defining property of Q_k, we can show (iv): for any k, let Q_k^* be the parent of Q_k. Then

$$\lambda \leq \frac{1}{|Q_k|}\int_{Q_k} |f| \leq \frac{2^n}{|Q_k^*|}\int_{Q_k^*} |f| \leq 2^n \lambda$$

where the last inequality comes from the fact Q_k^* belongs to one of the 'good' cubes.

Similary we can prove (ii): by the key stopping time property, $|Q_k| \leq \frac{1}{\lambda}\int_{Q_k}|f|$. Thus, since the Q_k's are non-overlapping,

$$\sum_{k=1}^{\infty}|Q_k| = \frac{1}{\lambda}\int_{\cup_k Q_k}|f| \leq \frac{1}{\lambda}\,\|f\|_1$$

Next is the decomposition of f: define the 'good' function g by

$$g = \begin{cases} f & \text{on } \mathbb{R}^n \setminus \bigcup_{k=1}^{\infty} Q_k =: F \\ \fint_{Q_k} f & \text{in } Q_k \end{cases}$$

Given $x \in F$, then we may find a sequence of (dyadic) cubes $\{C_k\}$ in F containing x in the interior with $\ell(C_k) \to 0$. Because $\fint_{C_k}|f| < \lambda$, Lebesgue Differentiation theorem (Theorem 1.3) gives that $|f(x)| < \lambda$, i.e. $|f| < \lambda$ a.e. on F. Moreover, by (iv), $\left|\fint_{Q_k} f\right| \leq 2^n\lambda$, which gives an uniform bound on $\bigcup_{k=1}^{\infty} Q_k$. Combining the two bounds, we conclude $\|g\|_\infty \leq 2^n\lambda$.
At the same time,

$$\begin{aligned}\|g\|_1 &= \int_F |g| + \sum_{k=1}^{\infty}\int_{Q_k}|g| \\ &\leq \int_F |f| + \sum_{k=1}^{\infty}\int_{Q_k}|f| = \|f\|_1\end{aligned}$$

These two results prove (i).
Last to show (iii), naturally we set the 'bad' function b to be

$$b = f - g = \begin{cases} 0 & \text{on } F \\ f - \fint_{Q_k} f & \text{on } Q_k \end{cases}$$

Define $b_k = \left(f - \fint_{Q_k} f\right)\chi_{Q_k}$. Clearly then the b_k's decompose b and supp $b_k \subseteq Q_k$. To complete the proof, we compute that

$$\int b_k = \int_{Q_k} f - \int_{Q_k}\left(\fint_{Q_k} f\right) = \int_{Q_k} f - |Q_k|\fint_{Q_k} f = 0$$

□

Proof of Theorem 5.10, continued: Now we're ready to prove the w(1, 1) bound. Fix $\lambda > 0$ and suppose $f \in L^1 \cap L^2$ (which is dense in L^1. We need L^2 because apriori T is defined only there). Using decomposition of f,

$$|\{|Tf| > \lambda\}| \leq \left|\left\{|Tb| > \frac{\lambda}{2}\right\}\right| + \left|\left\{|Tg| > \frac{\lambda}{2}\right\}\right|$$

The 'good' part can be easily approximated: by Markov inequality,

$$\left|\left\{ |Tg| > \frac{\lambda}{2} \right\}\right| \le \frac{4}{\lambda^2} \int |Tg|^2 \le \frac{C}{\lambda^2} \int |g|^2$$

However again by part (i) of the lemma above, g itself is bounded, so

$$\left|\left\{ |Tg| > \frac{\lambda}{2} \right\}\right| \le \frac{C}{\lambda^2} \|g\|_\infty \int |g| \le \frac{C}{\lambda^2} \cdot 2^n \lambda \cdot \|f\|_1 = \frac{C}{\lambda} \|f\|_1$$

which is in the wanted form. It remains to bound the other 'distribution'. To do so, set $\widetilde{A} = \bigcup_{k=1}^{\infty} 5\sqrt{n} Q_k$ and $\widetilde{F} = \mathbb{R}^n \setminus \widetilde{A}$. Then

$$|\widetilde{A}| \le (5\sqrt{n})^n \sum_{k=1}^{\infty} |Q_k| \le \frac{C}{\lambda} \|f\|_1$$

by the lemma part (ii) for some constant C depending on n. Therefore,

$$\left|\left\{ |Tb| > \frac{\lambda}{2} \right\}\right| = \left|\left\{ x \in \widetilde{A} \,\middle|\, |Tb| > \frac{\lambda}{2} \right\}\right| + \left|\left\{ x \in \widetilde{F} \,\middle|\, |Tb| > \frac{\lambda}{2} \right\}\right|$$

The first expression is bounded by $|\widetilde{A}|$, which is already in the wanted form. It then remains to bound the second one.
We begin by observing that since we choose f to be in $L^1 \cap L^2$, b is also in such class. The lemma above gives a decomposition of b, so each b_k also lies in the same class. Thus because T is L^2-bounded, we may write

$$Tb = \sum_{k=1}^{\infty} Tb_k$$

Because $\widetilde{F} \subseteq \mathbb{R}^n \setminus 5\sqrt{n} Q_k$ for all k, an application of the Markov inequality gives

$$\begin{aligned} \left|\left\{ x \in \widetilde{F} \,\middle|\, |Tb| > \frac{\lambda}{2} \right\}\right| &\le \frac{2}{\lambda} \int_{\widetilde{F}} |Tb(x)|\, dx \\ &\le \frac{2}{\lambda} \sum_{k=1}^{\infty} \int_{\mathbb{R}^n \setminus 5\sqrt{n} Q_k} |Tb_k(x)|\, dx \\ &\le \frac{2}{\lambda} \sum_{k=1}^{\infty} \int_{\mathbb{R}^n \setminus 5\sqrt{n} Q_k} \left| \int_{5\sqrt{n} Q_k} K(x,y)\, b_k(y)\, dy \right| dx \end{aligned} \tag{9}$$

where we may express in terms of the kernel because the support of b_k and the domain of x-integration are separate.
Let y_k to be the center of Q_k and let $y \in Q_k$. Because $x \in \mathbb{R}^n \setminus 5\sqrt{n} Q_k$, from triangle inequality we are assured $|x - y| \ge 2|y - y_k|$.

From the assumption of the C-Z kernel,

$$|K(x,y) - K(x,y_k)| \le C \frac{|y-y_k|^\alpha}{|x-y_k|^{n+\alpha}}$$

Since b_k averages to 0, $\int K(x,y_k) b_k(y)\,dy = 0$. Thus,

$$\begin{aligned}
\left|\int K(x,y)b_k(y)\,,dy\right| &= \left|\int_{5\sqrt{n}Q_k} [K(x,y) - K(x,y_k)]\, b_k(y)\,dy\right| \\
&\le C \int_{\substack{|x-y_k| \\ \ge 2|y-y_k|}} \frac{|y-y_k|^\alpha}{|x-y_k|^{n+\alpha}} |b_k(y)|\,dy
\end{aligned}$$

Applying the estimate above to (9) and Fubini's theorem,

$$\begin{aligned}
\left|\left\{x \in \widetilde{F} \,\middle|\, |Tb| > \frac{\lambda}{2}\right\}\right| &\le \frac{C}{\lambda} \sum_{k=1}^\infty \int_{5\sqrt{n}Q_k} |b_k(y)| \int_{\substack{|x-y_k| \\ \ge 2|y-y_k|}} \frac{|y-y_k|^\alpha}{|x-y_k|^{n+\alpha}}\,dx\,dy \\
&\le \frac{C}{\lambda} \sum_{k=1}^\infty \int_{5\sqrt{n}Q_k} |b_k(y)|\,dy \\
&\le \frac{C}{\lambda} \|f\|_1
\end{aligned}$$

where we use the fact

$$\begin{aligned}
\int_{\substack{|x-y_k| \\ \ge 2|y-y_k|}} \frac{|y-y_k|^\alpha}{|x-y_k|^{n+\alpha}}\,dx &= |y-y_k|^\alpha \int_{S^{n-1}} \int_{2|y-y_k|}^\infty \frac{1}{\rho^{n+\alpha}} \rho^{n-1}\,d\rho\,dS^{n-1} \\
&= C_n |y-y_k|^\alpha \,[-\rho^{-\alpha}]_{2|y-y_k|}^\infty \\
&= 2^\alpha C_n =: C_{n,\alpha}
\end{aligned}$$

This gives the bound in the wanted form, so the proof is complete.

□

The theorem allows us to reduce our consideration for SIO to the question of whether the given operator is L^2-bounded, as the other L^p-bound then follows. In the following, we will give some criterions pertaining to SIO of the convolution type. The non-convolution type will be considered later on.

The general set-up is as follows: consider $K(x) = \dfrac{\Omega(x)}{|x|^n}$ where $\Omega(\lambda x) = \Omega(x)$ for $\lambda > 0$ (homogenous of degree zero). Moreover, $\Omega \in L^1(S^{n-1})$ and $\int_{S^{n-1}} \Omega(\theta)\,d\sigma(\theta) = 0$. Define then the operator

$$Tf(x) := \text{p.v.} \int_{\mathbb{R}^n} \frac{\Omega(x-y)}{|x-y|^n} f(y)\,dy$$

To expand on the definition,

$$Tf(x) = \text{p.v. } K * f(x) = \lim_{\substack{\varepsilon\to 0\\ R\to 0}} K_{\varepsilon,R} * f(x)$$

where $K_{\varepsilon,R}(x) = K(x)\cdot \chi_{\{\varepsilon<|x|<R\}}$. Then we can prove this little proposition:

Proposition 5.12. *Suppose $\widehat{Tf} = m(\xi)\hat{f}(\xi)$. Then $T : L^2 \to L^2$ iff $m \in L^\infty$.*

Proof. One direction is easy: assuming $m \in L^\infty$, Plancherel formula gives

$$\|Tf\|_2 = \left\|\widehat{Tf}\right\|_2 = \left\|m\hat{f}\right\|_2 \le \|m\|_\infty \left\|\hat{f}\right\|_2 = \|m\|_\infty \|f\|_2$$

and thus $\|T\|_{\le} \|m\|_\infty$.
For the other direction, suppose $T : L^2 \to L^2$ but $m \notin L^\infty$. The for each N, there is a set $E_N \subseteq \mathbb{R}^n$ so that $|E_N| > 0$ and $|m| \ge N$ on E_N. In fact, it suffices if we consider when the E_N's are bounded (otherwise, consider its part contained in some large ball centered at the origin).
Because the E_N's are bounded, the function $\widehat{f_N} = \chi_{E_N}$ belong to L^2. Therefore, we may apply again Plancherel formula to see

$$\|Tf_N\|_2 = \left\|\widehat{Tf_N}\right\|_2 = \left\|m\widehat{f_N}\right\|_2 \ge N\,\|f_N\|_2$$

This holds for all N, so T cannot be L^2-bounded, a contradiction.

□

Note. In fact, $\|T\|_{=} \|m\|_\infty$.

The previous proposition, while simple, is not very practical, as we're hardly ever been assured the existence of Fourier multiplier m. However, we'll show that if such a multiplier exists, then T is L^2-bounded. Before that, we'll need to define a new class of functions:

Definition 5.13. The space $L\log L(X)$ is defined as follows: for any f in this space,

$$\int_X |f|\log(2+|f|) < \infty$$

Note. While the condition is slightly weaker than L^1, since logarithm grows slower than any polynomial power, this space is not contained in any L^p for $p > 1$. In fact, it lies between L^1 and L^p for any power $p > 1$.
Moreover, we have the following fact: given $E \subseteq \mathbb{R}^n$ compact,

$$L^1(E) \supsetneq L\log L(E) \supsetneq L^p(E) \qquad p > 1$$

The following theorem makes the condition on Fourier multiplier precise:

Theorem 5.14. *Suppose $Tf = p.v.\ K * f$ where $K(x) = \dfrac{\Omega(x)}{|x|^n}$ and Ω is homogenous of degree 0. If (i) Ω is odd and $\Omega \in L^1(S^{n-1})$ or (ii) $\Omega \in L\log L(S^{n-1})$, then $\widehat{Tf} = m\hat{f}$ with $m \in L^\infty$, and hence T is L^2-bounded and $\|T\|_{2\to 2} = \|m\|_\infty$.*

The proof relies on this lemma, that gives an explicit expression for m:

Lemma 5.15. *For $K(x) = \dfrac{\Omega(x)}{|x|^n}$ with Ω satisfying the condition of the theorem, the Fourier multiplier $m(\xi) = p.v.\ \widehat{K}(\xi)$ exists and has the expression*

$$m(\xi) = \int_{S^{n-1}} \left[\frac{-\pi i}{2} sgn(\xi \cdot y) + \log\left(\frac{1}{|\xi \cdot y|}\right)\right] \Omega(y)\, d\sigma(y)$$

for $\xi \in S^{n-1}$. Thus $m(\xi)$ is homogenous of degree zero.

Proof of Lemma. Consider

$$m(\xi) := \lim_{\varepsilon \to 0;\, R \to \infty} \widehat{K}_{\varepsilon,R}(\xi)$$

with

$$\widehat{K}_{\varepsilon,R}(\xi) = \int_{\varepsilon < |\xi| < R} \Omega(y)\, |y|^{-n}\, e^{-2\pi i \xi \cdot y}\, dy$$

Writing the 'kernel' in polar coordinate (and using the fact Ω is homogenous),

$$\begin{aligned}
\widehat{K}_{\varepsilon,R}(\xi) &= \int_{S^{n-1}} \Omega(\theta) \int_{\varepsilon}^{R} e^{-2\pi i \rho\omega \cdot r\theta} \frac{dr}{r}\, d\sigma(\theta) \\
&= \int_{S^{n-1}} \Omega(\theta) \underbrace{\int_{\varepsilon}^{R} [e^{-2\pi i \rho\omega \cdot r\theta} - \cos(2\pi \rho r)] \frac{dr}{r}}_{I_{\varepsilon,R}(\rho\omega,\theta)}\, d\sigma(\theta) \qquad (10)
\end{aligned}$$

where we inserted an integral which evaluates to zero. To see this, note that since Ω is odd, its integral over S^{n-1} yields zero. Since $\dfrac{\cos(2\pi\rho r)}{r}$ is integrable,

$$\int_{S^{n-1}} \Omega(\theta) \int_{\varepsilon}^{R} \cos(2\pi\rho r) \frac{dr}{r}\, d\sigma(\theta) = 0$$

We'll show that $I_{\varepsilon,R}$ gives the wanted expression. First consider its imaginary part:

$$\operatorname{Im} I_{\varepsilon,R}(\rho\omega,\theta) = -\int_{\varepsilon}^{R} \frac{\sin(2\pi\rho\omega \cdot r\theta)}{r}\, dr$$

which is uniformly bounded on $\rho\omega$ and θ. Moreover it converges, as $\varepsilon \to 0$ and $R \to \infty$, to

$$-\int_0^{\infty} \frac{\sin t}{t} \operatorname{sgn}(\omega \cdot \theta)\, dt = -\frac{\pi}{2} \operatorname{sgn}(\omega \cdot \theta) = -\frac{\pi}{2} \operatorname{sgn}(\xi \cdot y)$$

Next we'll compute the real part. Let $h(t) = \cos(2\pi t)$ and set $\lambda = \rho|\omega \cdot \theta|$. Then we may write

$$\operatorname{Re} I_{\varepsilon,R} = \int_{\varepsilon}^{R} \frac{h(\lambda r) - h(\rho r)}{r}\, dr$$

We claim that

$$\lim_{\substack{\varepsilon\to 0\\ R\to\infty}} \int_\varepsilon^R \frac{h(\lambda r)-h(\rho r)}{r}\,dr = h(0)\,\log\left(\frac{\rho}{r}\right)$$

We may rescale and break the integrals so that

$$\begin{aligned}
\mathrm{Re}\, I_{\varepsilon,R} &= \int_\varepsilon^R \frac{h(r)-h(\rho r/\lambda)}{r}\,dr \\
&= \int_\varepsilon^1 \frac{h(r)}{r}\,dr + \int_1^\infty \frac{h(r)}{r}\,dr - \left(\int_\varepsilon^{\lambda/\rho} \frac{h\left(\frac{\rho}{\lambda}r\right)}{r}\,dr + \int_{\lambda/\rho} \frac{h\left(\frac{\rho}{\lambda}r\right)}{r}\,dr\right)
\end{aligned}$$

By rescaling again with $r \mapsto \dfrac{\lambda}{\rho}r$, the last integral becomes

$$\int_{\lambda/\rho}^\infty \frac{h(\rho r/\lambda)}{r}\,dr = \int_1^\infty \frac{h(r)}{r}\,dr$$

and thus we obtain a cancellation. It then remains to consider the case when ε is on the neighbourhood of zero. From integrating by parts,

$$\int_\varepsilon^1 \frac{h(r)}{r}\,dr - \int_{\varepsilon\rho/\lambda}^1 \frac{h(r)}{r}\,dr = h(r)\,\log r\Big|_\varepsilon^1 - \int_\varepsilon^1 h'(r)\,\log r\,dr - \left(h(r)\log r\Big|_{\varepsilon\rho/\lambda}^1 + \int_{\varepsilon\rho/\lambda}^1 h'(r)\log r\,dr\right)$$

Because $|h'(r)| \sim r$ and any polynomial power dominates log, the integrals are absolutely convergent. In particular in the limit $\varepsilon \to 0$, both converges to $\int_0^1 h'(r)\,\log r\,dr$, so we obtain a cancellation.

The only thing left to consider is

$$-h(\varepsilon)\,\log\varepsilon + h\left(\frac{\rho}{\lambda}\varepsilon\right)\,\log\left(\frac{\rho}{\lambda}\varepsilon\right) = \left(h\left(\frac{\rho}{\lambda}\varepsilon\right) - h(\varepsilon)\right)\,\log\varepsilon + h\left(\frac{\rho}{\lambda}\varepsilon\right)\,\log\frac{\rho}{\lambda}$$

In the first term, because the term in parentheses vanishes with order ε (because h is continuous) and $\varepsilon\,\log\varepsilon \to 0$ as $\varepsilon \to 0$, the whole first term vanishes in the limit. Similarly, continuity of h gives that h in the second term simply tends to 1. This gives the claim.

Combining the real and part of (10), we get that

$$m(\xi) = \lim_{\substack{\varepsilon\to 0\\ R\to\infty}} \int_{S^{n-1}} \Omega(\theta)(\mathrm{Re}\, I_{\varepsilon,R} + i\mathrm{Im}\, I_{\varepsilon,R})\,d\sigma\theta = \int_{S^{n-1}} \left[-i\frac{\pi}{2}\mathrm{sgn}(\xi\cdot y) + \log\left(\frac{1}{|\xi\cdot y|}\right)\right]\,\Omega(\theta)\,d\sigma(\theta)$$

□

Proof of Theorem. Assuming the lemma above, if Ω is odd and in $L^1(S^{n-1})$, then

$$\int_{S^{n-1}} \log\left(\frac{1}{|\xi\cdot y|}\right)\,\Omega(y)\,dy = 0$$

because the integrand is odd (because $|\xi \cdot y|$ is even). Thus $\|m\|_\infty \le \frac{\pi}{2}\|\Omega\|_1$, i.e. $m \in L^\infty$. Use Proposition 5.12 to get the wanted conclusion.
Now suppose that $\Omega \in L\log L(S^{n-1})$. It suffices to show that the multiplier m is bounded. In particular, since the Orlicz space $L\log L$ sits in L^1,

$$\int_{S^{n-1}} \operatorname{sgn}(\xi \cdot y)\, \Omega(y)\, d\sigma(y)$$

is a bounded function of ξ. Then it suffices to show that

$$\int_{S^{n-1}} \log \frac{1}{|\xi \cdot y|} \Omega(y)\, d\sigma(y) \tag{11}$$

is bounded for all ξ. To simplify matters, we may assume after some rotation that $\xi = e_1$. So we'll show

$$\int_{S^{n-1}} \log \frac{1}{|y_1|} \Omega(y)\, d\sigma(y) = \sum_{k=0}^{\infty} \int_{E_k} \log \frac{1}{|y_1|}\, \Omega(y)\, d\sigma(y)$$

where

$$E_0 = \{y \in S^{n-1} \mid |\Omega(y)| \le 1\} \qquad E_k = \{y \in S^{n-1} \mid 2^{k-1} < |\Omega(y)| \le 2^k\}$$

is bounded for all ξ.
First note that $y \ne 0$, so $\log \dfrac{1}{|y_1|}$ is integrable. Since Ω itself is bounded,

$$\int_{E_0} \log \frac{1}{|y_1|} \Omega(y)\, d\sigma(y) \le C$$

for some $C > 0$.
Now we consider when $k \ge 1$. Observe that since $\Omega \in L\log L$,

$$\sum_{k=2}^{\infty} C\, |E_{k-1}| (k-1)\, 2^{k-1} \le \int_{|\Omega|>1} |\Omega(y)| \log(2 + |\Omega(y)|)\, d\sigma(y) < \infty$$

for some universal constant $C > 0$. Hence the series $\sum_{k=1}^{\infty} |E_k|\, k \cdot 2^k$ converges.
Consider now on each E_k when $\dfrac{1}{|y_1|} \le 2^{2k}$. Because $|y_1| \ge 1$ (i.e. $\log \frac{1}{|y_1|} \ge 0$) and $\log \dfrac{1}{|y|} \le Ck$ (because of our choice of $|y_1|$), then

$$\sum_{k=1}^{\infty} \int_{\substack{E_k \cap \\ \{|y_1|^{-1} \le 2^{2k}\}}} \log \frac{1}{|y_1|}\, |\Omega(y)|\, d\sigma(y) \le \sum_{k=1}^{\infty} |E_k|\, k\, 2^k < \infty$$

On the other hand, if $\dfrac{1}{|y_1|} > 2^{2k}$,

$$\int_{\substack{E_k \cap \\ \{|y_1|^{-1} > 2^{2k}\}}} \log \frac{1}{|y_1|}\, |\Omega(y)|\, d\sigma(y) \le 2^k \int_{\{|y_1| < 2^{-2k}\}} \log \frac{1}{|y_1|}\, d\sigma(y) \le -C \cdot 2^k \int_0^{2^{-2k}} \log y_1\, dy_1 = C\, k\, 2^{-k}$$

Summing for all k, we conclude that the sum is finite. From these two cases, we conclude that (11) is uniformly bounded, hence $m \in L^\infty$.

□

6 Fourier Multiplier

Consider an operator defined as $TfL = (m\widehat{f})^\vee$. Formally Tf is a convolution operator. Then from previous section, we know that T is L^2-bounded iff $m \in L^\infty$. In this section, we'll consider the L^p-boundedness for $p \neq 2$. As a motivating example, we'll consider when $m = \chi_B$ where B is the unit ball. Then Tf defined as above is L^2-bounded but not on any other L^p-spaces (This is a highly non-trivial result of Charles Fefferman).
The following theorem gives a necessary and sufficient condition for this question.

Theorem 6.1 (Hörmander Multiplier)**.** *Let $m \in L^\infty$ with $\|m\|_\infty \leq A < \infty$. Suppose $(Tf)^\wedge = m\widehat{f}$ with the condition that $m \in C^k$ away from the origin for some $k > n/2$. Assume also that we have the following estimate for m:*

$$\sup_{0<R<\infty} \left(R^{2|\alpha|-n} \int_{R\leq|\xi|\leq 2R} \left| \left(\frac{\partial}{\partial \xi}\right)^\alpha m(\xi) \right|^2 d\xi \right)^{1/2} \leq A$$

for all multi-indices $|\alpha| \leq k$. Then $T : L^p \to L^p$ for all $1 < p < \infty$ and

$$\|Tf\|_p \leq C_p A \ \|f\|_p$$

For $p = 1$, T is of type $w(1,1)$.

Note. The estimate (6.1) can be seen as implying that

$$\left(⨍_{R\leq|\xi|\leq 2R} \left| R^{|\alpha|} \left(\frac{\partial}{\partial \xi}\right)^\alpha m(\xi) \right|^2 d\xi \right)^{1/2} \leq A$$

uniformly in R. In particular, this estimate holds if

$$\left| \left(\frac{\partial}{\partial \xi}\right)^\alpha m(\xi) \right| \leq C|\xi|^{-|\alpha|}$$

Some examples of $m(\xi)$ with this property:

(i) If m is homogenous of degree zero (e.g. as in SIOs of the convolution type).

(ii) If $m(\xi) = |\xi|^{is}$ for $s \in \mathbb{R}$.

Proof. As in the Calderon-Zygmund theorem (Theorem 5.10), we note that it suffices to furnish a weak $(1,1)$ estimate, as the other cases follow by interpolation. Let $K := \check{m}$ be the kernel of T. We want to show that K satisfies the Hörmander condition

$$\int_{|x-y|>2y} |K(x-y) - K(x)|\, dx \leq M \tag{12}$$

for some M and all y.
Let $\varphi_j = \varphi_j(|\xi|)$ be a smooth function with compact support on $2^{j-2} \leq |\xi| < 2^{j+2}$ so that $\sum_{j=-\infty}^{\infty} \varphi_j(|\xi|) = 1$ (obviously we avoid the case $\xi = 0$). To construct such function, let $\psi \in C_c^\infty\left(\frac{1}{4}, 4\right)$ such that $0 \leq \psi \leq 1$ and $\psi \equiv 1$ on $\left[\frac{1}{2}, 2\right]$ (we may take any standard smoothing function with this property).
Set $\psi_j(\xi) = \psi\left(\frac{|\xi|}{2^j}\right)$ for $\xi \in \mathbb{R}^n$. Note that for any $\xi \in \mathbb{R}^n$, $\sum_{j=-\infty}^{\infty} \psi_j(\xi)$ is positive and finite, because only finitely many $\psi_j(\xi)$ is non-zero (but there's always at least one which is non-zero). Then we can define

$$\varphi_j(\xi) = \frac{\psi_j(\xi)}{\sum_{k=-\infty}^{\infty} \psi_k(\xi)}$$

This allows use to use the Paley-Littlewood decomposition of m into its spectrum. Set $m = \sum_{k=-\infty}^{\infty} m_j$ where $m_j = m \cdot \varphi_j$. Let m^N the partial sum of the m_j's from $-N$ to N. Clearly m^N is also of class C_c^k (because m is C^k and the φ_j's are C_c^∞. In particular, this implies $m^N \in L^1 \cap L^2$.
We may then analogously define $K_j := \widecheck{m_j}$ and K^N to be its finite sum. We want to show that K^N satisfies (12) uniformly in N, i.e.

$$\sup_N \int_{|x|\geq 2|y|} |K^N(x-y) - K^N(x)|\, dx \leq C\, A$$

Notice that $\|m^N\|_\infty \leq \|m\|_\infty \leq A$. Then by the same argument as Theorem 5.10 (because K^N is a Calderon-Zygmund kernel), if we define $T^N f := K^N * f$,

$$\left|\left\{|T^N f| > \lambda\right\}\right| \leq \frac{CA}{\lambda} \|f\|_1$$

uniformly in

□

7 Fractional Integral and Sobolev Spaces

Recall that for $f \in \mathcal{S}$, the Foruier transform turns differentiation into multiplication, i.e. $\left(\frac{\partial}{\partial \xi_j} f\right)^\wedge(\xi) = 2\pi i \xi_j \hat{f}(\xi)$. In particular, since it plays well with multiple differentiation, we have $(-\Delta f)^\wedge(\xi) = 4\pi^2 |\xi|^2 \hat{f}(\xi)$. This observation motivates the following definition.

Definition 7.1. For $f \in \mathcal{S}$ and $\beta \geq 0$ (and possibly larger class of function, whenever it makes sense), the fractional power of the Laplacian is defined by

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

When $\beta > 0$, $(-\Delta)^{\beta/2}$ is called the fractional differentiation operator. On the other hand, as in Fundamental Theorem of Calculus, we'll call this operator to be fractional integration operator if $0 > \beta > -n$. These two are (for now) formally inverse of each other. For convenience, we'll write $\beta = -\alpha$, so whenever $0 < \alpha < n$, we may define the fractional integral operator as

$$I_{\alpha} f := (-\Delta)^{\alpha/2} f$$

This first lemma will help us explore the properties of this fractional integral operator.

Lemma 7.2. *Let* $0 < \alpha < n$. *Then*

$$(|x|^{\alpha-n})^{\wedge}(\xi) = C_{n,\alpha}|\xi|^{-\alpha}$$

in the sense of tempered distribution, i.e. for all $\varphi \in \mathcal{S}$,

$$\int_{\mathbb{R}^n} |x|^{\alpha-n}\widehat{\varphi}(x)\, dx = C_{n,\alpha} \int_{\mathbb{R}^n} |\xi|^{-\alpha}\varphi(\xi)\, d\xi$$

Note. Ultimately this lemma says that we may 'toss' FT for functions not in L^1 or L^2, up to some constant $C_{n,\alpha}$.

Proof. Recall the formula for FT of Gaussian functions:

$$\int_{\mathbb{R}^n} e^{-\pi\delta|x|^2}\widehat{\varphi}(x)\, dx = \delta^{-n/2} \int_{\mathbb{R}^n} e^{-\pi|\xi|^2/\delta}\, \varphi(\xi)\, d\xi \tag{13}$$

Multiply both sides by $\delta^{\beta-1}$, where $\beta = \dfrac{n-\alpha}{2}$, then integrate in δ with respect to $d\delta$. The LHS of the (13) will give

$$\begin{aligned}
\int_0^{\infty} \delta^{\beta} \int_{\mathbb{R}^n} e^{-\pi\delta|x|^2}\widehat{\varphi}(x)\, dx \,\frac{d\delta}{\delta} &= \int_{\mathbb{R}^n} \widehat{\varphi}(x) \int_0^{\infty} \delta^{\beta} e^{-\pi\delta|\xi|^2}\,\frac{d\delta}{\delta}\, dx && \text{(Fubini)}\\
&= \pi^{-\beta} \int_{\mathbb{R}^n} \widehat{\varphi}(x)\,|x|^{-2\beta} \int_0^{\infty} e^{-\delta}\delta^{\beta}\,\frac{d\delta}{\delta}\, dx && \left(\delta \mapsto \frac{\delta}{\pi|x|^2}\right)\\
&= \frac{\Gamma\left(\frac{n-\alpha}{2}\right)}{\pi^{\beta}} \int_{\mathbb{R}^n} \widehat{\varphi}(x)\,|x|^{\alpha-n}\, dx
\end{aligned}$$

On the other hand, the RHS of (13) gives

$$\delta^{-n/2} \int_0^{\infty} \delta^{\beta} \int_{\mathbb{R}^n} e^{-\pi|\xi|^2/\delta}\varphi(x)\, dx\, d\delta = \int_{\mathbb{R}^n} \varphi(\xi) \int_0^{\infty} \delta^{-\alpha/2} e^{-\pi|\xi|^2/\delta}\,\frac{d\delta}{\delta}\, d\xi \qquad \text{(Fubini)}$$

$$= \pi^{-\alpha/2} \int_{\mathbb{R}^n} \varphi(\xi)\, |\xi|^{-\alpha} \int_0^\infty \delta^{-\alpha/2} e^{-1/\delta} \frac{d\delta}{\delta}\, d\xi \quad (\delta \mapsto \pi\delta|\xi|^2)$$

$$= \frac{\Gamma\left(\frac{\alpha}{2}\right)}{\pi^{\alpha/2}} \int_{\mathbb{R}^n} \varphi(\xi)\, |\xi|^{-\alpha}\, d\xi$$

Setting the expression for RHS and LHS to be equal, we get the wanted expression with

$$C_{n,\alpha} = \frac{\Gamma((n-\alpha)/2)}{\Gamma(\alpha/2)}\, \pi^{\beta-\alpha/2}$$

□

This lemma will help us to get an explicit expression for the operator I_α, which is in the following proposition.

Proposition 7.3. *Let* $0 < \alpha < n$. *Then for* $f \in \mathcal{S}$,

$$I_\alpha f(x) = \frac{1}{C(n,\alpha)} \int_{\mathbb{R}^n} |x-y|^{\alpha-n}\, f(y)\, dy \tag{14}$$

in the sense of distribution. Here $C(n,\alpha) = \pi^{n/2} 2^\alpha \dfrac{\Gamma(\alpha/2)}{\Gamma(\,(n-\alpha)/2)}$.

Proof. Let $f, \varphi \in \mathcal{S}$ and define

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

$$\begin{aligned}
\int_{\mathbb{R}^n} I_\alpha f(x)\, \widehat{\varphi}(x)\, dx &= \int_{\mathbb{R}^n} f(y) \int_{\mathbb{R}^n} \frac{\widehat{\varphi}(x)}{|x-y|^{n-\alpha}}\, dx\, dy && \text{(Fubini)} \\
&= \int_{\mathbb{R}^n} f(y) \int_{\mathbb{R}^n} \frac{\widehat{\varphi}(x-y)}{|x|^{n-\alpha}}\, dx\, dy && \text{(Convolution property)} \\
&= C_{n,\alpha} \int_{\mathbb{R}^n} f(y) \int_{\mathbb{R}^n} |\xi|^{-\alpha} e^{-2\pi i \xi\cdot y} \varphi(\xi)\, d\xi\, dy && \text{(Lemma 7.2 + FT of shift)} \\
&= C_{n,\alpha} \int_{\mathbb{R}^n} \varphi(\xi)\, |\xi|^{-\alpha} \int_{\mathbb{R}^n} f(y) e^{-2\pi i \xi\cdot y}\, dy\, d\xi && \text{(Fubini)} \\
&= C_{n,\alpha} \int_{\mathbb{R}^n} \varphi(\xi)\, |\xi|^{-\alpha}\, \widehat{f}(\xi)\, d\xi && \text{(Definition of FT)} \\
&= (2\pi)^\alpha C_{n,\alpha} \int_{\mathbb{R}^n} \varphi(\xi)\, (I_\alpha f)^\wedge(\xi)\, d\xi && \text{(Definition of } I_\alpha)
\end{aligned}$$

From this short computation, we conclude

$$\int_{\mathbb{R}^n} T_\alpha f(x)\, \widehat{\varphi}(x), dx = (2\pi)^\alpha C_{n,\alpha} \int_{\mathbb{R}^n} \varphi(\xi)\, (I_\alpha f)^\wedge(\xi)\, d\xi \tag{15}$$

Now set $g = (I_\alpha f)^\wedge(\xi)$ (which is integrable, because $\widehat{f} \in \mathcal{S}$, which means it beats any polynomial power) and let $h = \widehat{\phi}$. Continuing from (15),

$$\begin{aligned}
\int_{\mathbb{R}^n} T_\alpha f(x)\,\widehat{\varphi}(x)\,dx &= (2\pi)^\alpha C_{n,\alpha} \int_{\mathbb{R}^n} \check{h}(x)\,g(x)\,dx && \text{(Fourier inversion)} \\
&= (2\pi)^\alpha C_{n,\alpha} \int_{\mathbb{R}^n} h(\xi)\,\check{g}(\xi)\,d\xi && \text{(Proposition 3.12 for inverse FT)} \\
&= (2\pi)^\alpha C_{n,\alpha} \int_{\mathbb{R}^n} \hat{\varphi}\, I_\alpha f(\xi)\,d\xi
\end{aligned}$$

because $\hat{\varphi} \in \mathcal{S}$ is arbitrary, we conclude from the last expression

$$T_\alpha f(x) = (2\pi)^\alpha C_{n,\alpha} I_\alpha f(x)$$

and we obtain $C(n, \alpha) = (2\pi)^\alpha C_{n,\alpha}$.

□

For completeness, we will note some basic properties of fractional integrals which proofs are omitted.

Proposition 7.4 (Properties of Fractional Integrals). *(a)* $I_\alpha(I_\beta f) = I_{\alpha+\beta}$ *for* $f \in \mathcal{S}$ *and* $\alpha, \beta > 0$, $\alpha + \beta < n$.

(b) $-\Delta I_\alpha f = -I_\alpha \Delta f = I_{\alpha-2} f$ *for* $n \geq 3$ *and* $2 \leq \alpha \leq n$.

Up to now, we have considered fractional integrals as a kind of formal object, applied on class of functions that have very rapid decay (of class $\mathcal{S}$). However, I_α is an integration operator, so it is natural to consider its action on L^p-functions. This following theorem gives the definitive answer to this question.

Theorem 7.5 (Hardy-Littlewood-Sobolev). *Let* $0 < \alpha < n$, $1 \leq p < q < \infty$ *with* $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}$. *Then*

(i) if $f \in L^p$, *then* $I_\alpha f(x)$ *converges absolutely a.e. on* $\mathbb{R}^n$.

(ii) if $p > 1$, *then* $\left\|I_\alpha f\right\|_q \leq C_{p,\alpha,n} \|f\|_p$.

(iii) if $p = 1$, *then* I_α *is a* $w(1, q)$*-operator on* L^1.

Note. One might wonder about the relation between p and q, which is not quite the usual dual. To see where this comes from, consider the dilation operator τ_δ defined as

$$\tau_\delta f(x) = f(\delta x)$$

Then

$$\begin{aligned}
(\tau_{\delta^{-1}} I_\alpha \tau_\delta f)(x) &= \frac{1}{C(n, \alpha)} \int_{\mathbb{R}^n} \frac{1}{|\delta^{-1} x - y|^{n-\alpha}} f(\delta y)\,dy \\
&= \delta^{-\alpha} I_\alpha f(x) && (y \mapsto y/\delta)
\end{aligned}$$

Moreover, we have two simple observations that

$$\left\|\tau_p f\right\|_p \qquad \left\|\tau_{\delta^{-1}} I_\alpha f\right\|_q = \delta^{n/q}\left\|I_\alpha f\right\|_q$$

Combining these results, we conclude that for $f \in L^p$ and $q > 0$,

$$\delta^{-\alpha}\left\|I_\alpha f\right\|_q = \delta^{-n/p}\cdot\delta^{n/q}\left\|I_\alpha f\right\|_q$$

Simplifying the exponent on δ shows that the stated relation between p and q is necessary. In the theorem, we will prove that it is actually sufficient.

Proof. First note that (i) follows from (ii) or (iii), because if $I_\alpha f$ belongs to L^q or weak L^1, then $I_\alpha f$ is finite a.e., so the integral is absolutely convergent.
WLOG $f \geq 0$ (otherwise just consider the positive and negative parts). Let $R > 0$ to be a number chosen later. Then we may split (14) as

$$C(n,\alpha)\, I_\alpha f(x) = \underbrace{\int_{|x-y|\leq R} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy}_{A_R} + \underbrace{\int_{|x-y|>R} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy}_{B_R}$$

Note that we can write A_R as a convolution, e.g. $A_R = R^\alpha(\varphi_R * f)(x)$ where $\varphi_R(x) = R^{-n}\varphi\left(\frac{x}{R}\right)$ and $\varphi(x) = |x|^{\alpha-n}\chi_{|x|\leq 1}$. Observe that $\varphi \in L^1$ (because $-n < \alpha - n < 0$ and φ has compact support), radial, and decreasing. This gives us an approximation to the identity, so Theorem 2.3 gives that

$$|\varphi_R * f| \leq C \cdot Mf \tag{16}$$

uniformly on $\mathbb{R}$. In particular, if $p > 1$, (16) gives (via Corollary 1.5, take p-th power, use the bound, then take p-th root) $|A_R| \leq C\,\|f\|_p$, as wanted.
For B_R, we can approximate as follows

$$\begin{aligned}
B_R &\leq \|f\|_p \left(\int_{|x-y|>R} |x-y|^{(\alpha-n)p'}\, dy\right)^{1/p'} && \text{(Hölder's inequality)}\\
&= \|f\|_p \left(\int_{|x-y|>R} |x-y|^{(\alpha-n)p'}\, dy\right)^{1/p'} && \text{(Given relation on}\alpha)\\
&\leq R^{-n/q}\,\|f\|_p \left(\int_{|x-y|>R} |x-y|^{-n}\, dy\right)^{1/p'} && (\,|x-y| > R\,)\\
&\leq CR^{-n/q}\,\|f\|_p && (17)
\end{aligned}$$

We have obtained $\|f\|_p$ estimate for A_R and B_R. Now we're going to optimize this bound to determine R. Set (16) and (17) to be equal. Ignoring constants,

$$R^\alpha\ \|Mf\|_p = R^{-n/q}\,\|f\|_p \qquad \Rightarrow \qquad R = \left(\frac{\|Mf\|_p}{\|f\|_p}\right)^{-p/n}$$

and hence

$$I_\alpha f(x) \lesssim \|f\|_p \cdot R^{-n/q} \lesssim \|f\|_p^{1-p/q} \cdot Mf(x)^{p/q}$$

which, if integrated, will give

$$\int_{\mathbb{R}^n} |I_\alpha f(x)|^q \, dx \le C \|f\|_p^{q-p} \|Mf\|_p^p \le C \|f\|_p^q$$

if $p > 1$ by Corollary 1.5. If $p = 1$,

$$(I_\alpha f(x))^q \lesssim \|f\|_1^{q-1} \cdot Mf(x)$$

and thus

$$\begin{aligned} \left|\{|I_\alpha f| > \lambda\}\right| &\le \left|\left\{ Mf(x) \ge C \frac{\lambda^q}{\|f\|_1^{q-1}} \right\}\right| && \text{(the estimate above)} \\ &\le C \left(\frac{\|f\|_1}{\lambda}\right)^q && \text{(w(1, 1) bound of } Mf) \end{aligned}$$

This completes the proof.

□

As a counterpart to the fractional integration, we may talk about the fractional differentitation. Let $D^\alpha = (-\Delta)^{\alpha/2}$ for $0 < \alpha < n$. Sometimes the class $\mathcal{S}$ is too strong, as it requires functions which derivatives decay faster than any polynomial. On the other hand, some control of the decay is still needed. Starting from $\mathcal{S}$, we may then define the following class:

Definition 7.6. The homogenous Sobolev space $\dot{L}^p_\alpha$ is defined as the completion of $\mathcal{S}$ with respect to the norm $\|f\|_{\dot{L}^p_\alpha} := \|D^\alpha f\|_p$ for $1 < p < \infty$.

Becuase $D^\alpha = (I_\alpha)^{-1}$, formally we may think that $\dot{L}^p_\alpha = I_\alpha(L^p)$, i.e. if $g = D^\alpha f$, $g \in L^p$ iff $f \in \dot{L}^p_\alpha$. Then observe that the HLS theorem (Theorem 7.5) implies that

$$\left\|I_\alpha g\right\|_q \le C_{p,\alpha} \|g\|_p \Leftrightarrow \|f\|_q \le C_{p,\alpha} \|D^\alpha f\|_p = C_{p,\alpha} \|f\|_{\dot{L}^p_\alpha}$$

with α, p, and q is in the same set-up of the theorem. Now, the second inequality implies implies that if f is in $\dot{L}^p_\alpha$, then f is also in L^q. In other words, we obtain an embedding of Sobolev space $\dot{L}^p_\alpha \hookrightarrow L^q$. In particular, eventhough f is an equivalence class (module constant, say), this embedding allows us to choose an appropriate representation of its elements, that is those which are also in L^q. This has implication for Riesz transformation:

Proposition 7.7. $(R_j f)^\wedge(\xi) = -i \dfrac{\xi_j}{|\xi|} \widehat{f}(\xi)$

Proof. Recall that for odd homogenous Ω, Theorem 5.14 and its associated lemma gives an explicit expression for the multiplier m of the convolution operator p.v.$\frac{\Omega(x)}{|x|^n} * f$. For R_j, $\Omega(x) = C_n x_j$, which clearly satisfies Theorem 5.14. It then follows, with same theorem, that

$$(R_j f)^\wedge(\xi) = m(\xi)\,\widehat{f}(\xi)$$

and thus it suffices (from the expression for m) to show that for $\xi \in S^{n-1}$,

$$-i\xi_j = -i\frac{\pi}{2}\int_{S^{n-1}} \operatorname{sgn}(\omega\cdot\xi)\,(C_n\omega_j)\,d\sigma(\omega)$$

In particular, since the multiplier m is the FT of the kernel of the operator, we need to show that p.v.$\left(\frac{x_j}{|x|^{n+1}}\right)^\wedge = C\frac{\xi_j}{|\xi|}$ in the sense of tempered distribution. Let $\varphi \in \mathcal{S}$. By definition,

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

Let η be a radial and smooth cut-off function so that $\eta \cong 1$ on $B_1(0)$. Then we may split the integral above as

$$\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} \frac{x_j}{|x|^{n+1}}\,\varphi(x)\,dx = \lim_{\varepsilon\to 0}\left[\underbrace{\int_{|x|>\varepsilon}\frac{x_j}{|x|^{n+1}}\,\varphi(x)\,\eta(x)\,dx}_{\text{I}} + \underbrace{\int_{|x|>\varepsilon}\frac{x_j}{|x|^{n+1}}\,\varphi(x)\,(1-\eta(x))\,dx}_{\text{II}}\right]$$

We'll begin by estimating I. Note that because η is radial, it is even, so $\int_{|x|>\varepsilon}\frac{x_j}{|x|^{n+1}}\eta(x)\,dx = 0$ (because x_j is odd). Thus

$$\text{I} = \lim_{\varepsilon\to 0}\int_{|x|>\varepsilon}\frac{x_j}{|x|^{n+1}}\,\eta(x)\,[\varphi(x)-\varphi(0)]\,dx$$

Because φ is smooth, MVT says $\varphi(x)-\varphi(0)\sim x$, so the degree of singularity goes down by one and now the integral is absolutely convergent. Thus we get that

$$\text{I} = \int_{\mathbb{R}^n}\frac{x_j}{|x|^{n+1}}\,\eta(x)\,[\varphi(x)-\varphi(0)]\,dx$$

Using the observation that $\frac{x_j}{|x|^{n+1}} = C_n'\frac{\partial}{\partial x_j}\frac{1}{|x|^{n-1}}$ and integration by parts, we may rewrite I as

$$\text{I} = \underbrace{-C_n'\int_{\mathbb{R}^n}\frac{|x|^{n-1}}{}\eta(x)\,\frac{\partial\varphi}{\partial x_j}\,dx}_{\text{I'}}\underbrace{-\,C_n'\int_{\mathbb{R}^n}\frac{1}{|x|^{n-1}}\,\frac{\partial\eta}{\partial x_j}\,[\varphi(x)-\varphi(0)]\,dx}_{\text{I''}}$$

Similarly,

$$II = \underbrace{-C_n' \int_{\mathbb{R}^n} \frac{1}{|x|^{n-1}} [1-\eta(x)] \frac{\partial \varphi}{\partial x_j} \, dx}_{\text{II'}} + \underbrace{C_n' \int_{\mathbb{R}^n} \frac{1}{|x|^{n-1}} \frac{\partial \eta}{\partial x_j} \varphi(x) \, dx}_{\text{II''}}$$

Combining these pieces appropriately,

$$\begin{aligned} \text{p.v.} \int_{\mathbb{R}^n} \frac{x_j}{|x|^{n+1}} \varphi(x)\, dx &= (\text{I'} + \text{II'}) + (\text{I''} + \text{II''}) \\ &= \underbrace{-C_n' \int_{\mathbb{R}^n} \frac{1}{|x|^{n-1}} \frac{\partial \varphi}{\partial x_j} \, dx}_{\text{III}} + \underbrace{C_n' \varphi(0) \int_{\mathbb{R}^n} \frac{1}{|x|^{n-1}} \frac{\partial \eta}{\partial x_j} \, dx}_{\text{IV}} \end{aligned}$$

Because η is even and radial, $\frac{\partial \eta}{\partial x_j}$ is odd (we'll get a factor of x_j, which is odd), so IV vanishes. Moreover, because $\left(\frac{1}{|x|^{n-1}}\right)^\wedge = C \frac{1}{|\xi|}$, Proposition 3.12 says

$$\text{III} = C \int_{\mathbb{R}^n} \frac{1}{|\xi|} \left(\frac{\partial \varphi}{\partial x_j}\right)^\vee (\xi) \, d\xi = C \int_{\mathbb{R}^n} \frac{1}{|\xi|} \xi_j \check{\varphi}(\xi) \, d\xi$$

In particular, if we let $\psi \in \mathcal{S}$ and set $\varphi = \hat{\psi}$, we have just shown that

$$\text{p.v.} \int_{\mathbb{R}^n} \frac{x_j}{|x|^{n+1}} \hat{\psi}(x) \, dx = C \int_{\mathbb{R}^n} \frac{\xi_j}{|\xi|} \psi(\xi) \, d\xi$$

and thus, again from Proposition 3.12 we conclude

$$\left(\frac{x_j}{|x|^{n+1}}\right)^\wedge = C \frac{\xi_j}{|\xi|}$$

in the sense of tempered distribution, as claimed. In particular, if $\xi \in S^{n-1}$ (so $|\xi| = 1$), the formula for multiplier from Theorem 5.14 gives that

$$C\xi_j = m(\xi) = -i\frac{\pi}{2} C_n \int_{S^{n-1}} \operatorname{sgn}(\xi \cdot \omega)\, \omega_j \, d\sigma(\omega)$$

Observe that this reuslt is independent of the choice of ξ and is rotation invariant, so we may take for simplicity $j = 1$ and $\xi = e_1$ (so $\xi_1 = 1$). Thus,

$$i \cdot C \cdot \xi_1 = \frac{\pi}{2} C_{n,\alpha} \int_{S^{n-1}} |\omega_1| \, d\sigma(\omega)$$

But the integral evaluates to $2 \cdot \pi^{(n-1)/2}/\Gamma((n+1)/2)$. Combined with $C_{n,\alpha} = \Gamma((n+1)/2)/\pi^{(n+1)/2}$, we conclude that $C = -i$, as wanted.

□

Corollary 7.8. $R_j = \frac{\partial}{\partial x_j} I_1$. *In particular,* $\sum_{j=1}^{n} R_j^2 = -Id$.

This corollary shows that R_j (or rather, R_j^2), decomposes f. The makes the comment about Sobolev embedding precise:

Corollary 7.9. *Let* $1 < p < q < \infty$ *so that* $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$. *Then*

$$\|f\|_q \le C_{p,n} \|\nabla f\|_p \approx \|f\|_{\dot{L}_1^p}$$

Proof. Set $D = D^1$ and let $f \in \mathcal{S}$ and $g = Df$ (so $f = I_1 g$). By H-L-S theorem (Theorem 7.5),

$$\|f\|_q = \|I_1 g\|_q \le C_{n,p} \|g\|_p$$

Using the previous corollary,

$$\|f\|_q = C_{n,p} \left\| \sum_{j=1}^{n} R_j^2 g \right\|_p \lesssim \sum_{j=1}^{n} \left\| R_j^2 g \right\|_p \lesssim \sum_{j=1}^{n} \left\| R_j g \right\|_p$$

where we used boundedness of R_j at the last inequality. From the previous corollary again,

$$\|f\|_q \le C \sum_{j=1}^{n} \left\| \frac{\partial}{\partial x_j} I_1 g \right\|_p = C \sum_{j=1}^{n} \left\| \frac{\partial}{\partial x_j} f \right\|_p \approx \|\nabla f\|_p$$

This shows I_1 and gradient are comparable. On the other hand,

$$\|\nabla f\|_p = \|\nabla I_1 Df\|_p = \|RDf\|_p \le \|Df\|_p = \|f\|_{\dot{L}_1^p}$$

This gives the equivalence with the Sobolev norm. □

Note. If we let $W^{1,p}$ to be the completion of $\mathcal{S}$ with respect to the norm $\|f\|_{W^{1,p}} = \|\nabla f\|_p$. This is what is usually known as Sobolev space. By the last corollary, we can see that $W^{1,p} \cong \dot{L}_1^p$.

8 Littlewood-Paley Theorem

Let $\psi \in C_c^\infty(B_1(0))$ be a real, radial function with average 0. We can see that for some $C, \alpha > 0$.

$$|\hat{\psi}(\xi)| \le C \min(|\xi|^\alpha, |\xi|^{-\alpha})$$

In other words, because $\hat{\psi} \in \mathcal{S}$, we can control the decay of $\hat{\psi}$ by $|\xi|^\alpha$ near the origin and by $|\xi|^{-\alpha}$ far from it. Note here that we write ψ as if it is a function on $\mathbb{R}$. Of course this is justified, because ψ is radial.

Now observe that by MVT, $|\widehat{\psi}(\xi)| = |\widehat{\psi}(\xi) - \widehat{\psi}(0)| \leq C|\xi|$, because $\widehat{\psi} \in \mathcal{S}$. Combining this observation with the bound above, we get that if ψ is non-trivial,

$$0 < \int_0^\infty |\widehat{\psi}(t)|^2 \frac{dt}{t} = C < \infty$$

In particular, normalizing with $\widehat{\psi}/\sqrt{c}$, we may assume that the integral evaluates to 1. From this formula, we set $\psi_t(x) = t^{-n}\psi\left(\frac{x}{t}\right)$ and define $Q_t f := \psi_t * f$.

Proposition 8.1. $\int_0^\infty Q_t^2 \cdot \, dt = Id$ *in* L^2*-operator topology, i.e. for all* $f \in L^2$,

$$\lim_{\varepsilon \to 0} \left\| \int_\varepsilon^{1/\varepsilon} Q_t^2 f \, dt - f \right\|_2 = 0$$

Proof. Fix $f \in L^2$. We'll estimate the limit above as follows:

$$\begin{aligned}
\left\| \int_\varepsilon^{1/\varepsilon} Q_t^2 f \frac{dt}{t} - f \right\|_2 &= \left\| \int_\varepsilon^{1/\varepsilon} (Q_t^2 f)^\wedge \frac{dt}{t} - \widehat{f} \right\|_2 && \text{(Plancherel in } \xi \text{ for } L^2) \\
&= \left\| \widehat{f}(\xi) \left[\int_\varepsilon^{1/\varepsilon} |\widehat{\psi}(t\xi)|^2 \frac{dt}{t} - 1 \right] \right\|_2 && \text{(FT of convolution is multiplication)} \\
&= \left\| \widehat{f}(\xi) \left[\int_{\varepsilon/|\xi|}^{|\xi|\varepsilon} |\widehat{\psi}(t)|^2 \frac{dt}{t} - 1 \right] \right\|_2
\end{aligned}$$

But we have set the integral over the whole positive-half to be 1, so we should conclude that the expression above vanishes in the limit.

□

Note. This proposition is sometimes called Calderon reproducing formula or continuous Littlewood-Paley decomposition. Note here that when we use Q_t, we implicitly choose ψ. In what follows, we will assume that $\int_0^\infty |\widehat{\psi}(t)|^2 \frac{dt}{t} = 1$.

Define the vertical square function for a ψ as above to be

$$g(f) = g_\psi(f) := \left(\int_0^\infty |Q_t|^2 \frac{dt}{t} \right)^{1/2}$$

Note that g is actually an operator. Also, we again assume an implicit function ψ associated with Q_t (and hence g). We'll show that g is an isometry in L^2.

Proposition 8.2. *Let* ψ *be as above and* $f \in L^2$*. Then the operator* $f \mapsto g_\psi(f)$ *is* L^2*-bounded and* $\left\| g_\psi(f) \right\|_2 = \|f\|_2$.

Proof. The proof is by simple computation:

$$\begin{aligned}
\|g(f)\|_2^2 &= \int_{\mathbb{R}^n}\int_0^\infty |Q_t f(x)|^2 \frac{dt}{t}\,dx \\
&= \int_0^\infty \int_{\mathbb{R}^n} |Q_t f(x)|^2\,dx\,\frac{dt}{t} && \text{(Fubini)} \\
&= \int_0^\infty \int_{\mathbb{R}^n} |(Q_t f)^\wedge(\xi)|^2\,d\xi\,\frac{dt}{t} && \text{(Plancherel)} \\
&= \int_0^\infty \int_{\mathbb{R}^n} |\hat{\psi}(t|\xi|)\,\hat{f}(\xi)|^2\,d\xi\,\frac{dt}{t} && \text{(FT of convolution is multiplication)} \\
&= \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \int_0^\infty |\hat{\psi}(t|\xi|)|^2 \frac{dt}{t}\,d\xi && \text{(Tonelli)} \\
&= \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \int_0^\infty |\hat{\psi}(t)|^2 \frac{dt}{t}\,d\xi && \left(t \mapsto \frac{t}{|\xi|}\right) \\
&= \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2\,d\xi
\end{aligned}$$

where the last line uses the normalization of ψ. This completes the proof.

□

We have considered the case when ψ is a nice function (radial, etc.) and operator Q_t is of convolution type. Next we'll consider the more general case, that is let ψ_t to satisfy the following conditions:

Definition 8.3. The function ψ_t satisfies the Littlewood-Paley condition for some $\alpha > 0$ if

(i) $|\psi_t(x,y)| \le C\dfrac{t^\alpha}{(t+|x-y|)^{n+\alpha}}$

(ii) $|\psi_t(x,y+h) - \psi_t(x,y)| \le C\dfrac{|h|^\alpha}{(t+|x-y|)^{n+\alpha}}$ if $|h| \le t$.

Compared the the condition for C-Z kernel, (ii) says that Hölder continuity only applies in the second variable locally, instead when the points are far away. Note here that we don't put any other condition on ψ_t.

Now we define the operator $\Theta_t f(x) := \int_{\mathbb{R}^n} \psi_t(x,y)\,f(y)\,dy$ for ψ_t satisfying L-P condition. The following theorem shows that we can basically reduce the consideration to the convolution type.

Theorem 8.4. *Let Θ_t be defined as above. Suppose that $\Theta_t 1 = 0$ $\left(\text{i.e. } \int_{\mathbb{R}^n} \psi_t(x,y)\,dy = 0\right)$. Then*

$$\int_{\mathbb{R}^n}\int_0^\infty |\Theta_t f(x)|^2 \frac{dt}{t}\,dx \le C\,\|f\|_2^2$$

i.e. the operator $g_\theta f := \left(\int_0^\infty |\Theta_t f(x)|^2 \frac{dt}{t}\right)^{1/2}$ *is L^2-bounded.*

Proof. Note that by (i) on L-P condition, if we replace ψ_t on Θ_t by its size bound, the estimate will be of convolution type, so Theorem 2.3 says $\sup_{t>0}|\Theta_t f| \le C \cdot Mf$. In particular, this means (i) $\Theta_t f$ is uniformly integrable for all $t > 0$ and $f \in L^2(\mathbb{R}^n)$ and (ii) Θ_t is L^2-bounded. Hence if we let $Q_t f = \varphi_t * f$, where φ is of class $C_c^\infty(B_1(0))$, radial, and averages zero, Proposition 8.1 allows us to write

$$\Theta_t f = \Theta_t\left(\int_0^\infty Q_s^2 f \,\frac{ds}{s}\right) = \int_0^\infty \Theta_t(Q_s^2 f)\,\frac{ds}{s}$$

where we may bring Θ_t in the integral because $\int_0^\infty Q_s^2 f\,\frac{ds}{s}$ is convergent in L^2 and Θ_t is L^2-bounded.

$$\begin{aligned}\int_{\mathbb{R}^n}\int_0^\infty |\Theta_t f(x)|^2\,\frac{dt}{t}\,dx &= \int_{\mathbb{R}^n}\int_0^\infty \left|\int_0^\infty \Theta_t(Q_s^2 f)(x)\,\frac{ds}{s}\right|^2 \frac{dt}{t}\,dx \\ &\le \int_{\mathbb{R}^n}\int_0^\infty \left|\int_0^\infty \min\left(\frac{s}{t},\frac{t}{s}\right)^{\beta/2}\cdot\max\left(\frac{s}{t},\frac{t}{s}\right)^{\beta/2}|\Theta_t(Q_s^2 f)(x)|\,\frac{ds}{s}\right|^2\frac{dt}{t}\,dx\end{aligned}$$

for β to be chosen. Note that this insertion is justified because for any given s and t, the minimum and maximum will cancel. If we apply C-S inequality in the s-integral above,

$$\begin{aligned}\int_{\mathbb{R}^n}\int_0^\infty |\Theta_t f(x)|^2\,\frac{dt}{t}\,dx &\le \int_{\mathbb{R}^n}\int_0^\infty \underbrace{\left(\int_0^\infty \min\left(\frac{s}{t},\frac{t}{s}\right)^{\beta}\frac{ds}{s}\right)}_{\text{constant } C_\beta}\left(\int_0^\infty \max\left(\frac{s}{t},\frac{t}{s}\right)^{\beta}|\Theta_t(Q_s^2 f)(x)|^2\,\frac{ds}{s}\right)\frac{dt}{t}\,dx \\ &= C_\beta\int_0^\infty\int_0^\infty \max\left(\frac{s}{t},\frac{t}{s}\right)^{\beta}\int_{\mathbb{R}^n}|\Theta_t Q_s^2 f|^2\,dx\,\frac{ds}{s}\,\frac{dt}{t}\end{aligned}$$

where the last equality uses Tonelli theorem. We claim now that there exists $\gamma > 0$ so that

$$\|\Theta_t Q_s\|_{2\to 2} \le \min\left(\frac{s}{t},\frac{t}{s}\right)^{\gamma}$$

Proof of Claim. Suppose $s < t$. Recall that $Q_s f := \varphi_s * f$ with φ_s is the usual approximation to identity with average zero and normalization $\int_0^\infty |\widehat{\varphi}_s(z)|^2\,\frac{dz}{z} = 1$. From the definition, we may write

$$\begin{aligned}\Theta_t Q_s f(x) &= \int_{\mathbb{R}^n}\psi_t(x,z)\,Q_s f(z)\,dz \\ &= \int_{\mathbb{R}^n}\psi_t(x,z)\int_{\mathbb{R}^n}\varphi_s(z-y)\,f(y)\,dy\,dz \\ &= \int_{\mathbb{R}^n} f(y)\underbrace{\int_{\mathbb{R}^n}\psi_t(x,z)\,\varphi_s(z-y)\,dz}_{\Phi_{s,t}(x,y)}\,dy \qquad \text{(Fubini)}\end{aligned}$$

Because φ has average zero, we may insert it to $\Phi_{s,t}$ to get

$$\left|\Phi_{s,t}(x,y)\right| = \left|\int_{\mathbb{R}^n} [\psi_t(x,z) - \psi_t(x,y)]\, \varphi_s(z-y)\, dz\right|$$

Note that because $\operatorname{supp}\varphi_s \subseteq B_s(0)$, the only relevant area for integration is when $|z-y| < s$. Because φ is also bounded by 1, the integration turns practically into and an averaging integral (because φ_s gives a factor of s^{-n}). If we replace ψ_t with the bound on L-P condition (ii),

$$\begin{aligned}|\Phi_{s,t}(x,y)| &\lesssim \fint_{|y-z|<s} \frac{s^\alpha}{(t+|x-y|)^{n+\alpha}}\, dz \\ &\leq \left(\frac{s}{t}\right)^\alpha \frac{t^\alpha}{(t+|x-y|)^{n+\alpha}}\end{aligned}$$

Applying this estimate back to the main expression,

$$\begin{aligned}|\Theta_t Q_s f(x)| &\leq C\left(\frac{s}{t}\right)^\alpha \int_{\mathbb{R}^n} \frac{t^\alpha}{(t+|x-y|)^{n+\alpha}}\, |f(y)|\, dy \\ &\leq C\left(\frac{s}{t}\right)^\alpha \cdot Mf(x) && \text{(Theorem 2.3)}\end{aligned}$$

When $f \in L^2$, the maximal operator is bounded, so

$$\left\|\Theta_t Q_s f\right\|_2 \leq C\left(\frac{s}{t}\right)^\alpha \|f\|_2$$

In this case we may take $\gamma = \alpha$ as in the L-P condition.

□

Continuing from the last line, using the claim to get rid of $\Theta_t Q_s$,

$$\begin{aligned}\int_{\mathbb{R}^n}\int_0^\infty |\Theta_t f(x)|^2 \frac{dt}{t}\, dx &\leq C_\beta \int_0^\infty \int_0^\infty \max\left(\frac{s}{t},\frac{t}{s}\right)^\beta \min\left(\frac{s}{t},\frac{t}{s}\right)^{2\gamma} \int_{\mathbb{R}^n} |Q_s f(x)|^2\, dx\, \frac{dt}{t}\,\frac{ds}{s} \\ &= C_\gamma \int_0^\infty \left(\int_0^\infty \min\left(\frac{s}{t},\frac{t}{s}\right)^\gamma \frac{dt}{t}\right) \int_{\mathbb{R}^n} |Q_s f(x)|^2\, dx\, \frac{ds}{s} && \text{(Take } \beta = \gamma\text{)} \\ &= C_\gamma^2 \int_{\mathbb{R}^n}\int_0^\infty |Q_s f(x)|^2 \frac{ds}{s}\, dx && \text{(Tonelli)} \\ &= C_\gamma^2 \|f\|_2^2 && \text{(Proposition 8.2)}\end{aligned}$$

□

We may define an alternative to the vertical square function by taking the integration over a conical section, instead of vertically letting t to range while keeping everything the same. Indeed, define the conical square function as

$$Sf(x) = S_\psi f(x) = \left(\int_0^\infty \int_{|x-y|<t} |\Theta_t f(y)|^2\, dy\, \frac{dt}{t^{n+1}}\right)^{1/2}$$

To see the relation to the vertical square function, observer that by inserting constants and taking t^n, we get measure of the ball $|x-y|<t$,

$$Sf(x)^2 \approx \int_0^\infty \fint_{|x-y|<t} |\Theta_t f(y)|^2 \, dy \, \frac{dt}{t}$$

So instead taking integration of $|Q_t f|^2/t$, we are taking integral over its averages, adding some degree of regularity. In fact, we'll show that these two are, as far as L^2-operators are concerned, equivalent.

Proposition 8.5. *Let $\Theta_t f(x)$ be as in the set-up of L-P theorem. Then*

$$\left\|S_\psi f\right\|_2 \approx \left\|g_\psi f\right\|_2 \lesssim \|f\|_2$$

Proof. The proof is straight calculation:

$$\begin{aligned}
\left\|S_\psi f\right\|_2^2 &\approx \int_{\mathbb{R}^n} \int_0^\infty \fint_{|x-y|<t} |\Theta_t f(y)|^2 \, dy \, \frac{dt}{} dx \\
&= \int_{\mathbb{R}^n} \int_0^\infty |\Theta_t f(y)|^2 \fint_{|x-y|<t} \, dy \, \frac{dt}{t} \, dx \\
&= \int_{\mathbb{R}^n} \int_0^\infty |\Theta_t f(y)|^2 \, dy \, \frac{dt}{t} && \text{(Average integral over } y \text{ gives 1)} \\
&= \left\|g_\psi(f)\right\|_2^2
\end{aligned}$$

□

8.1 Vector-valued SIO and L^p Extension of Littlewood-Paley Theory

Assume that H_1 is a separable Hilbert space and let f takes value at H_1. To be precise, for each $x \in \mathbb{R}^n$ we consider the mapping

$$\mathbb{R}^n \to H_1 \;:\; x \mapsto f(x) = \sum_j \alpha_j(x) \, e_j$$

where e_j is an orthonormal basis for H_1. Suppose also that for each pair $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $x \neq y$, the kernel $K(x, y)$ is a bounded linear mapping from H_1 to H_2, another separable Hilbert space. Then we may consider $Tf(x) : \mathbb{R}^n \to H_2$ given by

$$Tf(x) = \int K(x, y) \, f(y) \, dy$$

Definition 8.6. Let $\|K(x, y)\|$ denotes the operator norm on $\mathcal{L}(H_1, H_2)$ (for fixed $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$). Then call $K(x, y)$ to be a standard vector-valued C-Z kernel if

(i) $\|K(x,y)\| \le C\,\|x-y\|^{-n}$.

(ii) $\|K(x+h,y)-K(x,y)\| + \|K(x,y+h)-K(x,y)\| \le C\,\dfrac{|h|^\alpha}{|x-y|^{n+\alpha}}$ for $|h| \le \frac{1}{2}|x-y|$. Alternatively we may split the condition in two:

$$\|K(x+h,y)-K(x,y)\| \le C\,\frac{|h|^\alpha}{|x-y|^{n+\alpha}} \qquad \|K(x,y+h)-K(x,y)\| \le C\,\frac{|h|^\alpha}{|x-y|^{n+\alpha}}$$

provided $|x-y| \ge h$.

Analogously, we may define the Littlewood-Paley condition for vector-valued functions similarly by replacing $|\cdot|$ with the Hilbert space operator norm $\|\cdot\|$. As parallel to the C-Z condition, it might sometimes be convenient also to assume continuity on the first variable and that ψ_t averages to 0, i.e. $\Theta_t 1 = 0$.
The following theorem is a version of the Calderon-Zygmund theorem (Theorem 5.10) for vector-valued functions.

Theorem 8.7. *Suppose T is a vector-valued SIO and $T : L^2(\mathbb{R}^n, H_1) \to L^2(\mathbb{R}^n, H_2)$ is a bounded operator. Then T is of type $w(1,1)$ and also L^p-bounded for $1 < p < \infty$. Moreover, if the kernel $K(x,y)$ of T satisfies the C-Z size and smoothness condition in the second variable (but not necessarily the first), then T is of type $w(1,1)$ and also L^p-bounded for $1 < p \le 2$.*

Note that eventhough it is not a corollary of the aforementioned theorem, its proof will follow the same outline with appropriate modifications.
The point of using this result about vector-valued function is to obtain an extension of the Littlewood-Paley theory for L^p, where $p \ne 2$. Of course the case $p = 2$ has been dealt with in the previous section. This following theorem gives the extension result.

Theorem 8.8 (Littlewood-Paley for L^p)**.** *Let Q_t and Θ_t be defined as in the previous section with the associated function ψ. Then $g_\psi f$ and $S_\psi f$ are each of type $w(1,1)$ and also L^p-bounded for $1 < p \le 2$. Moreover, if $\psi_t(x,y)$ (defined as before) also satisfies the L-P smoothness condition on the first variable, then $g_\psi f$ and $S_\psi f$ are L^p-bounded for $1 < p < \infty$.*

Proof. In light of the vector-valued C-Z theorem, it suffices to cast the functions $f \mapsto g(f)$ and $f \mapsto S(f)$ as vector-valued SIO with kernels that satisfy the vector-valued C-Z condition. This may happen when we make appropriate choices for the Hilbert spaces H_1 and H_2. Indeed, let H_1 to be the complex space and

$$H_2 = \left\{ h(t) \,\middle|\, \int_0^\infty |h(t)|^2 \frac{dt}{t} < \infty \right\} = L^2\left((0,\infty) \,\middle|\, \frac{dt}{t}\right)$$
$$\widetilde{H_2} = \left\{ h(x,t) \,\middle|\, \int_0^\infty \int_{|z|<t} |h(z,t)|^2\,dz\,\frac{dt}{t^{n+1}} < \infty \right\}$$

We claim that the kernels associated with $g_\psi f$ and $S_\psi f$ are in $\mathcal{L}(\mathbb{C}, H_2)$ and $\mathcal{L}(\mathbb{C}, \widetilde{H_2})$, respectively.

We begin with the square vertical function. For any given $f : \mathbb{R}^n \to \mathbb{C}$, set $Tf(x) = \{\Theta_t f(x)\}_{t>0}$, i.e. $(Tf(x))(t) = \Theta_t f(x)$. Obsever that $Tf(x)$ belongs to H_2. Indeed,

$$\|Tf(x)\|_{H_2} = \left(\int_0^\infty |\Theta_t f(x)|^2 \frac{dt}{t}\right)^{1/2} = gf(x)$$

However, Theorem 8.4 gives that $g(f)$ is in L^2, seen as a function of x. Thus $Tf(x)$ is finite a.e., i.e. $Tf(x) \in H_2$ for a.e. x.
Observe here that $\psi_t(x,y) \in H_2$ for each fixed x, y. Morever, this means we may associate H_2 with the operator space $\mathcal{L}(\mathbb{C}, H_2)$ (to see this, note that $f(y) \in \mathbb{C}$, so $\psi_t(x,y) f(y) \in H_2$ iff $\psi_t(x,y) \in H_2$, where in the first, $\psi_t(x,y)$ is seen as operator from $\mathbb{C}$ to H_2). This means, $\Theta_t f(x) = \int_{\mathbb{R}^n} \psi_t(x,y) f(y)\, dy$ belongs to H_2 with kernel $K(x,y)(t) = \psi_t(x,y)$.
Now we'll show that $K(x,y)$ satisfies the vector-valued C-Z kernel condition. We begin with the size condition: we may write

$$\|K(x,y)\|_{H_2}^2 = \underbrace{\int_0^{|x-y|} |\psi_t(x,y)|^2 \frac{dt}{t}}_{\text{I}} + \underbrace{\int_{|x-y|}^\infty |\psi_t(x,y)|^2 \frac{dt}{t}}_{\text{II}}$$

First we bound I:

$$\begin{aligned}
\int_0^{|x-y|} |\psi_t(x,y)|^2 \frac{dt}{t} &\lesssim \int_0^{|x-y|} \left(\frac{t^\alpha}{(t+|x-y|)^{n+\alpha}}\right)^2 \frac{dt}{t} && \text{(L-P condition on } \psi_t) \\
&\lesssim \frac{1}{|x-y|^{2(n+\alpha)}} \int_0^{|x-y|} t^{2\alpha-1}\, dt && (|x-y| \text{ dominates } t \text{ in denominator}) \\
&\approx \frac{1}{|x-y|^{2n}} && \text{(Evaluate the integral)}
\end{aligned}$$

Next for II,

$$\begin{aligned}
\int_{|x-y|}^\infty |\psi_t(x,y)|^2 \frac{dt}{t} &\lesssim \int_{|x-y|}^\infty \left(\frac{t^\alpha}{(t+|x-y|)^{n+\alpha}}\right)^2 \frac{dt}{t} && \text{(L-P condition on } \psi_t) \\
&\lesssim \int_{|x-y|}^\infty t^{-2n-1}\, dt && (t \text{ dominates } |x-y| \text{ in denominator}) \\
&\approx \frac{1}{|x-y|^{2n}}
\end{aligned}$$

Combining I and II and taking square root will give the size condition for C-Z kernel.
Now we try to obtain the Hölder continuity condition. Suppose $|x-y| \geq 2|h|$. We'll first show continuity on the second variable.

$$\|K(x,y+h) - K(x,y)\|_{H_2}^2 = \underbrace{\int_0^{|h|} |\psi_t(x,y+h) - \psi_t(x,y)|^2 \frac{dt}{t}}_{\text{I}} + \underbrace{\int_{|h|}^\infty |\psi_t(x,y+h) - \psi_t(x,y)|^2 \frac{dt}{t}}_{\text{II}}$$

Again, we'll bound them separately. First for I,

$$\begin{aligned} \mathrm{I} &\lesssim \int_0^{|h|} \left(\frac{t^\alpha}{(t+|x-y-h|)^{n+\alpha}} \right)^2 \frac{dt}{t} \\ &\quad + \int_0^{|h|} \left(\frac{t^\alpha}{(t+|x-y|)^{n+\alpha}} \right)^2 \frac{dt}{t} && \begin{pmatrix}\text{L-P condition on } \psi_t \\ \text{and } \triangle \text{ inequality}\end{pmatrix} \\ &\lesssim \left(\frac{1}{|x-y-h|^{2(n+\alpha)}} + \frac{1}{|x-y|^{2(n+\alpha)}} \right) \int_0^{|h|} t^{2\alpha-1}\,dt && (|x-y| \text{ dominates } t) \\ &\approx \frac{1}{|x-y|^{2(n+\alpha)}} |h|^{2\alpha} && \begin{pmatrix}|h| \text{ is small, so} \\ |x-y| \approx |x-y-h|\end{pmatrix} \end{aligned}$$

Now for II,

$$\begin{aligned} \mathrm{II} &\lesssim \int_{|h|}^{\infty} \left(\frac{|h|^\alpha}{(t+|x-y|)^{n+\alpha}} \right)^2 \frac{dt}{t} && \begin{pmatrix}\text{L-P Hölder condition and} \\ |x-y| \approx |x-y-h|\end{pmatrix} \\ &\lesssim \frac{|h|^{2\alpha}}{|x-y|^{2n+\alpha}} \int_{|h|}^{\infty} t^{-\alpha-1}\,dt && \begin{pmatrix}\text{Split denominator to } t^{\alpha/2} \\ \text{and } |x-y|^{n+\alpha/2}\end{pmatrix} \\ &\approx \frac{|h|^{\alpha}}{|x-y|^{2n+\alpha}} \end{aligned}$$

Note here that taking square root will give the bound to be on the order $\frac{|h|^{\alpha/2}}{|x-y|^{n+\alpha/2}}$. This is acceptable for the C-Z condition, as we may choose to let the exponent to be $\alpha/2$ in place of α (alternatively, choose the highest between bound of I and II).
Continuity on the first variable is shown in same way. Here we'll assume that L-P condition gives Hölder continuity on the first variable. This will give us the C-Z continuity condition and hence our proof for $g_\psi f$ is complete.
Now we'll repeat the same proof for $S_\psi f$. By change of variable with $z = y - x$, we can write

$$Sf(x) = \left(\int_0^\infty \int_{|z|<t} |\Theta_t f(x+z)|^2\,dz\,\frac{dt}{t^{n+1}} \right)^{1/2}$$

As before, we may view the kernel $\psi_t(x+z, y)$ for fixed $x, y \in \mathbb{R}^n$ to take values in $\widetilde{H_2}$, so $\Theta_t f(x+z)$ defines an element of $\mathcal{L}(\mathbb{C}, \widetilde{H_2})$, which again can be identified with $\widetilde{H_2}$. Thus the vector-valued SIO is associated with vector-valued kernel $K(x,y)(z,t) = \psi_t(x+z, y)$ (note here that z is not related with $x - y$ again).
Now we're ready to check the conditions for C-Z kernel. We'll begin with the size condition.

$$\begin{aligned} \|K(x,y)\|_{\widetilde{H_2}}^2 &= \int_0^\infty \int_{|z|<t} |\psi_t(x+z,y)|^2\,dz\,\frac{dt}{t^{n+1}} \\ &= \underbrace{\int_0^{|x-y|/2} \int_{|z|<t} |\psi_t(x+z,y)|^2\,dz\,\frac{dt}{t^{n+1}}}_{\mathrm{I}} + \underbrace{\int_{|x-y|/2}^{\infty} \int_{|z|<t} |\psi_t(x+z,y)|^2\,dz\,\frac{dt}{t^{n+1}}}_{\mathrm{II}} \end{aligned}$$

Observe that in I, $|z|$ is small compared to $|x-y|$ (in fact $|z| < \frac{|x-y|}{2}$), so we may $|x+z-y| \approx |x-y|$. Hence

$$\begin{aligned}
\mathrm{I} &\lesssim \int_0^{|x-y|/2} \int_{|z|<t} \left(\frac{t^\alpha}{(t+|x+z-y|)^{n+\alpha}} \right)^2 dz \, \frac{dt}{t^{n+1}} && \text{(L-P (i) for } \psi_t) \\
&\lesssim \frac{1}{|x-y|^{2(n+\alpha)}} \int_0^{|x-y|/2} t^{2\alpha-1} \cdot \underbrace{t^{-n} \int_{|z|<t} dz}_{\text{Constant size of } B_t(0)} \, dt && (|x-y| \approx |x+z-y|) \\
&\approx |x-y|^{-2n}
\end{aligned}$$

On the other hand, for II, t dominates $|x+z-y|$, so

$$\begin{aligned}
\mathrm{II} &\lesssim \int_{|x-y|/2}^{\infty} \int_{|z|<t} \left(\frac{t^\alpha}{(t+|x+z-y|)^{n+\alpha}} \right)^2 dz \, \frac{dt}{t^{n+1}} && \text{(L-P(i) for } \psi_t) \\
&\lesssim \int_{|x-y|/2}^{\infty} t^{-2n-1} \cdot \underbrace{t^{-n} \int_{|z|<t} dz}_{\text{Constant size of } B_t(0)} \, dt && \begin{pmatrix} \text{Remove } |x+z-y| \text{ from} \\ \text{the denominator} \end{pmatrix} \\
&\approx |x-y|^{-2n}
\end{aligned}$$

Combining I and II and taking square root will give the C-Z size condition.
Now we'll show C-Z continuity for ψ_t. As before, it suffices to show this for the y-variable, as the on for x follows exactly the same proof. We may write

$$\begin{aligned}
\|K(x,y+h) - K(x,y)\|_{H_2}^2 &= \underbrace{\int_0^{|h|} \int_{|z|<t} |\psi_t(x+z, y+h) - \psi_t(x+z, y)|^2 \frac{dt}{t}}_{\mathrm{I}} \\
&\quad + \underbrace{\int_{|h|}^{\infty} \int_{|z|<t} |\psi_t(x+z, y+h) - \psi_t(x+z, y)|^2 \frac{dt}{t}}_{\mathrm{II}}
\end{aligned}$$

Suppose $|x-y| \geq 2|h|$. Then $\frac{1}{2}|x-y| \leq |x-y-h| \leq 3|x-y|$, i.e. $|x-y-h| \approx |x-y|$. Then we can bound I as follows: by L-P size condition on ψ_t,

$$\begin{aligned}
\mathrm{I} &\lesssim \int_0^{|h|} \int_{|z|<t} \left(\frac{t^{2\alpha}}{(t+|x+z-y-h|)^{2(n+\alpha)}} + \frac{t^{2\alpha}}{(t+|x+z-y|)^{2(n+\alpha)}} \right) dz \, \frac{dt}{t^{n+1}} && \begin{pmatrix} \text{L-P size condi-} \\ \text{tion on each } \psi_t \end{pmatrix} \\
&\lesssim \int_0^{|h|} t^{2\alpha-1} \cdot t^{-n} \int_{|z|<t} \left(\frac{1}{|x+z-y-h|^{2(n+\alpha)}} + \frac{1}{|x+z-y|^{2(n+\alpha)}} \right) dz \, dt && \begin{pmatrix} \text{Remove } t \text{ from} \\ \text{the denominator} \end{pmatrix} \\
&\lesssim \frac{1}{|x-y-h|^{2(n+\alpha)}} \int_0^{|h|} t^{2\alpha-1} \cdot \underbrace{t^{-n} \int_{|z|<t} dz}_{\text{Volume constant of } B_t(0)} \, dt && (|z| < |h|)
\end{aligned}$$

$$\approx \frac{|h|^{2\alpha}}{|x-y|^{2(n+\alpha)}} \qquad (|x-y-h| \approx |x-y|)$$

For II, because $|h| \leq t$, we're going to use the L-P continuity condition.

$$\text{II} \lesssim \int_{|h|}^{\infty} \int_{|z|<t} \left(\frac{|h|^{\alpha}}{(t+|x+z-y|)^{n+\alpha}} \right)^2 dz\, \frac{dt}{t^{n+1}} \qquad \text{(L-P continuity for } \psi_t\text{)}$$

Here we stop and consider two possible cases. Observe that if $|h| \leq t \leq \dfrac{|x-y|}{2}$, then $|x+z-y| \approx |x-y|$ (because $|z| < |x-y|/2$) and thus we can approximate part of above as

$$\frac{|h|^{2\alpha}}{|x-y|^{2n+\alpha}} \int_{|h|}^{|x-y|/2} t^{-\alpha-1} \cdot \underbrace{t^{-n} \int_{|z|<t} dz}_{\text{Volume constant of } B_t(0)} \; dt = \frac{|h|^{\alpha}}{|x-y|^{2n+\alpha}} \tag{18}$$

Again here we may take the exponent to be $\beta/2$. Now for the remaining part,

$$|h|^{2\alpha} \int_{|x-y|/2}^{\infty} \int_{|z|<t} \frac{1}{(t+|x+z-y|)^{2(n+\alpha)}}\, dz\, dt \leq |h|^{2\alpha} \int_{|x-y|/2}^{\infty} t^{-2n-2\alpha-1} \cdot \underbrace{t^{-n} \int_{|z|<t} dz}_{\text{Volume constant of } B_t(0)} \; dt$$

$$\lesssim \frac{|h|^{2\alpha}}{|x-y|^{2(n+\alpha)}}$$

Compare the two parts to choose the appropriatae exponent.

□

8.2 Discrete Littlewood-Paley Theory

Recall that in the convolution case, $Q_t f = \psi_t * f$ with ψ_t is a average-zero bump function. In fact, if ψ is such function, then we know that $\widehat{\psi} \in \mathcal{S}$, radial, vanishes at the origin, and $\widehat{\psi}(\xi) \sim \min(|\xi|, |\xi|^{-n})$. Because of the rapid decay of $\widehat{\psi}$, we can deduce that most of its 'mass' is concentrated on some shell. We will translate this idea to the discrete case.

Suppose $\psi \in \mathcal{S}$, $\psi_j(x) = 2^{jn}\psi(2^j x)$, and ψ have average zero. We'll obtain this function by defining $\psi = \check{\varphi}$, where $\varphi \in C_c^{\infty}(1/4, 4)$, and $\sum_{j\in\mathbb{Z}} \varphi_j^2(\xi) = 1$.

Now we need to build such φ. To do so, we'll begin with $\widetilde{\varphi} \in C_c^{\infty}(1/4 < |\xi| < 4)$, so that $0 \leq \widetilde{\varphi} \leq 1$, $\widetilde{\varphi}$ is radial, and $\widetilde{\varphi} \equiv 1$ on $\{3/8 \leq |\xi| \leq 3\}$. Define then for $\xi \neq 0$

$$S(\xi) = \sum_{j\in\mathbb{Z}} \widetilde{\varphi}^2(2^{-j}\xi)$$

Observe here that for each ξ, there are only finitely many annuli that contains ξ, so $S(\xi)$ is always finite. On the other hand, because $\widetilde{\varphi}$ is continuous on a compact set (and because each ξ is contained in at least one dyadic annulus), $S(\xi) > 0$. Thus, we may define

$$\varphi_j(\xi) = \frac{\widetilde{\varphi}^2(2^{-j}\xi)}{S(\xi)} \qquad \varphi(\xi) = \frac{\widetilde{\varphi}(\xi)}{S(\xi)}$$

and hence $\varphi_j(\xi) = \varphi(2^{-j}\xi)$. It's easy to check that ψ will then have the appropriate property (and by FT property, ψ_j will have the correct constant multiplier).
In this setting, define $Q_j f := \psi_j * f$. Following the same outline as in the continuous case, we can obtain similar result:

Theorem 8.9. *(a) Calderon reproducing formula:* $\sum_{j\in\mathbb{Z}} Q_j^2 = -Id$ *in* L^2*-operator topology.*

(b) Define discrete L-P square function by $Sf(x) := \left(\sum_{j\in\mathbb{Z}} |Q_j f(x)|^2\right)^{1/2}$. *Then* S *is* L^2*-bounded and* $\|Sf\|_2 = \|f\|_2$.

(c) The map $f \mapsto Sf$ *is of* $w(1,1)$*-type and* L^p*-bounded for* $1 < p < \infty$.

The last thing we will discuss is this fact: not only the square functions are L^p-bounded, their L^p-norms are also equivalent to the L^p-norm of the function f. We have shown the upper bounded via vector-valued C-Z theorem, so it now remains to show the lower bound.

Proposition 8.10. *Set* $Q_t f = \psi_t * f$, *where* ψ_t *is the usual bump function that satisfies the L-P condition (but* ψ_t *is not necessarily radial). Suppose also that* Q_t *satisfies the Calderon Reproducing Formula (Proposition 8.1). Then for* $1 < p < \infty$

$$\|f\|_p \le C \left\|g_\psi f\right\|_p \qquad \|f\|_p \le C \left\|S_\psi f\right\|_p$$

Proof. We have shown that for $p = 2$, this is an equality (Proposition 8.2). It then remains to show the bound for other values of p. Choose $h \in L^{p'}$ so that $\|h\|_{p'} = 1$. In light of the functional definition of the L^p-norm, it suffices to show that

$$\left|\int_{\mathbb{R}^n} f\, h\right| \le C\ \|g(f)\|_p$$

as one can then take supremum over all such h to obtain $\|f\|_p$. In fact, one may reduce the case further to consider only on the dense cases $f \in L^p \cap L^2$ and $h \in L^{p'} \cap L^2$. By reproducing formula,

$$\begin{aligned}\int_{\mathbb{R}^n} f\, h &= \int_{\mathbb{R}^n}\int_0^\infty Q_t^2 f\,\frac{dt}{t}\,h\,dx \\ &= \int_0^\infty \int_{\mathbb{R}^n} Q_t^2 f\,h\,dx\,\frac{dt}{t} && \text{(Tonelli)}\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \int_{\mathbb{R}^n} Q_t f \cdot Q_t^* h \, dx \, \frac{dt}{t} && \begin{pmatrix}\text{View } x\text{-integral as} \\ L^2\text{-inner product}\end{pmatrix} \\
&= \int_{\mathbb{R}^n} \int_0^\infty Q_t f \cdot Q_t^* h \, \frac{dt}{t} \, dx && \text{(Fubini)} \\
&\leq \int_{\mathbb{R}^n} \left(\int_0^\infty |Q_t f|^2 \, \frac{dt}{t} \right)^{1/2} \left(\int_0^\infty |Q_t h|^2 \, \frac{dt}{t} \right)^{1/2} dx && \begin{pmatrix}\text{Cauchy-Schwarz} \\ \text{on } t\text{-integral}\end{pmatrix} \\
&\leq \left\| g_\psi f \right\|_p \left\| g_{\widetilde{\psi}} h \right\|_{p'} && \begin{pmatrix}\text{Hölder on } x \\ Q_t \, *\end{pmatrix} \\
&\leq \left\| g_\psi f \right\|_p \left\| g_{\widetilde{\psi}} h \right\|_{p'} \\
&\leq C \left\| g_\psi f \right\|_p && \begin{pmatrix} g_\psi \text{ is bounded} \\ \text{and } \|h\|_{p'} = 1\end{pmatrix}
\end{aligned}
$$

For $S_\psi f$ we cam argue similarly to get

$$
\begin{aligned}
\int_{\mathbb{R}^n} f \, h &= \int_0^\infty \int_{\mathbb{R}^n} Q_t f(y) \cdot Q_t^* h(y) \, \frac{dt}{t} \, dy \\
&= \int_0^\infty \int_{\mathbb{R}^n} Q_t f(y) \cdot Q_t^* h(y) \, dy \fint_{|x-y|<t} dx \, \frac{dt}{t} && \begin{pmatrix}\text{The average integral} \\ \text{evaluates to 1}\end{pmatrix} \\
&= \underbrace{C_n}_{\text{Volume constant}} \int_{\mathbb{R}^n} \int_0^\infty \int_{|x-y|<t} Q_t f(y) \, Q_t^* h(y) \, dy \, \frac{dt}{t^{n+1}} \, dx && \text{(Fubini)} \\
&\lesssim \int_{\mathbb{R}^n} \left(\int_0^\infty \int_{|x-y|<t} |Q_t f(y)|^2 \, dy \, \frac{dt}{t^{n+1}} \right)^{1/2} && \begin{pmatrix}\text{Cauchy-Schwarz for} \\ \text{cone integral}\end{pmatrix} \\
&\qquad \left(\int_0^\infty \int_{|x-y|<t} |Q_t^* h(y)|^2 \, dy \, \frac{dt}{t^{n+1}} \right)^{1/2} dx \\
&\lesssim \left\| S_\psi f \right\|_p \left\| S_\psi h \right\|_{p'} && \begin{pmatrix}\text{Hölder on the} \\ x\text{-integral}\end{pmatrix} \\
&< C \left\| S_\psi f \right\|_p && (S_\psi \text{ is bounded})
\end{aligned}
$$

In both cases, taking supremum over all such h will give the wanted result.

□

9 Bounded Mean Oscillator (BMO)

Define the space of bounded mean oscillator (BMO) to be the space of equivalence classes of functions $f \in L^1_{\text{loc}}$, modulo constants, so that the BMO norm defined as

$$
\|f\|_* := \sup_Q \fint_Q (f(x) - f_Q) \, dx
$$

is finite. Here Q is a cube and f_Q is the average of f in Q.
Observe here that this space is non-empty. For example, consider the space of bounded functions modulo constant L^∞/C with norm

$$\|f\|_{L^\infty/C} = \inf_{C\in\mathbb{R}^n} \|f - C\|_\infty$$

Then this space is naturally contained in BMO. On the other hand, this inclusion is strict. In fact, we'll show later that the typical element of BMO has logarithm growth (and hence the space contains more than bounded functions).
Also note that if C_Q is a constant,

$$\inf_{C_Q} \fint_Q |f - C_Q| \le \fint_Q |f - f_Q|$$

On the other hand

$$\fint_Q |f - f_Q| = \fint_Q [(f - C_Q) - (f_Q - C_Q)] \le 2 \fint_Q |f - C_Q|$$

and hence

$$\fint_Q |f - f_Q| \le 2 \inf_{C_Q} \fint_Q |f - C_Q|$$

Thus we may give an equivalent definition for the BMO norm:

$$\|f\|_* \approx \sup_Q \inf_{C_Q} \fint |f(x) - C_Q|\, dx$$

This equivalence definition makes clear the need to consider the space of functions modulo constant: if f is constant, it is equal to its average and thus $\|f\|_* = 0$. To make the BMO norm to be a true norm, we are forced to consider only when the constant is zero, hence taking congruence.
The following lemma gives the fundamental fact about BMO:

Lemma 9.1 (John-Nirenberg)**.** *Let $f \in BMO(\mathbb{R}^n)$. Then there exist positive constants C_1 and C_2, dependent only on n, so that for all $\lambda > 0$ and cubes Q*

$$\left|\left\{x \in Q \,\middle|\, |f(x) - f_Q| > \lambda\right\}\right| \le C_1 \exp\left(-\frac{C_2\lambda}{\|f\|_*}\right) \cdot |Q|$$

Proof. Define

$$T := \sup_Q \fint_Q \exp\left(\frac{C_2}{\|f\|_*}|f(x) - f_Q|\right) dx$$

We claim that for appropriate choice of C_2, T is bounded by C_1 for some $C_1 > 0$ and the lemma will then follow. Indeed, once we've shown that T is bounded, we can write for each cube Q

$$C_1|Q| \ge \int_Q \exp\left(\frac{C_2}{\|f\|_*}|f(x) - f_Q|\right) dx \qquad \text{(Definition of } T\text{)}$$

$$\geq A^{\lambda}\left|\left\{\exp\left(\frac{C_2}{\|f\|_*}|f(x)-f_Q|\right) > A^{\lambda}\right\}\right| \qquad \text{(Markov for any } A>0\text{)}$$
$$= \exp\left(\frac{C_2\lambda}{\|f\|_*}\right)|\{|f(x)-f_Q|>\lambda\}| \qquad \left(\text{Set } A=\exp\left(\frac{C_2}{\|f\|_*}\right)\right)$$

Dividing by the exponent on the last line on both sides will give the result. Then it remains to show that T is bounded for appropriate constants $C_1, C_2 > 0$ (which will be the constants in the lemma). First we consider the case when f is bounded, so T is finite (by $e^{2C_2\|f\|_\infty}$). We'll find the appropriate constants here that works uniformly for all bounded f, so by extension it works for general f. By supremum property, we may choose a cube Q so that

$$⨍_Q \exp\left(\frac{C_2}{\|f\|_*}|f(x)-f_Q|\right)dx \geq \frac{1}{2}T \tag{19}$$

If we assume, WLOG, that $\|f\|_* = 1$, we may fix $t > 1$ (which will be determined later) so that $\|f\|_* = 1 < t$. Moreover, this means

$$⨍_Q |f(x)-f_Q|\,dx \leq 1 < t \tag{20}$$

Now we'll apply the stopping time argument. Starting from Q, we'll divide keep dividing it into 2^n equally-sized cubes until we hit a cube that violates (20). In other words, we'll obtain a collection of cubes $\{Q_j\}_{j=1}^{\infty}$ so that

$$⨍_{Q_j} |f(x)-f_Q|\,dx \geq t \tag{21}$$

Note that by Lebesgue Differentiation Theorem (Theorem 1.3), (20) and (21) implies

$$|f(x)-f_Q| \leq t \qquad \text{for} \qquad x \in Q \setminus \left(\bigcup_{j=1}^{\infty} Q_j\right)$$

Moreover, we can estimate the size of the bad sets:

$$\left|\bigcup_{j=1}^{\infty} Q_j\right| = \sum_j |Q_j| \leq \frac{1}{t}\sum_{j=1}^{\infty}\int_{Q_j}|f(x)-f_Q|\,dx = \frac{1}{t}\int_{\bigcup_{j=1}^{\infty}Q_j}|f(x)-f_Q|\,dx$$

If we bound the last integral crudely by extending the domain to Q, we may conclude that

$$\left|\bigcup_{j=1}^{\infty} Q_j\right| \leq \frac{|Q|}{t}⨍_Q |f(x)-f_Q|\,dx \leq \frac{|Q|}{t} \tag{22}$$

because the integral is bounded by the BMO norm. Some more observations: if Q_j* is the parent of Q_j, then maximality of the 'bad' cubes gives

$$⨍_{Q_j*} |f(x)-f_Q| \leq t \tag{23}$$

Now note that for each Q_j, we have

$$\begin{aligned}
|f_{Q_j} - f_{Q_j^*}| &\le \fint_{Q_j} |f(x) - f_{Q_j^*}|\,dx && \text{(Definition of } f_{Q_j}\text{)} \\
&\le 2^n \fint_{Q_j^*} |f(x) - f_{Q_j^*}|\,dx && (|Q_j^*| = 2^n|Q_j|) \\
&\le 2^n \|f\|_* < 2^n \cdot t
\end{aligned}$$

Similarly, by application of (23)

$$|f_{Q_j^*} - f_Q| \le \fint_{Q_j^*} |f(x) - f_Q|\,dx \le t$$

Combining these two results,

$$|f_{Q_j} - f_Q| \le |f_{Q_j} - f_{Q_j^*}| + |f_{Q_j^*} - f_Q| \le (2^n + 1)t \tag{24}$$

Now we're ready to bound T. From (19),

$$\frac{1}{2}T \le \underbrace{\frac{1}{|Q|}\int_{Q\setminus\cup_{j=1}^\infty Q_j} e^{C_2 t}\,dx}_{\text{I}} + \underbrace{\frac{1}{|Q|}\sum_{j=1}^\infty \int_{Q_j} \exp(C_2|f(x) - f_Q|)\,dx}_{\text{II}}$$

Bounding the 'good' part (I) is easy: because the integrand is constant,

$$\frac{1}{|Q|}\int_{Q\setminus\cup_{j=1}^\infty Q_j} e^{C_2 t}\,dx \le \fint_Q e^{C_2 t}\,dx = e^{C_2 t}$$

On the other hand, note that by (24)

$$|f(x) - f_Q| \le |f(x) - f_{Q_j}| + |f_{Q_j} - f_Q| \le |f(x) - f_{Q_j}| + (2^n + 1)t$$

Thus,

$$\begin{aligned}
\frac{1}{|Q|}\sum_{j=1}^\infty \int_{Q_j} \exp(C_2|f(x) - f_Q|)\,dx &\le \frac{1}{|Q|}\sum_{j=1}^\infty |Q_j| e^{C_2(2^n+1)t} \underbrace{\fint_{Q_j} \exp(C_2|f(x) - f_{Q_j}|)\,dx}_{T} \\
&\le \frac{1}{|Q|} e^{C_2(2^n+1)t}\, T \sum_{j=1}^\infty |Q_j| \\
&\le e^{C_2(2^n+1)t} \cdot \frac{T}{t}
\end{aligned}$$

where we use (22) in the last line. Now we'll determine the constants. Set $t = 16$ (the discussion so far holds for any $t > 1$) and choose C_2 small enough so that $\exp[C_2(2^n + 1)16] \leq 4$. Then the expression above is bounded by $T/4$. Putting everything together, we conclude from (19)

$$\frac{1}{2}T \leq e^{16C_2} + \frac{1}{4}T$$

Thus letting $C_1 = 4e^{16C_2}$ will give the wanted bound for T.
Now for the general case, we may replace f by f_N given by

$$f_N = \begin{cases} f & |f| \leq N \\ N & |f| > N \end{cases}$$

□

As a corollary, we present another result that's also often called the John-Nirenberg lemma:

Corollary 9.2 (Also John-Nirenberg)**.** *Let* $1 < p < \infty$. *Then*

$$\|f\|_{BMO_p} := \sup_Q \left(\fint_Q |f(x) - f_Q|\, dx \right)^{1/p} \leq C_{p,n}\ \|f\|_*$$

Proof. Observe that on one direction, Jensen inequality allows us to write

$$\fint_Q |f(x) - f_Q|\, dx \leq \left(\fint_Q |f(x) - f_Q|^p\, dx \right)^{1/p}$$

So $\|f\|_* \leq \|f\|_{\mathrm{BMO}_p}$. On the other hand,

$$\begin{aligned} \fint_Q |f(x) - f_Q|\, dx &= \frac{1}{Q} p \int_0^\infty \lambda^{p-1} |\{|f(x) - f_Q| > \lambda\}|\, d\lambda && \text{(Distributional defintion)} \\ &< \frac{p\, C_1}{|Q|} \int_0^\infty \lambda^{p-1} \exp\left(-\frac{C_2 \lambda}{\|f\|_*}\right) d\lambda && \text{(Lemma 9.1)} \\ &= \frac{C_1\, p}{|Q|}\ \|f\|_*^p \int_0^\infty \lambda^{p-1} e^{-C_2\lambda}\, d\lambda && (\lambda \mapsto \lambda\, \|f\|_*) \\ &= C\, \|f\|_*^p \end{aligned}$$

Take the supremum over all Q and take p-th root to get the wanted result.

□

Note. Eventhough the lemma and the corollary above omits the case when $p = 1$, we can obtain another characterization of BMO, via the proof of the corollary, if we have an analogue of Lemma 9.1. To be precise, if for all cubes Q we have for some $\beta > 0$ the weak bound

$$|\{|f(x) - f_Q| > \lambda\}| \leq C\, e^{-\lambda/\beta}$$

then $f \in$ BMO and $\|f\|_* \leq C\beta$.

There's another topic that we may analyze using the space BMO. Consider a measure on $\mathbb{R}^{n+1}$ that doesn't vanish on the boundary (as subset of $\mathbb{R}^n$). This will motivate our discussion of the Carleson measure. For more complete discussion, see [2] p. 158.

Definition 9.3. A non-negative Borel measure μ defined on $\mathbb{R}^{n+1}_+ = \mathbb{R}^n \times (0,\infty)$ is a Carleson measure if there exists a constant $C_0 < \infty$ so that for all cubes $Q \subseteq \mathbb{R}^n$, we have

$$\mu(R_Q) \le C_0\,|Q| \qquad \text{where} \qquad R_Q := Q \times (0, \ell(Q))$$

The best constant C_0 is sometimes denoted by $\|\mu\|_C$, the Carleson norm. Alternatively we may define the norm as

$$\|\mu\|_C = \sup_Q \frac{\mu(R_Q)}{|Q|}$$

The following proposition is the fundamental characterization of the Carleson measure with respect to BMO.

Proposition 9.4 (Fefferman-Stein)**.** *Suppose $\Theta_t f(x) = \int_{\mathbb{R}^n} \psi_t(x,y)\, f(y)\, dy$, where $\{\psi_t\}_{t=0}^\infty$ satisfies the L-P conditions (i.e. it belongs to the L-P class) with $\Theta_t 1 = 0$ for all t (in particular, g_ψ is L^2-bounded). Then for $b \in BMO$,*

$$d\mu(x,t) := |\Theta_t b(x)|^2\, dx\, \frac{dt}{t}$$

is a Carleson measure and $\|\mu\|_C \le C \cdot \|b\|_*^2$ where C is a constant depending on n and the L-P condition.

Proof. We want to show that for any cube Q,

$$\frac{1}{|Q|} \iint_{R_Q} |\Theta_t b(x)|^2\, dx\, \frac{dt}{t} \le C \cdot \|b\|_*^2$$

uniformly in Q. Now observer that because $\Theta_t 1 = 0$, we may replace b by its constant shift and the wanted result will not change. In this view, set $\widetilde{b} = b - b_{2Q}$. Furthermore, set $b_0 = \widetilde{b} \cdot \chi_{2Q}$ and $b_k = \widetilde{b} \cdot \chi_{2^{k+1}Q \setminus 2^k Q}$. This way, we decomposes $\widetilde{b}$ so that $\widetilde{b} = \sum_{k=0}^\infty b_k$.

By Minkowski inequality, if we decompose b,

$$\left(\frac{1}{|Q|} \iint_{R_Q} |\Theta_t b(x)|^2\, dx\, \frac{dt}{t} \right)^{1/2} \le \sum_{k=0}^\infty \underbrace{\left(\frac{1}{|Q|} \iint_{R_Q} |\Theta_t b_k(x)|^2\, dx\, \frac{dt}{t} \right)^{1/2}}_{I_k} \tag{25}$$

Bounding the zeroth term is easy. Because g_ψ is L^2-bounded,

$$I_0 \le \left(\frac{1}{|Q|} \int_{\mathbb{R}^n} \int_0^\infty |\Theta_t b_0(x)|^2\, \frac{dt}{t}\, dx \right)^{1/2} \qquad \text{(Extend the domain)}$$

$$\lesssim \left(\frac{1}{|Q|}\int_{\mathbb{R}^n} |b(x) - b_{2Q}|^2 \cdot \chi_{2Q}\, dx\right)^{1/2} \qquad (g_\psi \text{ (t-integral) is } L^2\text{-bounded})$$

$$\lesssim \|b\|_* \qquad \text{(Corollary 9.2)}$$

For the remaining part integrals, we will show that there exists an exponent $\gamma > 0$ so that $I_k \leq C \cdot 2^{-k\gamma}\ \|b\|_*$ for all $k \geq 1$. We begin by recalling the fact the definition

$$\Theta_t b_k(x) = \int_{2^{k+1}Q \setminus 2^k Q} \psi_t(x, y)[b(y) - b_{2Q}]\, dy$$

Now, for any $x \in Q$ and $y \in 2^{k+1}Q \setminus 2^k Q$, the distance $|x - y|$ is approximately $2^k \ell(Q)$. Thus, since ψ_t is of class L-P, applying the size condition to the expression above will gives us

$$|\Theta_t b_k(x)| \lesssim \int_{\substack{\{|x-y|\approx 2^k \ell(Q)\} \\ \cap 2^{k+1}Q}} \frac{t^\alpha}{(t + |x - y|)^{n+\alpha}} |b(y) - b_{2Q}|\, dy$$

$$\lesssim \left(\frac{t}{2\ell(Q)}\right)^\alpha ⨍_{2^{k+1}Q} |b(y) - b_{2Q}|\, dy \qquad \text{(Split the denominator)}$$

Observe that we may split the integrand above into dyadic 'steps', e.g.

$$|b(y) - b_{2Q}| \leq |b(y) - b_{2^{k+1}Q}| + \sum_{l=1}^{k} |b_{2^{l+1}Q} - b_{2^l Q}|$$

Observe that for any cube P,

$$|b_{2P} - b_P| = \left|⨍_P (b(y) - b_{2P})\, dy\right| \leq 2^n \left|⨍_{2P} (b(y) - b_{2P})\, dy\right| \leq 2^n \|f\|_*$$

where the last inequality is by definition of BMO. Thus,

$$|\Theta_t b_k(x)| \lesssim \left(\frac{t}{2^k \ell(Q)}\right)^\alpha \left[⨍_{2^{k+1}Q} |b(y) - b_{2^{k+1}Q}|\, dy + \sum_{l=1}^{k} |b_{2^{l+1}Q} - b_{2^l Q}|\right]$$

$$\lesssim \left(\frac{t}{2^k \ell(Q)}\right)^\alpha (1 + 2^n k)\, \|b\|_*$$

$$\lesssim \left(\frac{t}{\ell(Q)}\right)^\alpha 2^{-k\alpha}\, k\ \|b\|_*$$

and hence

$$(I_k)^2 \leq \int_0^{\ell(Q)} ⨍_Q \left(\frac{t}{\ell(Q)}\right)^{2\alpha} 2^{-2k\alpha}\, k^2\, dx\, \frac{dt}{t} \cdot \|b\|_*^2$$

$$= k^2\, 2^{-2k\alpha}\ \|b\|_*^2 ⨍_Q \int_0^{\ell(Q)} \left(\frac{t}{\ell(Q)}\right)^{2\alpha} \frac{dt}{t}\, dx \qquad \text{(Fubini)}$$

$$= \frac{1}{2\alpha} k^2 \, 2^{-k\alpha} \; \|b\|_*^2$$

Taking square root of both sides will give the wanted bound with $\gamma = \alpha$. Substituting this bound and the bound for I_0 to (25),

$$\left(\frac{1}{|Q|} \iint_{R_Q} |\Theta_t b(x)|^2 \, dx \, \frac{dt}{t} \right)^{1/2} \lesssim \|b\|_* + \sum_{k=1}^{\infty} k^2 \, 2^{-k\alpha} \; \|b\|_* \leq C \cdot \|b\|_*$$

where closer inspection will give C to be dependent only on α and ψ. Squaring both sides will give

$$\int_{R_Q} d\mu = \iint_{R_Q} |\Theta_t b(x)|^2 \, dx \, \frac{dt}{t} \leq C \cdot \|b\|_*^2 \; |Q|$$

and hence μ is indeed a Carleson measure which norm is bounded by BMO norm of b.

□

Note that the converse also holds: if $\|\mu\|_C$ is finite and b is 'non-degenerate' with appropriate decay, the function b must belong to BMO. See [2] Ch. 2 Sec. 4.2-4.4 for detail. We will come back to prove this later.
In turn, we can study more of the Carleson measures to get better understanding of BMO. For any given F on $\mathbb{R}^{n+1}_+$, we define

$$N_* F(x) = \sup_{\{|x-y|<t|\}} |F(y,t)| < \infty$$

and we define the class $\mathcal{N}$ to be the functions so that $N_* F \in L^1(\mathbb{R}^n)$. Note that here we might equivalently choose the aperture of the cone to be different than 1. It turns out that the space $\mathcal{N}$ is a dual of the space of Carleson measures $\mathcal{C}$. Since Proposition 9.4 gives that Carleson measures are associated with BMO, we can see there is a relation between BMO and the non-tangential maximal functions $\mathcal{N}$. This statement is made precise with the following result:

Lemma 9.5 (Carleson Embedding)**.** *Let μ be a Carleson measure. For all $0 < p < \infty$ and continuous $F \in \mathcal{N}$,*

$$\int_0^{\infty} \int_{\mathbb{R}^n} |F(y,t)|^p \, d\mu(y,t) \leq C_n \; \|\mu\|_C \int_{\mathbb{R}^n} (N_* F(x))^p \, dx$$

The proof of this lemma will rely on this following covering lemma:

Lemma 9.6 (Whitney Covering)**.** *Let $\Omega \subsetneq \mathbb{R}^n$, then there exists a collection $\mathcal{W} = \{Q_j\}_j$, where Q_j are non-overlapping closed dyadic cubes (so each is of size $c \cdot 2^k$ for some k) so that (i) $\Omega = \bigcup_{Q_j \in \mathcal{W}} Q_j$ and (ii) diam $Q_j \approx$ dist(Q_j, Ω^c) for all $Q_j \in \mathcal{W}$.*

Proof of Carleson Embedding Lemma. Because x^p is increasing, $N_* F^p = (N_* F)^p$, so it suffices to show the case when $p = 1$ as we can then apply this result to the function F^p. The proof is by showing F and $N_* F$ have the same distribution, i.e. for every $\lambda > 0$

$$\mu \left\{ (y,t) \in \mathbb{R}^{n+1}_+ \,\middle|\, |F(y,t)| > \lambda \right\} \leq C_n \; \|\mu\|_C \; |E_\lambda| \qquad \text{where } E_\lambda = \{x \in \mathbb{R}^n \mid N_* f(x)| > \lambda\}$$

Note E_λ is open: for any $x \in E_\lambda$, consider the cone Γ with vertex x in $\mathbb{R}^{n+1}_+$. This cone Γ is open, so for any $(y,t) \in \Gamma$ so that $|F(y,t)| > \lambda$ (exists because $N_* F(x) > \lambda$), there is a neighbourhood of (y,t) wholly contained in Γ so that $F > \lambda$ on this neighbourhood (because F is continuous). Moreover, if Γ' is another cone with vertex x', so long x' is sufficiently close, this neighbourhood of (y,t) will also be in Γ'. Thus, we have a whole neighbourhood of x so that $F > \lambda$ in this neighbourhood.
Let $\mathcal{W} = \{Q_j\}_j$ be the Whitney cover of E_λ. Given $(y,t) \in \mathbb{R}^{n+1}_+$ so that $|F(y,t)| > \lambda$, we must have $N_* F(x) > \lambda$ for the whole 'tent' $|x - y| < t$. This gives the first observation:

$$\{x \in \mathbb{R}^n \mid |x - y| < t\} \subseteq E_\lambda$$

In particular, if $y \in E_\lambda$, there is $Q_{j_0} \in \mathcal{W}$ so that $y \in Q_{j_0}$. From property of Whitney balls, $\operatorname{diam} Q_{j_0}$ is comparable to $\operatorname{dist}(Q_{j_0}, E_\lambda^c)$. In particular, $\operatorname{diam} Q_{j_0}$ is comparable to $\operatorname{dist}(y, E_\lambda^c)$, which in turn is at least t (otherwise we're back in cone $|x - y| < t$). Hence we get (i) $\ell(Q_{j_0}) \geq c\,t$ for some $c > 0$ and (ii) there exists an uniform constant K so that $(y,t) \in R_{KQ_{j_0}}$ (say take $K = 2/c$ for c as in (i)). This constant depends only on the Whitney balls).
Applying similar argument for each (y,t) and because the constant K is uniform,

$$\{(y,t) \mid |F(y,t)| > \lambda\} \subseteq \bigcup_{Q_j \in \mathcal{W}} R_{KQ_j}$$

The rest is just using property of the measure μ:

$$\begin{aligned}
\mu\left(\{(y,t) \mid |F(y,t)| > \lambda\}\right) &\leq \sum_{Q_j \in \mathcal{W}} \mu(R_{KQ_j}) \\
&\lesssim \|u\|_C \sum_{Q_j \in \mathcal{W}} |KQ_j| \qquad \text{(Property of Carleson measure)} \\
&\leq K^n \; \|u\|_C \; |E_\lambda|
\end{aligned}$$

because $\mathcal{W}$ decomposes E_λ. This shows the bound for distribution. Taking integral on both sides will give the wanted expression.

□

We may modify our proof for this lemma to get a more explicit duality. See [2] p. 59, Theorem 2(b): our Lemma 9.5 is part (a) and Corollary on p. 61.

10 Hardy Spaces H^p

In this section, we will consider the Hardy spaces H^p in $\mathbb{R}^n$. For simplicity, we will consider only the case when $\dfrac{n}{n+1} < p \leq 1$, while most results will only be presented for $p = 1$.

Definition 10.1. A function a is an H^p atom if there is a cube $Q \subseteq \mathbb{R}^n$ so that

(i) $\operatorname{supp} a \subseteq Q$.

(ii) $\|a\|_2 \leq |Q|^{\frac{p-2}{2p}}$.

(iii) $\int_Q a = 0$. In particular, a averages to 0 over $\mathbb{R}^n$.

We should make a small comment regarding the exponent in (ii) of the definition above. Note that for any p in consideration, we may apply Hölder inequality to exponent $2/p$ to get

$$\int_{\mathbb{R}^n} |a|^p \leq \|a\|_2^p \, |Q|^{\frac{2-p}{2}}$$

In particular, this says that the p-th norm is bounded by 1. On the other hand, we may replace condition (ii) with any other L^q bound for $q > 1$. The most important among them is the case $q = \infty$:

$$\|a\|_\infty \leq |Q|^{-1/p}$$

We can now define the Hardy space with respect to these atoms.

Definition 10.2. The Hardy space H^1 is the space of L^1 function f so that there exists a sequence $\{\lambda_k\} \in \ell^1$ and a sequence of H^1 atoms $\{a_k\}$ so that $f = \sum_{k=1}^{\infty} \lambda_k a_k$ converges in L^1 norm.

Note that H^1 is a Banach space with norm

$$\|f\|_{H^1} := \inf_{\{a_k\}} \sum_{k=1}^{\infty} |\lambda_k|$$

In other words, the infimum is taken over all atomic decompositions $f = \sum_k \lambda_k a_k$. Because $\|a_k\|_1 \leq 1$ for all k, we have that $\|f\|_1 \leq \|\{\lambda_k\}\|_1$. In particular, $\|f\|_1 \leq \|f\|_{H^1}$.
While we have $H^1 \subseteq L^1$, the space L^1 is strictly larger. Indeed, let $f = \sum_k \lambda_k a_k \in H^1$. By the absolute convergence of the decomposition, we may switch integration and summation to get

$$\int_{\mathbb{R}^n} f = \sum_{k=1}^{\infty} \lambda_k \int_{\mathbb{R}^n} f_k = 0$$

In particular, we have our first criterion of H^1 functions: its averages to zero on $\mathbb{R}^n$.
As one might have guessed, just as the space BMO is a substitute of L^∞, we can think the space H^1 as replacement of L^1. Indeed, the following theorem gives a stronger statement: not only they may 'replace' their respective spaces, but they are also dual like the spaces they're supposed to replace.

Theorem 10.3 (Fefferman)**.** $(H^1)^* = BMO$

The next theorem gives a way that H^1 space may naturally rise.

Theorem 10.4. *If T is a SIO associated with a standard C-Z kernel and T is L^2-bounded, then T maps H^1 to L^1.*

In the following, we present a series of characterizations of H^1.

Theorem 10.5. *Let $\{\psi_t\}_{t>0}$ be a collection of convolution-type L-P class functions with ψ be a bump function with average zero. Let $Q_t f = \psi_t * f$ so that the general Calderon reproducing formula (Proposition 8.1) applies. The H^1_ψ space is defined as*

$$H^1_\psi = \left\{ f \in L^1 \mid S_\psi f \in L^1 \right\}$$

where $S_\psi f$ is the vertical square function and with the norm $\|f\|_{H^1_\psi} := \left\|S_\psi f\right\|_1$. Then $H^1 = H^1_\psi$. In particular, $\left\|S_\psi f\right\|_1 \approx \|f\|_{H^1}$.

A small note about the reproducing formula is in order here. In Proposition 8.1, we presented a special case of Calderon reproducing formula when ψ is especially nice. This condition can be relaxed if we allow a more general result: let ψ be a radial function that satisfies L-P condition and $Q_t f = \psi_t * f$. Then the resolution of the identity

$$\int_0^\infty \widetilde{Q}_t Q_t \cdot \frac{dt}{t} = \mathrm{Id}$$

holds in L^2-operator norm. Here $\widetilde{Q}_t f = \widetilde{\psi}_t * f$, where $\widetilde{\psi}$ is a bump funtion supported on $B_1(0)$ and averages zero.
We should note two important cases here: first, if ψ is smooth, compactly supported, and radial, then we're back in the setting of Proposition 8.1 and $\widetilde{Q_t} = Q_t$, so we obtain the usual representation formula.
If ψ is merely radial (but not necessarily compactly supported, say), then $\widetilde{Q_t}$ satisfies the generalized reproducing formula whenever $\widetilde{\psi}$ is radial and $\widehat{\widetilde{\psi}}(t)\,\widehat{\psi}(t) \geq 0$ for all t with the product being strictly positive for at least one value of t. Indeed, the prototypical example (and the proof of this fact) we'll keep in mind will be encapsulated in the following proposition:

Proposition 10.6. *Set $\psi_t(x) = -t\,\partial_t P_t(x)$, where $P_t = c_n \dfrac{t}{(t^2 + |x|^2)^{1/2}}$ is the Poisson kernel on $\mathbb{R}^{n+1}_+$. Define $Q_t f := \psi_t * f$. Then there exists $\widetilde{\psi} \in C_c^\infty(B_1(0))$ which is radial and averages to 0 so that $\widetilde{Q}_t f := \widetilde{\psi}_t * f$ satisfies*

$$\int_0^\infty \widetilde{Q}_t Q_t \cdot \frac{dt}{t} = Id$$

in L^2-operator norm.

Proof. Recall that $\widehat{P}_t(\xi) = e^{-2\pi t|\xi|}$ (see, for example, [1] p. 62). Then $\widehat{\psi}_t(\xi) = 2\pi t|\xi|e^{-2\pi t|\xi|}$. Choose a non-trivial real-valued function $\widetilde{\psi}^0 \in C_c^\infty(B_{1/2}(0))$ that is radial and averages zero, and then set $\widetilde{\psi} = \widetilde{\psi}^0 * \widetilde{\psi}^0$. Then $\widetilde{\psi} \in C_c^\infty(B_1(0))$,radial, and also averages to 0. Moreover,

$$\widehat{\widetilde{\psi}}^2_t(\xi) = (\widehat{\widetilde{\psi}}^0_t(\xi))^2 = |\widehat{\widetilde{\psi}}^0(t|\xi|)|^2$$

Thus for nice enough f (say $f \in \mathcal{S}$),

$$\left(\int_0^\infty \widetilde{Q}_t Q_t f \,\frac{dt}{t}\right)^\wedge(\xi) = \int_0^\infty (\widetilde{Q}_t Q_t f)^\wedge(\xi)\,\frac{dt}{t}$$
$$= \int_0^\infty |\widehat{\widetilde{\psi}}^0(t|\xi|)|^2\, 2\pi|\xi|e^{-2\pi t|\xi|}\,\frac{dt}{t}\,\widehat{f}(\xi)$$

However, note that $2\pi|\xi|e^{-2\pi t|\xi|}$ integrates to a positive constant. Moreover, because $\widetilde{\psi}^0$ is chosen to be non-trivial, $|\widehat{\widetilde{\psi}}^0(t|\xi|)|$ is non-zero on a set of positive measure, so its integral is also strictly positive. Thus, for some constant $C > 0$, we may conclude that

$$\left(\int_0^\infty \widetilde{Q}_t Q_t f \,\frac{dt}{t}\right)^\wedge(\xi) = C\,\widehat{f}(\xi)$$

Normalize by C to obtain the general reproducing formula. □

Note that in the proposition, $\widetilde{\psi}$ satisfies all the requirement of the function ψ in Theorem 10.5. Thus we may take the general Calderon reproducing formula to be in the set-up of the previous proposition: we'll consider when $\widetilde{\psi}$ is a bump function and ψ is merely radial, where both satisfies the L-P condition.
The next characterization of atomic H^1 is also significant.

Theorem 10.7. *The non-tangential H^1 spaces is defined as*

$$H^1_{NT} = \{f \in L^1 \mid N_*(P_t f) \in L^1\}$$

where N_*F is the non-tangential maximal function (on the cone $\Gamma(x)$), P_t is the Poisson kernel on $\mathbb{R}^{n+1}_+$, and $P_t f = P_t * f$. The norm is given by $\|f\|_{H^1_{NT}} := \|N_*(P_t f)\|_1$. Then $H^1 = H^1_{NT}$. In particular, $\|N_*(P_t f)\|_1 \approx \|f\|_{H^1}$.

Theorem 10.8. *The Riesz H^1 space is defined as*

$$H^1_{Riesz} = \{f \in L^1 \mid R_j f \in L^1,\ 1 \le j \le n\}$$

where R_j is the j-th Riesz operator with the norm

$$\|f\|_{H^1_{Riesz}} = \|f\|_1 + \sum_{j=1}^n \left\|R_j f\right\|_1$$

Then $H^1 = H^1_{Riesz}$. In particular, $\|f\|_{H^1_{Riesz}} \approx \|f\|_{H^1}$.

We present the proofs of these statements below.

10.1 Proof of Theorem 10.3

We will begin by showing BMO $\subseteq (H^1)^*$. For the proof, we will rely on a Hölder-like inequality relating BMO and H^1. The basic version is in this following lemma:

Lemma 10.9. *Suppose a is an H^1 atom, then for all $b \in BMO$,*

$$\left| \int_{\mathbb{R}^n} ab \right| \le c_n \, \|b\|_*$$

Proof. The proof is just by calculation. Because a is of average zero, we may, as usual, insert a constant in the integral above. Indeed, if we insert $b_Q := \fint_Q b$, where Q is the cube associated with a,

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} a(x)\, b(x)\, dx \right| &= \left| \int_{\mathbb{R}^n} a(b - b_Q)\, dx \right| && \text{(Inserting constant)} \\
&\le \int_Q |a|\, |b - b_Q|\, dx \\
&\le \|a\|_2 \left(\int_Q |b - b_Q|^2\, dx \right)^{1/2} && \text{(Cauchy-Schwarz)} \\
&\le \left(\fint_Q |b - b_Q|^2\, dx \right)^{1/2} && (\|a\|_2 \le |Q|^{-1/2}) \\
&\le c_n \, \|f\|_*
\end{aligned}$$

where the last line we used the fact that BMO_2-norm (in the line before the last) is equivalent to the usual BMO-norm. ☐

Since we now know how BMO relates with H^1 atoms, it is natural to consider the following extension to atomic decomposition of $f \in H^1$:

Claim. *Suppose $f \in H^1$ has a finite atomic decomposition, i.e. $f = \sum_{k=1}^{N} \lambda_k a_k$. Then*

$$\left| \int_{\mathbb{R}^n} b\, f \right| \le c_n \, \|b\|_* \, \|f\|_{H^1}$$

Proof. It might seems like this is an easy consequence of the lemma above, because an application of triangle inequality gives

$$\left| \int_{\mathbb{R}^n} b\, f \right| \le c_n \, \|b\|_* \sum_{k=1}^{N} |\lambda_k|$$

from which we can approximate $\|f\|_{H^1}$. This is not quite right, however. The decomposition is not necessarily unique and the optimal bound (i.e. the norm) might be achieved by infinite sum.

Indeed, consider another decomposition $f = \sum_{k=1}^{\infty} \widetilde{\lambda}_k a_k$, chosen so that $\sum_{k=1}^{\infty} |\widetilde{\lambda}_k| \leq 2 \|f\|_{H^1}$. Also consider for now when $b \in L^\infty$. Then because $\|a_k\|_1 \leq 1$ for all k, the sum $\sum_{k=1}^{\infty} \widetilde{\lambda}_k a_k$ converges in L^1. In particular, dominated convergence gives

$$\int_{\mathbb{R}^n} b f = \sum_{k=1}^{\infty} \widetilde{\lambda}_k \int_{\mathbb{R}^n} b a_k$$

Thus by the lemma above,

$$\left| \int_{\mathbb{R}^n} b f \right| \leq c_n \sum_{k=1}^{\infty} |\widetilde{\lambda}_k| \, \|b\|_* \leq c_n \|b\|_* \|f\|_{H^1}$$

and hence this proves the claim.
The extension to BMO is not by density, however (because L^∞ is not dense in BMO, as L^∞ itself is not separable). Instead, given $b \in$ BMO a real-valued function, define

$$b_N(x) = \begin{cases} -N & b(x) \leq -N \\ b(x) & -N < b(x) \leq N \\ N & b(x) \geq N \end{cases}$$

Because f is assumed to have finite decomposition (from the assumption of the claim), f has compact support and thus f is in L^2. Thus because b is also in L^2_{loc}, dominated convergence gives

$$\left| \int_{\mathbb{R}^n} b_N \right| f \overset{n \to \infty}{\to} \left| \int_{\mathbb{R}^n} b f \right|$$

On the other hand, because $b_N \in L^\infty$, the previous case gives $\left| \int_{\mathbb{R}^n} b_N f \right| \leq c_n \|b_N\|_* \|f\|_{H^1}$. Thus the limit is bounded by

$$\left| \int_{\mathbb{R}^n} b f \right| \leq c_n \sup_N \|b_N\|_* \, \|f\|_{H^1} \tag{26}$$

It then remains to show that $\sup_N \|b_N\|_* \lesssim \|b\|_*$. To do so, we will use the following observation: if $b_1, b_2 \in$ BMO, then $\max(b_1, b_2)$ and $\min(b_1, b_2)$ are also in the class BMO and their norm is bounded by $c_n(\|b_1\|_* + \|b_2\|_*)$.
The proof comes from three facts. First, note that it suffices to work with the maximum, because $-\min(b_1, b_2) = \max(-b_1, -b_2)$. Second, if b_Q is any constant, triangle inequality gives $||b| - |b_Q|| \leq |b - b_Q|$. In particular, this means if $b \in$ BMO, so is $|b|$ and $\||b|\|_* \leq 2 \|b\|_*$. Third, we have the identity (that works for all reals)

$$\max(b_1, b_2) = \frac{1}{2} \left[(b_1 + b_2) + |b_1 - b_2| \right]$$

Combining these three facts,

$$\left\|\max(b_1,b_2)\right\|_* \le \frac{1}{2}(\left\|b_1\right\|_* + \left\|b_2\right\|_* + 2\left\|b_1-b_2\right\|_*) \le \frac{3}{2}(\left\|b_1\right\|_* + \left\|b_2\right\|_*)$$

Now we can complete this part of the proof. Let $b_1 = b$ and $b_2 \equiv N$ for some fixed N. Then $\widetilde{b_1} = \min(b,N)$ is BMO, by the observation. and $\left\|\widetilde{b_1}\right\|_* \le c(\left\|b_1\right\|_* + \left\|b_2\right\|_*)$. But the BMO norm doesn't see constants (which has oscillation zero), so $\left\|b_2\right\|_* = 0$ and $\left\|\widetilde{b_1}\right\|_* \le c\left\|b\right\|_*$. Now we'll repeat the same process with $\widetilde{b_1}$ in role of b_1 and $\widetilde{b_2} \equiv -N$. Then $\min(\widetilde{b_1},\widetilde{b_2}) = b_N$ is also BMO and

$$\left\|b_N\right\|_* \le c(\left\|\widetilde{b_1}\right\|_* + \left\|\widetilde{b_2}\right\|_*) = c\left\|\widetilde{b_1}\right\|_* \le c\left\|b_1\right\|_*$$

by previous bound. This bound is independent of N, so we conclude that $\sup_N \left\|b_N\right\|_* \le c\ \left\|b\right\|_*$ and hence (26) gives

$$\left|\int_{\mathbb{R}^n} b\,f\right| \le c_n\left\|b\right\|_*\left\|f\right\|_{H^1}$$

as wanted. This completes the claim. □

The extension to the whole H^1 is by continuity. Define a functional on the space of finite linear combination of H^1 atoms by taking

$$(\Lambda_b, f) = \int_{\mathbb{R}^n} b\,f$$

Such functional must be bounded (in fact $\left\|\Lambda_b\right\|_{(H^1)^*} \le c_n\left\|b\right\|_*$ by the claim). Then, because finite linear combination of atoms are dense in H^1, the functional Λ_b extends to another bounded linear functional Λ on H^1 with the same norm. This allows us to identify BMO with elements of $(H^1)^*$, hence BMO $\subseteq (H^1)^*$.

Now we'll prove the opposite direction. Let $\Lambda \in (H^1)^*$ be a bounded linear functional on H^1. We will construct a function b so that the realization of Λ is an integration against $b\,dx$. We begin by a construction b on a smaller space: given $Q \subseteq \mathbb{R}^n$, define the linear space

$$L^2_{Q,0} = \left\{ g \in L^2(Q) \,\Big|\, \int_{\mathbb{R}^n} g = 0 \right\}$$

Note here we implicitly assume the support of such function is in Q. Now let $g \in L^2_{Q,0}$. Then

$$a = \frac{g}{\|g\|_2}|Q|^{-1/2}$$

is a H^1 atom (averages 0 and $\|a\|_2 \le |Q|^{-1/2}$), so we can write $g = \lambda a$ where $\lambda = \|g\|_2\,|Q|^{1/2}$. Note that this means g itself is in H^1 (as it has an atomic decomposition) and $\|g\|_{H^1} \le \lambda$.

WLOG, after normalizing, we may assume $\|\Lambda\|_= 1$. This means, with $g \in L^2_{Q,0}$, that $|(\Lambda, g)| \le \|g\|_{H^1} \le \|g\|_2\,|Q|^{1/2}$. Because this inequality holds for all cube Q and $g \in L^2_{Q,0}$ (defined from such cube), we know that given any cube $Q \subseteq \mathbb{R}^n$, Λ defines a bounded linear functional on $L^2_{Q,0}$

with $\|\Lambda\|_{(L^2_{Q,0})^*} \le |Q|^{1/2}$. Moreover, because $L^2_{Q,0} \subseteq L^2$ is a Hilbert space, Riesz representation theorem gives an unique $F_Q \in L^2_{Q,0}$ so that we can realize Λ as

$$(\Lambda, g) = \int_{\mathbb{R}^n} F_Q\, g \qquad \forall g \in L^2_{Q,0}$$

and $\|F_Q\|_2 = \|\Lambda\|_{(L^2_{Q,0})^*} \le |Q|^{1/2}$.

With some modification, these F_Q's should be our candidate for the function b. We'll some observation beforehand. First, consider when we have two cubes $Q_1 \subseteq Q_2$ with the corresponding functions (constructed as above) F_{Q_1} and F_{Q_2}. Because $L^2_{Q_1,0} \subseteq L^2_{Q_2,0}$ and the realization of Λ is 'basically unique', we conclude that F_{Q_1} and F_{Q_2} gives rise to the same functional on Q_1, i.e. given $g \in L^2_{Q_1,0}$,

$$\int_{\mathbb{R}^n} (F_{Q_1} - F_{Q_2}) g = 0$$

Because g has average zero, we conclude that F_{Q_1} and F_{Q_2} can differ at most by constant.

Now we'll try to use this 'almost nesting' of the F_Q's to extend to the whole space. For the cube with side length R centered at 0 $Q_R = Q_R(0)$ (with $R \ge 1$), define $b_{Q_R} = F_{Q_R} + c_R$, where $c_R = -\fint_{Q_1} F_{Q_R}$. In other words we choose c_R so that $\fint_{Q_1} b_{Q_R} = \int_{Q_1} b_{Q_R} = 0$.

Let now $1 \le R \le R_2$, so $Q_1 \subseteq Q_{R_1} \subseteq Q_{R_2}$. By previous observation, $b_{Q_{R_1}}$ and $b_{Q_{R_2}}$ differs by at most a constant on Q_{R_1}. However, such constant is necessarily zero, because

$$\int_{Q_1} (b_{Q_{R_1}} - b_{Q_{R_2}}) = 0$$

(by choice). In particular, this is saying that for all $R_1 \le R_2$, given $x \in Q_{R_1}$, we know $b_{Q_{R_1}}(x) = b_{Q_{R_2}}(x)$. We can then define b by

$$b(x) := \lim_{R \to \infty} b_{Q_R}(x)$$

It then remains to show that b has all the wanted properties, i.e. $b \in \mathrm{BMO}$ and $\|b\|_* \le C = C \cdot \|\Lambda\|_{(H^1)^*}$. Fix a cube $Q \subseteq \mathbb{R}^n$. Because the sequence $b_{Q_R}(x)$ is eventantly constant, $b|_Q = b_{Q_R}|_Q$ whenever $Q_R \supseteq Q$. This means $b - c_Q = b_{Q_R} - c_Q = F_Q$ in Q (where F_Q comes from Riesz representation as discussed above and c_Q is chosen to be c_R for $R = \ell(Q)$). Then

$$\begin{aligned}
\frac{1}{|Q|} \int_Q |b - c_Q| &\le \left(\fint_Q |F_Q|^2 \right)^{1/2} && \text{(Hölder, pull out constant)} \\
&= |Q|^{-1/2} \|F_Q\|_2 \\
&\le C && (\|F_Q\|_2 \le C|Q|^{1/2})
\end{aligned}$$

The bound holds uniformly for all Q, so $b \in \mathrm{BMO}$ and $\|b\|_* \le C$. Hence we have a BMO function associated with Λ, so $(H^1)^* \subseteq \mathrm{BMO}$. This completes the proof.

10.2 Proof of Theorem 10.4

Recall that every function $f \in H^1(\mathbb{R}^n)$ can be represented as an infinite linear combination of H^1 atoms. We may then choose one that is close to its norm, i.e. we can find $\{\lambda_k\} \in \ell^1$ and $\{a_k\}$ H^1 atoms so that $f = \sum_{k=1}^{\infty} \lambda_k a_k$ and $\|f\|_{H^1} \approx \sum_{k=1}^{\infty} |\lambda_k|$.
Let T be a SIO with standard C-Z kernel $K(x, y)$ that is L^2-bounded. We begin by by showing that T commutes with infinite sums. Indeed, if $\{\lambda_k\}$ is an ℓ^1 sequence and $\{a_k\}$ is a sequence of H^1 atoms so that $\sum_k \lambda_k a_k \in H^1$, we claim that

$$T\left(\sum_{k=1}^{\infty} \lambda_k a_k\right) = \sum_{k=1}^{\infty} \lambda_k T a_k \tag{27}$$

First observe that since the linear operator T plays nice with finite sums, we may rewrite (27) as

$$T\left(\sum_{k=1}^{N-1} \lambda_k a_k + \sum_{k=N}^{\infty} \lambda_k a_k\right) = \sum_{k=1}^{N-1} \lambda_k T a_k + \sum_{k=N}^{\infty} \lambda_k T a_k \qquad \Rightarrow \qquad T\left(\sum_{k=N}^{\infty} \lambda_k a_k\right) = \sum_{k=N}^{\infty} \lambda_k T a_k$$

for any $N > 1$. By this observation, if we fix $\varepsilon > 0$, the set

$$\left\{\left|T\left(\sum_{k=1}^{\infty} \lambda_k a_k\right) - \sum_{k=1}^{\infty} \lambda T a_k\right| > \varepsilon\right\}$$

is equivalent with the set

$$\left\{\left|T\left(\sum_{k=N}^{\infty} \lambda_k a_k\right) - \sum_{k=N}^{\infty} \lambda T a_k\right| > \varepsilon\right\}$$

Thus, to get that infinite sums commute with T, it suffices to show that the last set has measure that tends to 0 (as $N \to \infty$) for all $\varepsilon > 0$. To do so, we'll split the set into two parts:

$$\left|\left\{\left|T\left(\sum_{k=N}^{\infty} \lambda_k a_k\right)\right| > \frac{\varepsilon}{2}\right\}\right| + \left|\left\{\left|\sum_{k=N}^{\infty} \lambda_k T a_k\right| > \frac{\varepsilon}{2}\right\}\right|$$

For the first one, note that because T has C-Z kernel, the C-Z theorem (Theorem 5.10) says T is of w(1, 1)-type. Applying the weak bound directly, we have

$$\left|\left\{\left|T\left(\sum_{k=N}^{\infty} \lambda_k a_k\right)\right| > \frac{\varepsilon}{2}\right\}\right| \leq \frac{C}{\varepsilon} \int_{\mathbb{R}^n} \left|\sum_{k=N}^{\infty} \lambda_k a_k\right|$$

which should tend to 0 as $N \to \infty$, because we assume the whole series converges in L^1 (because it's an element of H^1). This takes care of the first term.
For the second term, an application of Markov inequality gives

$$\left|\left\{\left|\sum_{k=N}^{\infty} \lambda_k T a_k\right| > \frac{\varepsilon}{2}\right\}\right| \leq \frac{2}{\varepsilon} \sum_{k=N}^{\infty} |\lambda_k| \int_{\mathbb{R}^n} |T a_k|$$

where we used Tonelli to switch sums and integration in the upper bound. It remains to show that the integral of $|Ta_k|$ is uniformly bounded. We'll use the following result:

Claim. *For all H^1 atoms a and SIO T as above, $\|Ta\|_1 \le C$, where the constant depends only on allowable parameters.*

Proof of Claim. Let a be a H^1 atom associated with cube Q. Then

$$\|Ta\|_1 = \underbrace{\int_{2Q} |Ta|}_{I_0} + \sum_{k=1}^{\infty} \underbrace{\int_{2^{k+1}Q \setminus 2^k Q} |Ta|}_{I_k}$$

We'll begin with the zeroth term:

$$\begin{aligned} I_0 &\le C_n\, |Q|^{1/2} \left(\int_{2Q} |Ta|^2 \right)^{1/2} && \text{(Cauchy-Schwarz)} \\ &\le C_n\ \|T\|_{2\to 2}\ \|a\|_2\ |Q|^{1/2} && (T \text{ is } L^2\text{-bounded}) \\ &\le C_n\ \|T\|_{2\to 2} = C && (\|a\|_2 \le |Q|^{-1/2}) \end{aligned}$$

For I_k, we'll try to give pointwise bound to the integrand. Let $x \in 2^{k+1}Q \setminus 2^k Q$ and let y_Q to be the center of Q. Because x and y_Q belongs to separate set,

$$\begin{aligned} |Ta(x)| &= \left| \int_Q K(x,y)\, a(y)\, dy \right| \\ &= \left| \int_Q (K(x,y) - K(x,y_Q))\, a(y)\, dy \right| && (a \text{ has average zero}) \\ &\lesssim \int_Q \frac{|y - y_Q|^\alpha}{|x - y_Q|^{n+\alpha}}\, |a(y)|\, dy && \text{(C-Z kernel continuity)} \\ &\lesssim \frac{(\ell(Q))^\alpha}{(2^k \ell(Q))^{n+\alpha}}\ \|a\|_1 && (|x - y_Q| \approx 2^k \ell(Q)) \\ &\le 2^{-k\alpha} (2^k \ell(Q))^{-n} && (\|a\|_2 \le 1) \end{aligned}$$

From this pointwise bound, we get that

$$I_k \lesssim 2^{-k\alpha}\, (2^k \ell(Q))^{-n}\, |2^{k+1}Q| \lesssim 2^{-k\alpha}$$

and hence

$$\|Ta\|_1 \lesssim C + \sum_{k=1}^{\infty} 2^{-k\alpha} = C$$

where if we trace the constants, C depends on n, C-Z kernel constant, and $\|T\|_{2\to 2}$. □

From this claim, we can see that

$$\left|\left\{\left|\sum_{k=N}^{\infty}\lambda_k T a_k\right| > \frac{\varepsilon}{2}\right\}\right| \le \frac{2}{\varepsilon}\sum_{k=N}^{\infty}|\lambda_k|\int_{\mathbb{R}^n}|Ta_k| \lesssim \frac{1}{\varepsilon}\sum_{k=N}^{\infty}|\lambda_k|$$

which tends to 0 as $N \to \infty$ because $\{\lambda_k\} \in \ell^1$. Thus, we conclude that

$$\left|\left\{\left|T\left(\sum_{k=N}^{\infty}\lambda_k a_k\right) - \sum_{k=N}^{\infty}\lambda T a_k\right| > \varepsilon\right\}\right| \le \left|\left\{\left|T\left(\sum_{k=N}^{\infty}\lambda_k\, a_k\right)\right| > \frac{\varepsilon}{2}\right\}\right| + \left|\left\{\left|\sum_{k=N}^{\infty}\lambda_k T a_k\right| > \frac{\varepsilon}{2}\right\}\right| \overset{N\to\infty}{\to} 0$$

and hence (27) holds and T commutes with infinite sum.

Now we can continue the proof of the theorem. Assuming the chosen atomic representation of f,

$$\begin{aligned}
\|Tf\|_1 &= \left\|\sum_{k=1}^{\infty}\lambda_k T a_k\right\|_1 && (27)\\
&\le \sum_{k=1}^{\infty}|\lambda_k|\ \|Ta_k\|_1 \\
&\le C\sum_{k=1}^{\infty}|\lambda_k| && \text{(Claim)}\\
&\approx C\ \|f\|_{H^1} && \text{(Choice of representation)}
\end{aligned}$$

Thus we conclude that T is a bounded linear map from H^1 to L^1.

10.3 Proof of Theorem 10.5

We start by showing $H^1 \subseteq H^1_\psi$ (This result is basically Thereom 3, p. 114 in [2], with slightly different proof). To do this, we need to show that given $f \in H^1$ with decomposition $f = \sum_{k=1}^{\infty}\lambda_k\, a_k$, chosen so that $\sum_{k=1}^{\infty}|\lambda_k| \le 2\|f\|_{H^1}$ (possible because the H^1 norm is the infimum over all such sequence $\{\lambda_k\}$), we have $\left\|S_\psi f\right\|_1 \lesssim \sum_k|\lambda_k| \lesssim \|f\|_{H^1}$.

To this purpose, it suffices to show two things: (i) the w(1, 1) bound for S_ψ implies $S_\psi\left(\sum_k \lambda_k a_k\right) \le \sum_k|\lambda_k|S_\psi a_k$; (ii) $\left\|S_\psi a\right\|_1 \le C$ uniformly for all H^1 atom a. We'll begin with (i): let a be a H^1 atom associated with the cube Q. Then we can break up the space $\mathbb{R}^n$ into shells like in the proof of Proposition 9.4:

$$\left\|S_\psi a\right\|_1 \le \underbrace{\int_{2Q}|S_\psi a|}_{I_0} + \sum_{k=k}^{\infty}\underbrace{\int_{2^{k+1}Q\setminus 2^kQ}|S_\psi a|}_{I_k}$$

We'll begin by bounding I_0: by Cauchy-Schwarz inequality,

$$I_0 \leq |2Q|^{1/2} \int_{2Q} |S_\psi a| \leq C\, |2Q|^{1/2}\, \|a\|_2$$

where we used the fact that S_ψ is L^2-bounded (Proposition 8.5). However by definition, $\|a\|_2 \leq |Q|^{-1/2}$, so we conclude that I_0 is bounded by some dimensional constant.
To bound I_k, we begin by
Here we should pause the proof for a detour, because our current tools is insufficient to give (a neat) proof of the opposite inclusion. For this purpose, we will introduce the notion of tent spaces:

Definition 10.10. Let $0 < p < \infty$. The tent space T_2^p is a space of function $F(x,t)$ so that if $\mathcal{A}F \in L^p(\mathbb{R}^n)$, where

$$\mathcal{A}F(x) = \left(\int_0^\infty \int_{|x-y|<t} |F(y,t)|^2 \, dy \, \frac{dt}{t^{n+1}} \right)^{1/2}$$

The norm is given by $\|F\|_{T_2^p} := \|\mathcal{A}F\|_p$.

Note that if $F(x,t) = Q_t f(x) = \psi_t * f(x)$ (as before), then $\mathcal{A}F(x) = S_\psi f(x)$. In particular, $f \in H^1_\psi$ iff $S_\psi f \in L^1$ iff $Q_t f \in T_2^1$.
We will approach this tent space with the view of atomic decomposition, like we did for the Hardy space H^1.

Definition 10.11. A T_2^1 atom is a function $A(x,t) \in L^2_{\text{loc}}(\mathbb{R}^{n+1}_+)$ so that there exists a cube $Q \subseteq \mathbb{R}^n$ so that

(i) $\operatorname{supp} A \subseteq R_Q$, where R_Q is the Carleson box (i.e. $R_Q = Q \times [0, \ell(Q)]$).

(ii) $\displaystyle\iint_{R_Q} |A(x,t)|^2 \, dx \, \frac{dt}{t} \leq \frac{1}{|Q|}$.

Note that (ii) is equivalent to saying $\|A\|_2 \leq |Q|^{-1/2}$, just like the H^1 atoms. We'll show several parallel result for atomic tent spaces, which should be familiar to us from the atomic Hardy space.

Lemma 10.12. $\displaystyle\int_0^\infty \int_{\mathbb{R}^n} |F(y,t)|\, |G(y,t)| \, dy \, \frac{dt}{t} \leq C_n \int_{\mathbb{R}^n} \mathcal{A}F(x)\, \mathcal{A}G(x) \, dx$

Proof. Recall that the average of a ball over itself is 1. Thus

$$\begin{aligned}
\int_0^\infty \int_{\mathbb{R}^n} |F(y,t)|\, |G(y,t)| \, dy \, \frac{dt}{t} &= \int_0^\infty \int_{\mathbb{R}^n} |F(y,t)|\, |G(y,t)| \fint_{|x-y|<t} dx \, dy \, \frac{dt}{t} \\
&= C_n \int_{\mathbb{R}^n} \int_0^\infty \int_{|x-y|<t} |F(y,t)|\, |G(y,t)| \, dy \, \frac{dt}{t^{n+1}} \, dx \qquad \text{(Tonelli)} \\
&\leq C_n \int_{\mathbb{R}^n} \left(\int_0^\infty \int_{|x-y|<t} |F(y,t)|^2 \, dy \, fracdtt^{n+1} \right)^{1/2}
\end{aligned}$$

$$\left(\int_0^\infty \int_{|x-y|<t} |G(y,t)|^2 \, dy \, fracdtt^{n+1}\right)^{1/2} dx \quad \text{(Cauchy-Schwarz)}$$

$$= C_n \int_{\mathbb{R}^n} \mathcal{A}F(x)\, \mathcal{A}G(x)\, dx$$

□

The next theorem is more substantial: it shows that the tent space can also be atomically decomposed, suggesting a parallel between T_2^1 and the atomic Hardy space.

Theorem 10.13 (Coifman-Mayer-Stein). *Every $F \in T_2^1$ has an atomic decomposition, i.e. given such F, there is a real number sequence $\{\lambda_k\} \in \ell^1$ and a sequence of T_2^1 atoms $\{A_k\}$ so that $F = \sum_{k=1}^\infty \lambda_k A_k$ and the series converges in T_2^1 with $\sum_k |\lambda_k| \le C_n \|F\|_{T_2^1}$. Moreover, if $F \in T_2^1 \cap T_2^2$, then $\sum_k \lambda_k A_k$ converges also in T_2^2.*

Proof. We begin by defining two sets:

$$\Omega_k = \{x \in \mathbb{R}^n \mid \mathcal{A}F > 2^k\} \qquad \Omega_k^* \{M(\chi_{\Omega_k}) > 1-\gamma\}$$

where $\gamma \in (0,1)$ is fixed (here the actual value doesn't really matter, as we're not optimizing γ). By the w(1, 1)-bound of Theorem 1.2 applied to χ_{Ω_k}, we know

$$|\Omega_k^*| \le \frac{c_n}{1-\gamma}|\Omega_k| \tag{28}$$

On the other hand, if defined this way, Ω_k and Ω_k^* form decreasing sequences and also Ω_k^* is open. By Whitney covering lemma (Lemma 9.6), for each k, there is a sequence of dyadic non-overlapping cubes $\{Q_j^k\}_j$ so that $\Omega_k^* = \bigcup_{j=1}^\infty Q_j^k$ and $\ell(Q_j^k) \approx \operatorname{dist}(Q_j^k, (\Omega_k^*)^c)$ (with constant between 1 and 4). Define now the tent set for Ω_k^* by

$$T(\Omega_k^*) := \left\{(x,t) \in \mathbb{R}_+^{n+1} \mid 0 < t \le 2\operatorname{dist}(x, (\Omega_k^*)^c)\right\}$$

Because $\{\Omega_k^*\}$ itself is decreasing, $T(\Omega_k^*)$ forms a decreasing sequence. Moreover, observe the following: if $x \in \Omega_k^*$, then the Whitney decomposition gives $x \in Q_j^k$ for some j. For such x, $\operatorname{dist}(x, (\Omega_k^*)^c) \le 4 \operatorname{diam} Q_j^k$ (by property of the Whitney cubes) and we conclude if $(x,t) \in T(\Omega_k^*)$,

$$0 < t \le 8 \operatorname{diam} Q_j^k = 8\sqrt{n}\,\ell(Q_j^k) \tag{29}$$

Now we're going to break up the tents into shells (because they are decreasing), which we will break further by the constituent Q_j^k. Define the set

$$T_j^k := (Q_j^k \times (0,\infty)) \bigcap (T(\Omega_k^*) \setminus T(\Omega_{k+1}^*))$$

By (29), we know $T_j^k \subseteq Q_j^k \times (0, \rho\,\ell(Q_j^k))$ for some ρ (in fact we may take $\rho = 8\sqrt{n}$).

Claim. $\sum_{k,j} \chi_{T_j^k}(x,t) = 1 a.e. \{(x,t) \subseteq \mathbb{R}^n \mid F(x,t) \neq 0\}$

Proof of claim. Note that for fixed k, the T_j^k's are non-overlapping, because $T_j^k \subseteq Q_j^k$, which are non-overlapping themselves. If we then put them together,

$$\bigcup_{j=1}^{\infty} T_j^k = \left(\Omega_k^* \times (0,\infty)\right) \bigcap (T(\Omega_k^*) \setminus T(\Omega_{k+1}^*)) = T(\Omega_k^*) \setminus T(\Omega_{k+1}^*)$$

Because $T(\Omega_k^*)$ is decreasing, we know by taking union over k that $\bigcup_{k,j} T_j^k = \bigcup_k T(\Omega_k^*)$. It then follows that

$$\sum_{k,j} \chi_{T_j^k} = \chi_{\cup_k T(\Omega_k^*)}$$

a.e. since the T_j^k's are non-overlapping. Thus to get the claim, it suffices to show $S := \{(x,t) \in \mathbb{R}^n \mid F(x,t) \neq 0\} \subseteq \bigcup_k T(\Omega_k^*)$.

The proof will involve stopping time argument (yet again). Begin by covering $\mathbb{R}_+^{n+1}$ with a grid of unit cubes $\mathbb{D}_0$. If $I \in \mathbb{D}_0$ so that

$$⨍_I |F|^2 = 0 \tag{30}$$

then remove $I \cap S$ from S. The set $I \cap S$ must be of measure zero, because if (30) holds, F is zero a.e. on I.

If (30) doesn't hold, then we keep sub-dividing I until we hit a cube $I' \subseteq I$ in $\mathbb{D}_m$ (for some m) so that $⨍_{I'} |F|^2 = 0$

Let $\{I_l\}$ be the collection of cubes collected from the process above. Because each $I_l \cap S$ has measure zero, so is their union. Now define the set

$$Z = \left(\bigcup_l (I_l \cap S)\right) \bigcup (\text{all cube boundaries in } \mathcal{D}_m \text{ for all } m)$$

This set is also of measure zero. Let $(x,t) \in S \setminus Z$. We'll show that $(x,t) \in T(\Omega_k^*)$ for some k.

The first observation is this: because $(x,t) \notin Z$, for all cubes in the grid $\mathbb{D}_m$ containing (x,t), the integral in (30) is strictly positive. Call this value τ.

Among these cubes containing (x,t) in some dyadic grid, choose one small enough (which we'll call I) so that for all y so that $|x-y| < t/2$, $I \subseteq \Gamma(y)$, where the cone $\Gamma(y) = \{(z,s) \mid |z-y| < s\}$. Moreover, for all such y,

$$\begin{aligned}
\mathcal{A}F(y) &= \left(\iint_{\Gamma(y)} |F(z,s)|^2 \, dz \, \frac{ds}{s^{n+1}}\right)^{1/2} \\
&\geq \left(\int_I |F(z,s)|^2 dz \, \frac{ds}{s^{n+1}}\right)^{1/2} \\
&= c_{n,I} \left(⨍_I |F|^2\right)^{1/2} && (s^{n+1} \approx t^{n+1} \text{ in } I) \\
&= c_{n,I} \tau^{1/2} > 0
\end{aligned}$$

where we choose $c_{n,I}$ to meet the correct proportion of I. Note here we can approximate $s \in I$ by t because I is away from the boundary of the half-plane.
Now choose k large enough so that $c_{n,I}\tau^{1/2} > 2^k$. This means for all $y \in B_{t/2}(x)$, $\mathcal{A}F(y) > 2^k$. In particular, this means $y \in \Omega_k$, so $\chi_{\Omega_k}(y) = 1$ and hence $M(\chi_{\Omega_k})(y) = 1 > 1-\gamma$. This gives $y \in \Omega_k^*$. This means $B_{t/2}(x) \subseteq \Omega_k^*$, so the distance $\operatorname{dist}(x, (\Omega_k^*)^c) \geq t/2$. In turn, this means $t \leq 2\operatorname{dist}(x, (\Omega_k^*)^c)$, and hence by definition $(x,t) \in T(\Omega_k^*)$. □

By this claim, we can write

$$F(x,t) = \sum_{k,j} F(x,t)\,\chi_{T_j^k}(x,t) \text{ a.e.}$$

Set $\lambda_j^k = c_0 \cdot 2^k |Q_j^k|$. We can define the 'primitive' atom $A_j^k = (\lambda_j^k)^{-1} \cdot F \cdot \chi_{T_j^k}$, so we can write a.e.

$$F(x,t) = \sum_{k,j} \lambda_j^k\, A_j^k(x,t)$$

By Dominated Convergence, this sum converges in T_2^1 (and also T_2^2 if $F \in T_2^1 \cap T_2^2$). Indeed,

$$\begin{aligned}
\sum_{k,j} |\lambda_j^k| &= c_0 \sum_k 2^k \sum_j |Q_j^k| \\
&= c_0 \sum_k 2^k |Q_k^*| && (\{Q_j^k\} \text{ is Whitney decomposition}) \\
&\leq \frac{c_n c_0}{1-\gamma} \sum_k 2^k |\Omega_k| && (28) \\
&\approx \frac{c_n c_0}{1-\gamma} \int_{\mathbb{R}^n} \mathcal{A}F && (\text{Distributional definition}) \\
&= \frac{c_n c_0}{1-\gamma} \|F\|_{T_2^1}
\end{aligned}$$

In other words, $\left\|\lambda_j^k\right\|_1$ is comparable to $\|F\|_{T_2^1}$ as hoped from a decomposition. It then remains to show that the A_j^k satisfies the definition of T_2^1 atoms.
Recall $T_j^k \subseteq R_{\rho Q_j^k}$. Then supp $A_j^k \subseteq R_{\rho Q_j^k}$, so we have the support condition. By definition of A_j^k,

$$\iint_{\mathbb{R}^{n+1}_+} |A_j^k|^2\, dx\, \frac{dt}{t} = \frac{1}{c_0^2}\, \frac{1}{2^{2k}}\, \frac{1}{|Q_j^k|^2} \iint_{T_j^k} |F(x,t)|^2\, dx\, \frac{dt}{t}$$

Claim. *There is a constant $b > 0$ so that*

$$\iint_{T_j^k} |F(x,t)|^2\, dx\, \frac{dt}{t} \leq \int_{\substack{bQ_j^k \cap \\ (\Omega_{k+1})^c}} (\mathcal{A}F(x))^2\, dx$$

Proof. Recall $T_j^k \subseteq [Q_j^k \times (0, \rho\,\ell(Q_j^k))] \cap (T(\Omega_{k+1}^*))^c$. Thus if $(y,t) \in T_j^k$, then $(y,t) \in (T(\Omega_{k+1}^*))^c$, i.e. $t > 2\operatorname{dist}(y, (\Omega_{k+1}^*)^c)$. Taking union over all such points, we have

$$T_j^k \subseteq \bigcup_{z \in (\Omega_{k+1}^*)^c} \left\{ (y,t) \mid |y-z| < \frac{t}{2} \right\}$$

With this definition, we claim that if (y,t) belongs to the union above,

$$\int_{(\Omega_{k+1})^c} \chi_{\{|x-y|<t\}}\, dx \ \geq c_{\gamma,n}\, t^n$$

In other words, $(\Omega_{k+1})^c$ is a fixed proportion of the ball $B_t(y)$. To see this, note that for each $z \in (\Omega^*_{k+1})^c$ we know $B_{t/2}(z) \subseteq B_t(y)$. Thus

$$|(\Omega_{k+1})^c \cap B_{t/2}(z)| \leq |(\Omega_{k+1})^c \cap B_t(y)| = \int_{(\Omega_{k+1})^c} \chi_{\{|x-y|<t\}}\, dx$$

On the other hand,

$$\frac{|B_{t/2}(z) \cap \Omega_{k+1}|}{|B_{t/2}(z)|} = ⨍_{B_{t/2}(z)} \chi_{\Omega_{k+1}} \leq (M\chi_{\Omega_{k+1}})(z) \leq 1-\gamma$$

by definition of $(\Omega^*_{k+1})^c$. It then follows that

$$|B_{t/2}(z) \cap (\Omega_{k+1})^c| \geq \gamma |B_{t/2}(z)| = c_{\gamma,n} t^n$$

Combining the two will give the wanted fact. We can then bound the integral in question:

$$\iint_{T^k_j} |F(y,t)|^2\, dy\, \frac{dt}{t} \leq \frac{1}{c_{\gamma,n}} \iint_{T^k_j} |F(y,t)|^2 \int_{(\Omega_{k+1})^c} \chi_{\{|x-y|<t\}}\, dx\, dy\, \frac{dt}{t^{n+1}} \qquad \text{(Fact just above)}$$

$$\leq \int_{(\Omega_{k+1})^c} \int_{Q^k_j} \int_0^{\rho\,\ell(Q^k_j)} |F(y,t)|^2 \chi_{\{|x-y|<t\}}\, \frac{dt}{t^{n+1}}\, dy\, dx \quad \text{(Fubini and definition of } T^k_j\text{)}$$

Now choose $b > 0$ so that $(\Omega_{k+1})^c \cap bQ^k_j$ is a set of positive measure. Continuing from above, we get then

$$\iint_{T^k_j} |F(y,t)|^2\, dy\, \frac{dt}{t} \leq \int_{\substack{(\Omega_{k+1})^c \cap \\ bQ^k_j}} \iint_{|x-y|<t} |F(y,t)|^2\, dy\, \frac{dt}{t^{n+1}}\, dx \approx \int_{\substack{(\Omega_{k+1})^c \\ \cap bQ^k_j}} \iint_{|x-y|<t} |\mathcal{A}F(x)|^2\, dx$$

This gives the claim.

□

We know in $(\Omega_{k+1})^c$, $\mathcal{A}F \leq 2^{k+1}$. Thus from the claim,

$$\iint_{T^k_j} |F|^2\, dx\, \frac{dt}{t} \lesssim c_n\, 2^{2k+2} |Q^k_j|$$

and hence

$$\iint_{\mathbb{R}^n} |A^k_j| 2\, dx\, \frac{dt}{t} \lesssim \frac{4c_n}{c_0^2} \frac{1}{|Q^k_j|} \leq \frac{1}{|Q^k_j|}$$

if $c_0 = \sqrt{4c_n}$. This gives the size condition and hence A^k_j is a T^1_2 atom.

□

Using this theorem (which is highly non-trivial), the rest of the proof comes in several small lemmas.

Lemma 10.14. *Let $\widetilde{Q_t} f := \widetilde{\psi}_t * f$ with $\widetilde{\psi}$ be a real, radial bump function that has average 0. Then*

$$(\Pi_{\widetilde{\psi}} F)(x) = \int_0^\infty \widetilde{Q_t} F(x,t)\,\frac{dt}{t} = \int_0^\infty \int_{\mathbb{R}^n} \widetilde{\psi}_t(x-y)\,F(y,t)\,dy\,\frac{dt}{t}$$

is a bounded operator from T_2^2 to L^2. The convergence of the integral should be understood in the weak sense.

Proof. Since the convergence of the integral is understood in the weak sense, we should see its action under 'inner product' with another L^2 element. Let $g \in L^2$ so that $\|g\|_2 = 1$. Then for $\varepsilon > 0$

$$\begin{aligned}
\int_{\mathbb{R}^n} g(x) \int_\varepsilon^{1/\varepsilon} (\widetilde{Q_t} F(\cdot,t))(x)\,\frac{dt}{t}\,dx &= \int_\varepsilon^{1/\varepsilon} \int_{\mathbb{R}^n} \widetilde{Q_t} g(x) \cdot F(x,t)\,dx\,\frac{dt}{t} && \left(\begin{array}{c}\text{Fubini \& } \widetilde{Q_t} \\ \text{is self-adjoint}\end{array}\right) \\
&\leq \int_{\mathbb{R}^n} \int_0^\infty |\widetilde{Q_t} g(x)| \cdot |F(x,t)|\,dx\,\frac{dt}{t} && \left(\begin{array}{c}\text{Fubini and intro-} \\ \text{duce absolute value}\end{array}\right) \\
&\lesssim \int_{\mathbb{R}^n} \mathcal{A}(\widetilde{Q_\cdot} g)(x) \cdot \mathcal{A}F(x)\,dx && \text{Lemma 10.12} \\
&\leq \left(\int_{\mathbb{R}^n} |\mathcal{A}(\widetilde{Q_\cdot} g)|^2\,dx \right)^{1/2} \left(\int_{\mathbb{R}^n} |\mathcal{A}F|^2\,dx \right)^{1/2} \\
&= \left\| S_{\widetilde{\psi}} g \right\|_2 \|F\|_{T_2^2} \\
&\leq C\,\|F\|_{T_2^2} && (S_\psi \text{ is } L^2\text{-bounded})
\end{aligned}$$

Taking limit as $\varepsilon \to 0$ and supremum over all such g will then gives

$$\left\| \Pi_{\widetilde{\psi}} F \right\|_2 \leq C\,\|F\|_{T_2^2}$$

□

The following lemma is standard, but we'll record it here without proof for future reference

Lemma 10.15. *Let X and Y be Banach spaces. Let $T : X \to Y$ be a bounded linear map and suppose $f = \sum_{k=1}^\infty f_k$ converges in X. Then $Tf = \sum_{k=1}^\infty T f_k$*

Assuming T_2^1 atoms exists (which we will get from Theorem 10.13 that is not yet proven), we might guess that the decomposition of $F(x,t) = \widetilde{Q_t} f(x)$ into T_2^1 atoms should give the H^1 decomposition of f. Indeed, the correspondence is given by the map $\Pi_{\widetilde{\psi}}$, as shown by the following lemma:

Lemma 10.16. *Let $\widetilde{\psi}$ be as in the previous lemma. If A is a T_2^1 atom supported in R_Q, then $\Pi_{\widetilde{\psi}} A = ca$, where a is an H^1 atom supported in $3Q$.*

Proof. □

This following lemma is a recast of previous lemma: not only $\Pi_{\widetilde{\psi}}$ is a bounded operator to L^2, it is also bounded to H^1. In particular, this should guarantee (as we see later) the correspondence between $F = \widetilde{Q_t} f$ and f through the map $\Pi_{\widetilde{\psi}}$.

Lemma 10.17. *Let $f \in T_2^1 \cap T_2^2$. Then $\left\|\Pi_{\widetilde{\psi}} F\right\|_{H^1} \leq C_{n,\widetilde{\psi}} \|F\|_{T_2^1}$.*

Proof. By Theorem 10.13, we can find atomic decomposition $F = \sum_k \lambda_k A_k$ where the sum converges both in T_2^1 and T_2^2. By Section 10.3, $\Pi_{\widetilde{\psi}}$ is a bounded operator, so Lemma 10.15 allows us to interchange sum and operator. Thus by previous lemma,

$$\Pi_{\widetilde{\psi}} F = \sum_k \lambda_k \Pi_{\widetilde{\psi}} A_k = c \sum_k \lambda_k a_k$$

where a_k is an H^1 atom. Because $\|a_k\|_{H^1} \leq 1$,

$$\left\|\Pi_{\widetilde{\psi}} F\right\|_{H^1} \leq c \sum_k |\lambda_k|$$

We can then choose the decomposition so that $\sum_k |\lambda_k| \leq 2 \|F\|_{T_2^1}$. This completes the proof.

□

The previous lemma relies on the class T_2^2. Of course, we'd like to extend this to the space T_2^1. This following lemma gives this extension by continuity.

Lemma 10.18. *$T_2^1 \cap T_2^2$ is dense in T_2^1.*

Proof. Let $F \in T_2^1$ and set

$$F_N = F \cdot \chi_{\{|F|<N\}} \cdot \chi_{\left\{\frac{1}{N}<t<N\right\}}$$

We can see F_N is in T_2^1 (it is just a restriction of F) and, by DCT, $F_N \to F$ in the same space. Moreover,

$$\begin{aligned}
\mathcal{A}F_N(x) &= \left(\int_0^\infty \int_{|x-y|<t} |F_N(y,t)|^2 \, dy \, \frac{dt}{t^{n+1}}\right)^{1/2} \\
&\leq c_N \cdot N \left(\int_{1/N}^N ⨍_{|x-y|<t} dy \, \frac{dt}{t}\right)^{1/2} && \text{(Bound for } F_N \text{ and Fubini)} \\
&= c_{n,N}
\end{aligned}$$

Thus if we take the T_2^2 norm of F_N,

$$\begin{aligned}
\|F_N\|_{T_2^2}^2 &= \int_{\mathbb{R}^n} (\mathcal{A}F_N(x))^2 \, dx \\
&\leq c_{n,N} \int_{\mathbb{R}^n} \mathcal{A}F_N(x) \, dx && \text{(Bound one of the } \mathcal{A}F_N) \\
&\leq c_{n,N} \|F\|_{T_2^1}
\end{aligned}$$

Hence F_N is also in T_2^2, i.e. we can approximate F by a sequence $\{F_N\}$ in $T_2^1 \cap T_2^2$. This is equal to being dense in T_2^1.

□

Corollary 10.19. $\left\|\Pi_{\widetilde{\psi}} F\right\|_{H^1} \leq C\ \|F\|_{T_2^1}$ *for all* $F \in T_2^1$.

Proof. We know the statement holds in $T_2^1 \cap T_2^2$. Because $\Pi_{\widetilde{\psi}}$ is bounded in T_2^1, we can extend to the whole T_2^1 by density. □

The next proposition gives what has been alluded before: the map $\Pi_{\widetilde{\psi}}$ connects F and f in the strong sense and thus we recover the reproducing formula.

Proposition 10.20. $\Pi_{\widetilde{\psi}} F = f$ *a.e. on* $\mathbb{R}^n$.

Proof. Set

$$\Pi^{\varepsilon}_{\widetilde{\psi}}(Q_t f) = \int_{\varepsilon}^{\infty} \widetilde{Q_t}(Q_t f)\,\frac{dt}{t} =:\ P_\varepsilon f$$

We'll begin by showing $P_\varepsilon f$ is basically an approximation to identity, so we can use the relevant theorems. In fact, as in Theorem 2.3, we'll show that

(i) $\sup_{\varepsilon>0} |P_\varepsilon f| \leq C \cdot Mf$

(ii) $P_\varepsilon f \to f$ as $\varepsilon \to 0$ a.e. for $f \in C_c^\infty$.

Let φ_ε denote the kernel of P_ε (so $P_\varepsilon f = \varphi_\varepsilon * f$). Because convolution is associative,

$$\varphi_\varepsilon(x) = \int_{\varepsilon}^{\infty} \widetilde{\psi}_t * \psi_t(x)\,\frac{dt}{t} =: \int_{\varepsilon}^{\infty} \widetilde{\widetilde{\psi}}_t(x)\,\frac{dt}{t}$$

Note that $\widetilde{\widetilde{\psi}}$ is also radial, satisfies L-P condition, and averages to 0 (convolution doesn't change any of these properties). Thus

$$\begin{aligned}
\varphi_\varepsilon(x) &= \int_{\varepsilon}^{\infty} t^{-n}\widetilde{\widetilde{\psi}}\left(\frac{x}{t}\right)\frac{dt}{t} \\
&= \int_{1}^{\infty} (\varepsilon t)^{-n}\widetilde{\widetilde{\psi}}\left(\frac{x}{t\varepsilon}\right)\frac{dt}{t} && (t \mapsto t\varepsilon) \\
&= \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)
\end{aligned}$$

where $\varphi = \varphi_1$ is defined in the usual way. We then find that φ_ε is also homogenous, so it really suffices to consider when $\varepsilon = 1$.
The wanted result should come from application of Theorem 2.3. Thus it suffices for us to show that $\varphi \in L^1$ and it is bounded by a radial, decreasing L^1 function. We know that φ is radial (because $\widetilde{\widetilde{\psi}}$ is), so because $\psi \in C_c^\infty$ (the kernel of Q_t) we get

$$|\varphi(x)| \leq \left\|\widetilde{\widetilde{\psi}}\right\|_\infty \int_1^\infty t^{-n-1}\,dt = c_n$$

but also

$$|\varphi(x)| = \left|\int_1^\infty \widetilde{\widetilde{\psi}}\left(\frac{|x|}{t}\right)\frac{dt}{t}\right|$$

$$
\begin{aligned}
&= \left| \int_0^1 \rho^n \widetilde{\widetilde{\psi}}(\rho|x|) \, \frac{d\rho}{\rho} \right| && (\rho = 1/t) \\
&= \left| \int_0^{|x|} \rho^{n-1} \widetilde{\widetilde{(}}\rho) \, d\rho \right| \, |x|^{-n} && \left(\rho \mapsto \frac{\rho}{|x|} \right) \\
&= c_n |x|^{-n} \left| \int_{|y|<|x|} \widetilde{\widetilde{\psi}}(y) \, dy \right| && \big(c_n = |B_1(0)| \big) \\
&= c_n |x|^{-n} \left| \int_{|y| \geq |x|} \widetilde{\widetilde{\psi}}(y) \, dy \right| && (\widetilde{\widetilde{\psi}} \text{ averages to zero}) \\
&= |x|^{-n} \left| \int_{|x|}^{\infty} \widetilde{\widetilde{\psi}}(\rho) \, \rho^{n-1} \, d\rho \right| \\
&\leq |x|^{-n} \int_{|x|}^{\infty} \rho^{-n-\alpha} \cdot \rho^{n-1} \, d\rho && (\text{L-P size condition}) \\
&\lesssim |x|^{-n-\alpha}
\end{aligned}
$$

From these two observations, we conclude

$$
|\varphi(x)| \leq \min\{1, |x|^{-n-\alpha}\} \lesssim \frac{1}{(1+|x|)^{n+\alpha}}
$$

Therefore $\varphi \in L^1$ and we can then take $C(1+|x|)^{-(n+\alpha)}$ as the LDM of φ, so Theorem 2.3 gives us (i). To show (ii), let $f \in C_c^\infty$. Convolution is commutative, so

$$
|\widetilde{Q}_t Q_t f(x)| = |Q_t \widetilde{Q}_t f(x)| = \left| \int_{\mathbb{R}^n} \psi_t(x-y) \int_{\mathbb{R}^n} \widetilde{\psi}_t(y-z)[f(z) - f(y)] \, dz \, dy \right|
$$

where we inserted a constant $f(y)$ in the last integral because $\widetilde{\psi}_t$ has average zero. Now we note the following: f is smooth, so $|f(z) - f(y)| \leq \|\nabla f\|_\infty |z-y|$. However, $\widetilde{\psi}_t$ is a bump function with support on $B_t(y)$, so $|z-y| \leq t$. Hence we conclude

$$
|\widetilde{Q}_t Q_t f(x)| \leq C\, t \, \|\nabla f\|_\infty \, \|\psi\|_1 \, \big\|\widetilde{\psi}\big\|_1
$$

On the other hand, using Young's convolution inequality and the fact $\widetilde{\psi}$ has compact support,

$$
\widetilde{Q}_t Q_t f(x)| \leq \|\psi_t\|_\infty \, \big\|\widetilde{\psi}\big\|_1 \, \|f\|_1 \lesssim t^{-n} \, \big\|\widetilde{\psi}\big\|_1 \, \|f\|_1
$$

and hence for all $x \in \mathbb{R}^n$,

$$
|\widetilde{Q}_t Q_t f(x)| \leq \min(t, t^{-n})
$$

and thus $\int_0^\infty \left|\widetilde{Q}_t Q_t f(x)\right| \frac{dt}{t}$ is finite a.e. x. In particular, $\int_0^\infty \widetilde{Q}_t Q_t f(x) | \frac{dt}{t}$ converges both pointwise and in L^2 (to f, through the reproducing formula). By uniqueness of limit, their limit must be the same, so $P_\varepsilon f \to f$ a.e. as $\varepsilon \to 0$. This gives (ii).

To get the proposition, we apply the same consideration as in Theorem 1.3: first define the function

$$\Lambda f(x) = \limsup_{\varepsilon\to 0} P_\varepsilon f(x) - \liminf_{\varepsilon\to 0} P_\varepsilon f(x)$$

We'll show that $\Lambda f \equiv 0$ a.e. Note that if $f \in H^1$, then $f \in L^1$. Then we can find a sequence $\{f_k\} \subseteq C_c^\infty$ so that $f_k \to f$ in L^1. By (ii), we know that $\Lambda f_k(x) = 0$ everywhere, so for any k,

$$0 \le \Lambda f(x) \le \Lambda(f - f_k)(x)$$

At the same time, $\Lambda f \le 2Mf$. Thus,

$$\begin{aligned} |\{\Lambda f > \varepsilon\}| &\le \left|\{\Lambda(f - f_k)\}\right| \\ &\le \left|\left\{M(f - f_k) > \frac{\varepsilon}{2}\right\}\right| \\ &\le \frac{2C_n}{\varepsilon}\left\|f - f_k\right\|_1 \end{aligned}$$

As $\left\|f - f_k\right\|_1 \to 0$, we conclude $|\{\Lambda f > \varepsilon\}| = 0$ for all $\varepsilon > 0$. In other words, the limit exists. It remains to compute the limit. Indeed, if we take the sequence $\{f_k\}$ as above,

$$\begin{aligned} \left|\lim_{\varepsilon\to 0}(P_\varepsilon f)(x) - f(x)\right| &= \left|\lim_{\varepsilon\to 0}(P_\varepsilon(f - f_k))(x) + (f_k - f)(x)\right| && (P_\varepsilon f_k - f_k \text{ vanishes in limit}) \\ &\le C \cdot M(f - f_k)(x) + |(f_k - f)(x)| && \text{(Use (i))} \end{aligned}$$

Hence for any $\gamma > 0$

$$\begin{aligned} \left|\left\{\left|\lim_{\varepsilon\to 0}(P_\varepsilon f)(x) - f(x)\right| > \gamma\right\}\right| &\le |\{M(f - f_k)(x) > \varepsilon/2\}| + |\{|f - f_k| > \varepsilon/2\}| \\ &\le \frac{2}{\gamma}(C_n + 1)\left\|f - f_k\right\|_1 && \text{(Theorem 1.2 and Markov)} \end{aligned}$$

Again, as $\left\|f - f_k\right\|_1 \to 0$, the measure is 0 (as the bound holds for all γ), so we conclude that $P_\varepsilon f \to f$ a.e. for any $f \in H^1$. However by DCT, we know $P_\varepsilon f \to \Pi_{\widetilde{\psi}} F$ a.e. Uniqueness of limit then gives the proposition. □

Now we are basically done. From the definitions, we have the following equivalences

$$f \in H^1_\psi \iff S_\psi f \in L^1 \iff F := Q_t f \in T^1_2$$

Combining Corollary 10.19 and Proposition 10.20, we know for any $f \in H^1$,

$$\|f\|_{H^1} = \left\|Pi_{\widetilde{\psi}}(Q_t f)\right\|_{T^1_2} \lesssim \left\|Q_t f\right\|_{T^1_2}$$

but says $\left\|Q_t f\right\|_{T^1_2}$ is finite and thus $\left\|Q_t f\right\|_{T^1_2} \approx \|f\|_{H^1_\psi}$. This means $H^1_\psi \subseteq H^1$ and completes the proof of Theorem 10.5.

We close this discussion with a straightforward corollary of Theorem 10.5.

Corollary 10.21. *The space H^1 is equal to H^1_Δ, where*

$$H^1_\Delta = \left\{ f \in L^1 \Big| \left(\int_0^\infty \int_{|x-y|<t} |(t\partial_t P_t * f)(y)|^2 \, dy \, \frac{dt}{t^{n+1}} \right)^{1/2} \in L^1 \right\}$$

Proof. Check that the kernel of the integral expression satisfies the hypothesis of Theorem 10.5. □

10.4 Proof of Theorem 10.7

Recall the definition of non-tangential maximal function: given u defined on $\mathbb{R}^{n+1}_+$, the maximal function is defined as $N_* F(x) = \sup_{|x-y|<t} |u(y,t)|$. Eventhough we function is defined only for cone of aperture 1, the following lemma says that maximal functions defined on cones of different aperture are equivalent.

Lemma 10.22. *Let $b > 0$ and define $N^b_* u(x) := \sup_{|x-y|<bt} |u(y,t)|$ for $u \in C(\mathbb{R}^{n+1}_+)$. Then for all $\widetilde{b} > b$ and for all $0 < p < \infty$,*

$$\int_{\mathbb{R}^n} (N^{\widetilde{b}}_* u(x))^p \, dx \le C \int_{\mathbb{R}^n} (N^b_* u(x))^p \, dx$$

Before we prove this lemma, there are three things to note. First, this lemma only give the non-trivial direction: for $\widetilde{b} > b$, we already have the pointwise bound $N^b_* u(x) \le N^{\widetilde{b}}_* u(x)$, as we're taking supremum over a larger set. The pointwise bound then implies the moment bound. Second, eventhough the maximum function is defined here for continuous function, we may replace u with a more general class if we replace the supremum with the essential supremum. Third, this inequality is only relevant when the RHS integral is finite.

Proof. Like in the proof of Theorem 10.13, we set $E_\lambda = \{N^b_* u > \lambda\}$ and $E^*_\lambda = \left\{ M\chi_{E_\lambda} > c \left(\frac{b}{\widetilde{b}} \right)^n \right\}$ with c_0 to be chosen. Now let $x \in \mathbb{R}^n \setminus E^*_\lambda$. Consider now $(y,t) \in \Gamma^{\widetilde{b}}(x)$, where the cone $\Gamma^{\widetilde{b}}(x) = \{(y,t) \,|\, |x-y| < \widetilde{b}t\}$.

Claim. *For $x \in \mathbb{R}^n \setminus E^*_\lambda$ and $(y,t) \in \Gamma^{\widetilde{b}}(x)$, $B_{bt}(y) \subseteq \mathbb{R}^n$ meets E^c_λ.*

Proof of Claim. Suppose not, so $B_{bt}(y) \subseteq E_\lambda$. Because $x \notin E^*_\lambda$,

$$c_0 \left(\frac{b}{\widetilde{b}} \right)^n \ge M\chi_{E_\lambda}(x) \ge \frac{|E_\lambda \cap B_{\widetilde{b}t}(x)|}{|B_{\widetilde{b}t}(x)|}$$

where the last inequality uses the definition of $M\chi_{E_\lambda}$ as supremum of the average over all such balls centered at x. Now note that because $y \in B_{\widetilde{b}t}(x)$, at least some portion of $B_{bt}(y)$ overlaps $B_{\widetilde{b}t}(x)$

(consider when $|x-y| = \widetilde{b}t$, say). This constant depends only on n, so we can denote this portion as c_n. Continuing from above,

$$c_0\left(\frac{b}{\widetilde{b}}\right)^n \geq \frac{|B_{bt}(y)\cap B_{\widetilde{b}t}(x)|}{|B_{\widetilde{b}t}(x)|} \geq c_n\,\frac{(bt)^n}{(\widetilde{b}t)^n} = c_n\left(\frac{b}{\widetilde{b}}\right)^n$$

Thus if we choose $c_0 < c_n$ (which only depends on n, the last expression is a contradiction. This proves the claim (and fixes c_0).

□

We can then summarize our finding so far: for each $x \in \mathbb{R}^n \setminus E_\lambda^*$ and for each $(y,t) \in \Gamma^{\widetilde{b}}(x)$, there exists $z \in \mathbb{R}^n$ so that $|y-z| < bt$ and

$$|u(y,t)| \leq N_*^b u(z) \leq \lambda$$

Moreover, the claim gives (via contraposition) that $\{N_*^{\widetilde{b}} u > \lambda\} \subseteq E_\lambda^*$. Hence

$$\begin{aligned}
\int_{\mathbb{R}^n} (N_*^{\widetilde{b}}u(x))^p\,dx &= p\int_0^\infty \lambda^{p-1}|\{N_*^{\widetilde{b}} > \lambda\}|\,d\lambda && \text{(Distributional definition)}\\
&\leq p\int_0^\infty \lambda^{p-1}|E_\lambda^*|\,d\lambda && (\{N_*^{\widetilde{b}}u > \lambda\} \subseteq E_\lambda^*)\\
&\leq pC\left(\frac{\widetilde{b}}{b}\right)^n \int_0^\infty \lambda^{p-1}|E_\lambda|\,d\lambda && \text{Theorem 1.2}\\
&= C\left(\frac{\widetilde{b}}{b}\right)^n \int_{\mathbb{R}^n} (N_*^b u(x))^p\,dx && \text{(Distributional definition)}
\end{aligned}$$

Taking $C = C(\widetilde{b}/b)^n$ will then suffice.

□

The next lemma gives a precise relation between the non-tangential maximal function and Hardy-Littlewood maximal function when the function u is reasonably nice.

Lemma 10.23. *Suppose $\varphi_t(x,y) \leq C\,\dfrac{t^\varepsilon}{(t+|x-y|)^{n+\varepsilon}}$ for some $\varepsilon > 0$. Set $P_t f(x) = \displaystyle\int_{\mathbb{R}^n} \varphi_t(x,y)\,f(y)\,dy$. Then for all $b > 0$, we have $N_b^*(P_t f)(x) \leq C\,Mf(x)$.*

Proof. Fix $x \in \mathbb{R}^n$ and $b > 0$. Then for $(y,t) \in \Gamma^b(x)$,

$$|P_t f(y)| = \left|\int_{\mathbb{R}^n} \varphi_t(y,z)\,f(z)\,dz\right| \leq \int_{\mathbb{R}^n} \frac{t^\varepsilon}{(t+|y-z|)^{n+\varepsilon}}\,|f(z)|\,dz$$

Now we'll break the whole space into two regimes: $|y-z| < 5bt$ or $|y-z| \geq 5bt$. In the first regime, triangle inequality gives $|x-z| < 6bt$. Thus if we consider the last integral in this regime,

$$\int_{|y-z|<5bt} \frac{t^\varepsilon}{(t+|y-z|)^{n+\varepsilon}}\,|f(z)|\,dz \leq \frac{1}{t^n}\int_{|x-z|<6bt} |f(z)|\,dz \leq c_{b,n} \fint_{|x-z|<6bt} |f(z)|\,dz \leq c_{b,n} Mf(x)$$

by definition of the H-L maximal function. This gives the wanted bound.
In the second regime, we get $|x - z| \approx |y - z| \geq 5bt$, so

$$\int_{|y-z|\geq 5bt} \frac{t^\varepsilon}{(t + |y - z|)^{n+\varepsilon}} |f(z)|\, dz \leq C \int_{\mathbb{R}^n} \frac{t^\varepsilon}{(t + |x - z|)^{n+\varepsilon}} |f(z)|\, dz$$

Now note that if we define

$$\varphi(x) = \frac{1}{(1 + |x|)^{n+\varepsilon}}$$

We can express the last integral as an approximation to the identity $\varphi_t * f(x)$ where φ_t is defined as usual. Thus by pointwise bound on Theorem 2.3,

$$\int_{|y-z|\geq 5bt} \frac{t^\varepsilon}{(t + |y - z|)^{n+\varepsilon}} |f(z)|\, dz \lesssim Mf(x)$$

Combining the two regime gives the lemma. □

We're now ready to prove the theorem. We'll begin by showing $H^1 \subseteq H^1_{\text{NT}}$. In fact, we'll show a slightly more general statement: $\left\|N_*(\varphi_t * f)\right\|_1 \leq C \|f\|_{H^1}$ whenever φ_t satisfies

(i) $|\varphi_t(x)| \leq C \dfrac{t^\varepsilon}{(t + |x|)^{n+\varepsilon}}$,

(ii) $|\varphi_t(x) - \varphi_t(y)| \leq C \dfrac{|x - y|^\varepsilon}{(t + |x|)^{n+\varepsilon}}$

for some $\varepsilon > 0$. Indeed, both (i) and (ii) are comparable to the L-P condition (except we don't require average zero). Note here that P_t satisfies (i) and (ii) with $\varepsilon = 1$.
We will follow the same strategy as in proof of Theorem 10.5 (showing $H^1 \subseteq H^1_\psi$). To utilize the atomic decomposition in H^1, we'll show that L^1-norm of maximal function of H^1 atoms are uniformly bounded, i.e. given an H^1 atom a, $\left\|N_*(\varphi_t * a)\right\|_1 \leq c$ uniformly for all a.
To do this, set $Q^* = 8Q$ and $R_j = 2^j Q \setminus 2^{j-1}Q$, where Q is the cube associated with a. Then we can do a splitting as before:

$$\left\|N_*(\varphi_t * a)\right\|_1 = \underbrace{\int_{Q*} N_*(\varphi_t * a)}_{\text{I}} + \sum_{j=4}^{\infty} \underbrace{\int_{R_j} N_*(\varphi_t * a)}_{\text{I}_j}$$

The first integral is quite simple:

$$\begin{aligned}
|I| &\leq |Q|^{1/2} \left(\int_{\mathbb{R}^n} (N_*(\varphi_t * a)(x))^2\, dx \right)^{1/2} && \text{(Cauchy-Schwarz)} \\
&\leq C|Q^*|^{1/2} \|Ma\|_2 && \text{(Lemma 10.23)} \\
&\leq C|Q|^{1/2} \|a\|_2 && (M \text{ is } L^2\text{-bounded}) \\
&\leq C && (|Q^*| = 8^n|Q| \text{ and } \|a\|_2 \leq |Q|^{-1/2})
\end{aligned}$$

For the I_j's we begin by giving a pointwise bound: if $x \in R_j$, then $N_*(\varphi_t * a)(x) \le C\, 2^{-j\varepsilon}\, Ma(x)$, where ε is the exponent from φ. To do this, it suffices to show that $|(\varphi_t * a)(y)| \le C\, 2^{-j\varepsilon}\, (Ma)(x)$ for $|x-y| < t$.
Fix $x \in R_j$ and let $|x-y| < t$. Consider first the case of $t \le \ell(Q)$. Because for $j \ge 4$ we have $\ell(Q) \le \frac{1}{8}2^{j-1}\ell(Q)$, we have that

$$|x-y| < t \le \ell(Q) \le \frac{1}{8}2^{j-1}\ell(Q)$$

In particular, for any $z \in Q$, x and y are relatively far from z, so $|z-y| \approx 2^j\ell(Q)$. Using this estimate, we can bound $|(\varphi_t * a)(y)|$ by

$$\begin{aligned}
|(\varphi_t * a)(y)| &\le C \int_{|y-z|\approx 2^j\ell(Q)} \frac{t^\varepsilon}{(t+|y-z|)^{n+\varepsilon}}\,|a(z)|\,dz && \text{(Condition (i) for } \varphi_t) \\
&\lesssim \frac{(\ell(Q))^\varepsilon}{(2^j\ell(Q))^\varepsilon} \fint_{|y-z|\le C\,2^j\ell(Q)} |a(z)|\,dz && (t \le \ell(Q)) \\
&\approx 2^{-j} \fint_{|x-z|\le C\,2^j\ell(Q)} |a(z)|\,dz && (|y-z| \approx |x-z|) \\
&\lesssim 2^{-j\varepsilon}\,(Ma)(x)
\end{aligned}$$

The second case is $t > \ell(Q)$. Let z_Q to be the center of Q. By inserting zero, we get

$$\begin{aligned}
|(\varphi_t * a)(y)| &= \left|\int_{\mathbb{R}^n} (\varphi_t(y-z) - \varphi_t(y-z_Q))\,a(z)\,dz\right| \\
&\lesssim \int_{\mathbb{R}^n} \frac{|z-z_Q|^\varepsilon}{(t+|y-z|)^{n+\varepsilon}}|a(z)|\,dz && \text{(Condition (ii) on } \varphi_t) \\
&\lesssim (\ell(Q))^\varepsilon \int_{\mathbb{R}^n} \frac{|a(z)|}{(t+|y-z|)^{n+\varepsilon}}\,dz && (|z-z_Q| \le \ell(Q))
\end{aligned}$$

Now there are two cases to consider here: whether x and z are sufficiently far or close. In the first, we may consider when $|x-z| > 5t$, so $|x-z| \approx |y-z|$. Thus

$$\ell(Q)^\varepsilon \int_{\mathbb{R}^n} \frac{|a(z)|}{(t+|y-z|)^{n+\varepsilon}}\,dz \lesssim (\ell(Q))^\varepsilon \int_{\mathbb{R}^n} \frac{|a(z)|}{(t+|x-z|)^{n+\varepsilon}}\,dz$$

On the other hand, in the second case $|x-z| \le 5t$, we know $|y-z| < 6t$, so t is much larger than $|y-z|$ and $|x-z|$. Hence

$$\frac{1}{(t+|y-z|)^{n+\varepsilon}} \approx \frac{1}{t^{n+\varepsilon}} \approx \frac{1}{(t+|x-z|)^{n+\varepsilon}}$$

and hence we can approximate

$$\ell(Q)^\varepsilon \int_{\mathbb{R}^n} \frac{|a(z)|}{(t+|y-z|)^{n+\varepsilon}}\,dz \approx \ell(Q)^\varepsilon \int_{\mathbb{R}^n} \frac{|a(z)|}{(t+|x-z|)^{n+\varepsilon}}\,dz$$

In either case the same estimate holds. Because $|x-z| \approx 2^j \ell(Q)$, we have

$$\begin{aligned}|(\varphi_t * a)(y)| &\lesssim (\ell(Q))^\varepsilon \int_{\mathbb{R}^n} \frac{|a(z)|}{(t+|x-z|)^{n+\varepsilon}}\,dz \\ &\lesssim \frac{(\ell(Q))^\varepsilon}{(2^j\ell(Q))^\varepsilon} \fint_{|x-z|\le C\,2^j\ell(Q)} |a(z)|\,dz \\ &\le 2^{-j\varepsilon}\,(Ma)(x)\end{aligned}$$

Using this pointwise bound, we know that

$$\begin{aligned}I_j &\le C2^{-j\varepsilon} \int_{2^jQ} Ma(x)\,dx \\ &\le C2^{-j\varepsilon}|2^jQ|^{\frac{\delta}{1+\delta}} \left(\int_{2^jQ} (Ma)^(1+\delta)\right)^{\frac{1}{1+\delta}} && \text{(Hölder with } 0<\delta<1)\end{aligned}$$

However we know M is $L^{1+\delta}$-bounded, so we can bound the last line with

$$\begin{aligned}I_j &\lesssim 2^{-j\varepsilon}|2^jQ|^{\frac{\delta}{1+\delta}} \left(\int_Q |a|^{1+\delta}\right)^{\frac{1}{1+\delta}} \\ &= 2^{-j\varepsilon}|2^jQ|^{\frac{\delta}{1+\delta}}|Q|^{\frac{1}{1+\delta}} \left(\fint_Q |a|^2\right)^{1/2} && \left(\text{Hölder: } \left(\fint_Q |a|^{1+\delta}\right)^{\frac{1}{1+\delta}} \le \left(\fint_Q |a|^2\right)^{1/2}\right) \\ &= 2^{-j\varepsilon}2^{jn\delta/(1+\delta)} && (\|a\|_2 \le |Q|^{-1/2})\end{aligned}$$

Now choose δ small enough so that $\dfrac{n\delta}{1+\delta} < \varepsilon$, so we have $I_j < 2^{-j\gamma}$ for some small γ (say $\varepsilon/2$). Thus the sum of the I_j's is just an universal constant. Combining with the bound for I will give the wanted uniform bound for all H^1 atom a.

To pass to the general case, we note that $\|\varphi_t\|_1$ is bounded uniformly for all n and ε. Thus by approximation to identity,

$$\sup_{t>0} \|\varphi_t * f\|_1 \le C\,\|f\|_1$$

We then conclude the map $f \mapsto \varphi_t * f$ is bounded uniformly (in t) in L^1.

Fix $\eta > 0$. Because convolution is linear on finite sums, we can know that

$$\left|\left\{\left|\left(\varphi_t * \sum_{k=1}^{\infty} \lambda_k a_k\right) - \sum_{k=1}^{\infty} \lambda_k(\varphi_t * a_k)\right| > \eta\right\}\right| = \left|\left\{\left|\left(\varphi_t * \sum_{k=1}^{\infty} \lambda_k a_k\right) - \sum_{k=1}^{\infty} \lambda_k(\varphi_t * a_k)\right| > \eta\right\}\right|$$

for all $N \ge 1$. However we can bound the set by Markov to get

$$\le \frac{1}{\eta}\left[\int_{\mathbb{R}^n} \left|\varphi_t * \sum_{k=N}^{\infty} \lambda_k a_k\right| + \sum_{k=N}^{\infty} |\lambda_k| \int_{\mathbb{R}^n} |\varphi_t * a_k|\right]$$

We know φ_t is uniformly bounded and the sum $\sum_{k=N}^{\infty} \lambda_k a_k$ converges in L^1, so we can bound the first term by Young's convolution inequality. For the second, we'll use the fact the map $\varphi_t * f$ is uniformly L^1-bounded for all $t > 0$. Thus,

$$\leq \frac{1}{\eta} \sum_{k=N}^{\infty} |\lambda_k| \, \|a_k\|_1 \lesssim \frac{1}{\eta} \sum_{k=N}^{\infty} |\lambda_k| \overset{N\to\infty}{\to} 0$$

Hence the equality

$$\varphi_t * \left(\sum_{k=1}^{\infty} \lambda_k a_k \right) = \sum_{k=1}^{\infty} \lambda_k (\varphi_t * a_k)$$

a.e. (in fact this holds everywhere, because for $f \in L^1$, $\varphi_t * f$ is continuous on $\mathbb{R}_+^{n+1}$). With all this observations, we can then conclude

$$\begin{aligned} N_* \left(\varphi_t * \left(\sum_{k=1}^{\infty} \lambda_k a_k \right) \right)(x) &= \sup_{|x-y|<t} \left| \varphi_t * \left(\sum_{k=1}^{\infty} \lambda_k a_k \right) \right| \\ &= \sum_{|x-y|<t} \left| \sum_{k=1}^{\infty} \lambda_k (\varphi_t * a_k) \right| \\ &\leq \sum_{k=1}^{\infty} |\lambda_k| N_*(\varphi_t * a_k)(x) \end{aligned}$$

Armed with this fact, passing to the general case is now easy: choose a representation of $f = \sum_k \lambda_k a_k \in H^1$ so that $\sum_k |\lambda_k| \leq 2 \|f\|_{H^1}$. Then

$$\begin{aligned} \|N_*(\varphi_t * f)\|_1 &= \sum_k |\lambda_k| \, \|N_*(\varphi_t * a_k)\|_1 && (L^1\text{-convergent series commute}) \\ &\leq C \sum_k |\lambda_k| && (N_* \text{ is uniformly bounded for } H^1 \text{ atoms}) \\ & C \|f\|_{H^1} \end{aligned}$$

This completes the proof for this direction, so $H^1 \subseteq H^1_{\mathrm{NT}}$.
Now we'll prove the other direction: we want to show

$$\|S_\Delta f\|_1 = \|f\|_{H^1_\Delta} \leq C \|N_*(P_t * f)\|_1 = C \|f\|_{H^1_{\mathrm{NT}}} \tag{31}$$

where P_t is the Poisson kernel extension to $\mathbb{R}_+^{n+1}$. Recall that

$$Sf = S_\Delta f = \left(\int_0^\infty \int_{|x-y|<t} |t \partial_t u(y,t)|^2 \, dy \, \frac{dt}{t^{n+1}} \right)^{1/2}$$

where $u(y,t) = (P_t * f)(y)$.
We have shown (Lemma 10.22) that the non-tangential maximal functions have equivalent L^p norms regardless of apertures. We'll consider one that'll be convenient later: fix some $b > 1$ and set $N_* \equiv N_*^b$.

WLOG we may take $f \in H^1_{\text{NT}}$. Otherwise replace f by $f_\varepsilon L = P_\varepsilon * f$. From Young's inequality, we know this approximation of identity is in L^2 whenever $f \in L^1$ (because $\varphi_t \in L^2$).
Similar to what we've seen, let $B_\lambda = \{N_* u > \lambda\}$ (remember here that $N_* \equiv N_*^b$) and $F_\lambda = (B_\lambda)^c$. Define the 'sawtooth region'

$$R = \bigcup_{x \in F_\lambda} \Gamma(x) \qquad R_b = \bigcup_{x \in F_\lambda} \Gamma^b(x)$$

and note that such R (and R^b) is the complement of the 'tent region' (cf. the set $T(\Omega_k^*)$ in the proof of Theorem 10.13). Moreover, we also observe that R is a special Lipschitz domain, because ∂R is the graph $\{(x,t) \,|\, t = \delta(x)\}$, where $\delta(x) = \text{dist}(x, F_\lambda)$, and $\delta(x)$ is Lipschitz with constant 1. Here we'll kmake two observations:

Claim. *(1)* $|u(y,t)| \le \lambda$ *in* $\overline{R_b}$.

(2) $|t\, \nabla u(y,t)| \le C\lambda$ *in* $\overline{R}$

Proof of Claim. The proof of (1) is straightforward: if $(y,t) \in \overline{R_b}$, we know it belongs to some cone $\Gamma^b(x)$ for $x \in F_\lambda$. But on this cone, the maximum $N_* u(x) \le \lambda$, so (1) should follow.
The argument for (2) needs more work. First recall that if $(y,t) \in \overline{\Gamma(x)}$ with $t > 0$, there exists a constant $c > 0$ (not necessarily unique), so that $B((y,t),c) \subseteq \Gamma^b(x)$. Like in (1), because $x \in F_\lambda$, for all $(z,\tau) \in \Gamma^b(x)$, we have $|u(z,\tau)| \le \lambda$.
Now we know that ∇u is harmonic in every compact set (u itself is harmonic), so by the Mean Value Property of harmonic functions,

$$\nabla u(y,t) = ⨌_{B_{c/2}((y,t))} \nabla u(z,\tau)\, dz\, d\sigma$$

Thus applying Jensen inequality to move to L^2-norms,

$$|\nabla u(y,t)| \le \left(⨌_{B_{c/2}((y,t))} |\nabla u|^2 \right)^{1/2}$$

For this expression to give the wanted bound, we'll need the following lemma, which we will not prove (see PDE books):

Proposition 10.24 (Cacciopoli inequality)**.** *Let u be harmonic in* $2B$ *where* $B = B_r(Y) \subseteq \mathbb{R}^d$*. Then there exists a constant* C_d *so that*

$$⨍_B |\nabla u|^2 \le C_d \frac{1}{r^2} ⨍_{2B} |u|^2$$

We'll use this proposition with $B = B_{c/2}((y,t))$. Because we choose c so that $B_c((y,t)) \subseteq \Gamma_b(x)$, where (1) holds,

$$|\nabla u(y,t)| \lesssim \left(⨌_{B_c((y,t))} |u|^2 \right)^{1/2} \le C\,\lambda$$

□

From this observation, we can now finish the proof of the theorem. We'll do so by relating the distribution of $Sf(= S_\Delta f)$ and $N_* u$. First observe that

$$\begin{aligned}
\int_{F_\lambda} (Sf(x))^2\,dx &= \int_{F_\lambda} \iint_{\Gamma(x)} |t\partial_t u(y,t)|^2\,dy\,\frac{dt}{t^{n+1}}\,dx \\
&\le \int_{F_\lambda} \iint_{\Gamma(x)} |\nabla u(y,t)|^2\,dy\,\frac{dt}{t^{n-1}}\,dx && \text{(Replace } \partial_t \text{ with the full gradient)} \\
&= \iint_R |\nabla u(y,t)|^2\,|\{x \in F_\lambda \mid (y,t) \in \Gamma(x)\}|\,dy\,\frac{dt}{t^{n-1}} && \text{(Tonelli)}
\end{aligned}$$

Notice here that switching the switching the domain of integration means taking the values of $x \in F_\lambda$ so that a give $(y,t) \in R$ belongs to $\Gamma(x)$ (which is guaranteed by definition of R, but such x is neither necessarily unique nor always in F_λ). For each $(y,t) \in \Gamma(x)$, by definition $|x-y| < t$, so we know

$$|\{x \in F_\lambda \mid (y,t) \in \Gamma(x)\}| \le |B_t(y)| = c_n t^n$$

Thus the last line can be bounded by

$$\lesssim \iint_R |\nabla u(y,t)|^2\,t\,dt\,dt$$

Moreover, we know u is harmonic (it is a convolution with harmonic function P_t), so $\frac{1}{2}\Delta(u^2) = |\nabla u|^2$. Also easily $\Delta t = 0$, so by Gauss-Green formula

$$\begin{aligned}
&= \iint_R \frac{1}{2}\Delta(u(y,t))^2\,t\,dy\,dt \\
&= \frac{1}{2}\int_{\partial R}\left(t\,\frac{\partial(u^2)}{\partial\nu} - u^2\,\frac{\partial t}{\partial \nu}\right)d\sigma
\end{aligned}$$

where ν is the outer unit normal to ∂R. Here note the boundary term vanishes: we take $f \in L^2$, so

$$|u(x,t)| + t\,|\nabla u(x,t)| \to 0 \qquad \text{as } |x|, |t| \to \infty$$

We'll split the domain of integration to ∂R^{B_λ} and ∂R^{F_λ}, the portions meeting B_λ and F_λ, respectively. Because F_λ is precisely the case when $t = 0$, the boundary and F_λ coincide. We can then write the integral (dropping the constant $\frac{1}{2}$) as

$$\int_{F_\lambda}\left(\underbrace{u\,\frac{\partial u}{\partial \nu}\,t}_{\text{I}} - \underbrace{u^2\,\frac{\partial t}{\partial\nu}}_{\text{II}}\right)d\sigma + \int_{\partial R^{B_\lambda}}\left(\underbrace{u\,\frac{\partial u}{\partial \nu}\,t}_{\text{III}} - \underbrace{u^2\,\frac{\partial t}{\partial\nu}}_{\text{IV}}\right)d\sigma$$

Because $t = 0$ on F_λ, the term I simply vanishes (this might not be strictly ture, but we can approximate $\frac{du}{d\nu}$ by using $f_\varepsilon = P_\varepsilon * f$ and takes the limit as $\varepsilon \to 0$). For III, we note that because A is Lipschitz, $\|\nabla A\|_\infty \le 1$, so

$$d\sigma(x) = \sqrt{1 + |\nabla A|^2}\,dx \le \sqrt{2}\,dx$$

Thus, replacing the normal derivative with the full gradient,

$$\begin{aligned}|\mathrm{III}| &\le \int_{\partial R^{B_\lambda}} t\,|u|\,|\nabla u|\,d\sigma \\ &\lesssim \lambda^2\,\sigma(\partial R^{B_\lambda}) && \text{(Claim (1) and (2))} \\ &\le \sqrt{2}\lambda^2\,|B_\lambda| && \text{(Equivalence of measure)}\end{aligned}$$

For the rest , we note that $|\partial t/\partial \nu| \le |\nabla t| = 1$. Replacing u by its non-tangential maximum, we can write II as

$$|\mathrm{II}| = \left|\int_{F_\lambda} u^2\,\frac{\partial t}{\partial \nu}\,d\sigma\right| \le \int_{F_\lambda} (N_* u)^2\,dx = \int_0^\lambda t\,|\{N_* u > t\}|\,dt$$

On the other hand, we can bound IV using Claim (1) and equivalence of surface measure and Lebesgue norm:

$$|\mathrm{IV}| \le \int_{\partial R^{B_\lambda}} |u|^2\,\left|\frac{\partial t}{\partial \nu}\right|\,d\sigma \le \lambda^2 |B_\lambda|$$

Putting all the pieces all together, we conclude

$$\int_{F_\lambda} (Sf(x))^2\,dx \le C\left(\lambda^2\,|\{N_* u > \lambda\}| + \int_0^\lambda t\,|\{N_* u > t\}|\,dt\right)$$

As mentioned above, we'll now relate the distribution of Sf and the distribution of $N_* u$ using the bound above.

$$\begin{aligned}|\{Sf > \lambda\}| &\le |\{x \in F_\lambda \mid Sf > \lambda\}| + |B_\lambda| \\ &\le \frac{1}{\lambda^2}\int_{F_\lambda} (Sf)^2\,dx + |B_\lambda| && \text{(Markov)} \\ &\lesssim |\{N_* u > \lambda\}| + \frac{1}{\lambda^2}\int_0^\lambda t\,|\{N_* u > t\}|\,dt && \text{(Recall the definition of } B_\lambda\text{)}\end{aligned}$$

Thus, if we integrate the distribution of Sf from 0 to ∞ and use the bound above,

$$\begin{aligned}\|Sf\|_1 &= \int_0^\infty |\{Sf > \lambda\}|\,d\lambda && \text{(Definition)} \\ &\lesssim \int_0^\infty |\{N_* u > \lambda\}|\,d\lambda + \int_0^\infty \frac{1}{\lambda^2}\int_0^\lambda t\,|\{N_* u > t\}|\,dt\,d\lambda \\ &= \left\|N_* u\right\|_1 + \int_0^\infty |\{N_* u > t\}|\,t\int_t^\infty \frac{1}{\lambda^2}\,d\lambda\,dt && \text{(Tonelli)} \\ &= C\,\left\|N_* u\right\|_1\end{aligned}$$

This is precisely the wanted estimate (31), so $H^1_{\mathrm{NT}} \subseteq H^1_\Delta$. This completes the proof.
One small comment about this proof of the second direction: if B_λ is empty for some λ, then $F_\lambda = \mathbb{R}^n$. Thus (31) says

$$\int_{\mathbb{R}^n} (Sf)^2 \lesssim \int_{\mathbb{R}^n} (N_* u)^2$$

Using this fact, by working with the distribution functions again, similar method will give us L^p estimates from the L^2 estimate presented here.

11 Fourier Restriction and Strichartz Estimates

11.1 Oscillatory Integrals

In this section, we will discuss the oscillatory integral of the first kind, as presented in Chapter VIII of [2]. The toy problem we'll have in mind is as follows: given real valued $\varphi \in C^\infty(\mathbb{R})$ and $\psi \in C_c^\infty(\mathbb{R})$, define

$$I(\lambda) := \int_{\mathbb{R}} e^{i\lambda\varphi(x)}\psi(x)\,dx$$

This $I(\lambda)$ arises quite naturally: Fourier transform is the ur-example of such oscillatory integral. The problem we will ask is as follows: what is the rate of decay of $I(\lambda)$ as $\lambda \to \infty$? In other words, what is the best $\beta > 0$ so that asymptotically we have

$$I(\lambda) \to C\,\lambda^{-\beta}$$

Here φ is called the phase functions and the point x_0 where $\varphi'(x_0) = 0$ is called the stationary point. Three basic principles should guide the intuition: first, the principle contribution to $I(\lambda)$ comes from the critical points of φ. Second, if $\varphi'(x_0) = 0$, then $|I(\lambda)| \lesssim \lambda^{-1/k}$ where k is the order of the first non-vanishing derivative at x_0. Third, the is an universal estimate for the decay of $I(\lambda)$ given in terms of a lower bound for some derivative of φ. In this section we will make this precise.

Proposition 11.1. *Let $\varphi \in C^\infty$ be a real-valued function and $\psi \in C_c^\infty$. Furthermore, suppose $\varphi'(x) \neq 0$ for all $x \in supp\,\psi$. Then $|I(\lambda)| \leq C_N \lambda^{-N}$ for all $N \geq 1$.*

Proof. Define $Df(x) = (i\lambda\varphi'(x))^{-1}\frac{d}{dx}f(x)$ (this is well-defined for our application, because $\varphi'(x) \neq 0$ in the relevant domain). By design, we have $D(e^{i\lambda\varphi(x)}) = e^{i\lambda\varphi(x)}$. Moreover, if we take the adjoint D^T to be in the sense so that

$$\int (Df)g = \int f(D^T g)$$

then we can compute that

$$D^T f(x) = -\frac{d}{dx}\left(\frac{f(x)}{-i\lambda\,\varphi'(x)}\right)$$

We've see that D is the identity for $e^{i\lambda\varphi(x)}$, so we may write for any $N \in \mathbb{N}$

$$I(\lambda) = \int_{\mathbb{R}} D^N(e^{i\lambda\varphi(x)})\,\psi(x)\,dx$$

We're working with smooth functions with compact support, so if we're integrating "by parts", tossing D to the other terms, the boundary terms will always vanish and we have

$$I(\lambda) = \int_{\mathbb{R}} e^{i\lambda\varphi(x)}\,(D^T)^N\psi(x)\,dx = i^N\lambda^{-N}\int_{\mathbb{R}} e^{i\lambda\varphi(x)}\,(\widetilde{D})^N\psi(x)\,dx$$

where $\widetilde{D} = -i\lambda D^T$. Because ψ is smooth with compact support and norm of the complex exponent is 1, we can simply take

$$C_N = \int_{\mathbb{R}} |(\widetilde{D})^N\psi(x)|\,dx$$

This gives the wanted constant for the rate of decay. □

The proof of the previous proposition is quite simple. However note that the obtained constant C_N depends on φ, ψ, and N. This is not necessarily the case if we have slightly different assumption. For example, the following proposition gives universal constant depending only on the chosen rate of decay.

Proposition 11.2 (Van der Corput Lemma)**.** *Suppose φ be a real-valued smooth function on a neighbourhood of $[a, b]$ and $|\varphi^{(k)}(x)| \geq 1$ on $[a, b]$ for some $k \geq 1$. Then*

$$\left| \int_a^b e^{i\lambda\varphi(x)}\, dx \right| \leq C_k \lambda^{-1/k}$$

provided either (i) $k = 1$ and φ' is monotone or (ii) $k \geq 2$.

Note. The monotonicity of φ' implies that φ is monotone. We'll see similar result later on with this assumption.

Proof. We begin with (i). Set again $Df = (i\lambda\varphi'(x))^{-1} \frac{d}{dx} f(x)$. As noted before, because D is identity for $e^{i\lambda\varphi(x)}$,

$$\int_a^b e^{i\lambda\varphi(x)}\, dx = \int_a^b D(e^{i\lambda\varphi(x)})\, dx = \int_a^b \frac{1}{i\lambda\varphi'(x)} \frac{d}{dx} e^{i\lambda\varphi(x)}\, dx$$

Again integrating 'by parts', we get that

$$\int_a^b e^{i\lambda\varphi(x)}\, dx = -\int_a^b e^{i\lambda\varphi(x)} \frac{d}{dx}\left(\frac{1}{i\lambda\varphi'(x)}\right) dx + \left.\frac{e^{i\lambda\varphi(x)}}{i\lambda\varphi'(x)}\right|_a^b$$

Because $|\varphi'(x)| \geq 1$ on $[a, b]$ (as $k = 1$), we know the boundary term is bounded by $\frac{2}{\lambda}$.
We make a small observation: if φ' is monotone and never zero, so is $1/\varphi'$. In particular, this means the derivative of $(\varphi')^{-1}$ is always positive or negative. In either case, the absolute value doesn't change anything. Thus, we can bound the integral above as

$$\begin{aligned}
\left| \int_a^b e^{i\lambda\varphi(x)} \frac{d}{dx}\left(\frac{1}{i\lambda\varphi'(x)}\right) dx \right| &\leq \frac{1}{\lambda} \int_a^b \left| \frac{d}{dx}\left(\frac{1}{\varphi'(x)}\right) \right| dx \\
&= \frac{1}{\lambda} \left| \int_a^b \frac{d}{dx}\left(\frac{1}{\varphi'(x)}\right) dx \right| && ((\varphi')^{-1} \text{ is monotone}) \\
&= \frac{1}{\lambda} \left| \frac{1}{\varphi'(b)} - \frac{1}{\varphi'(a)} \right| && \text{(FTC)} \\
&\leq \frac{2}{\lambda} && (|\varphi'| \geq 1)
\end{aligned}$$

Therefore, $I(\lambda)| \leq 4\lambda^{-1}$. We can then take $c_1 = 4$.
The proof of the higher order is by induction. Suppose that we know the k-th case for some $k \geq 1$. By replacing φ with $-\varphi$ if necessary, we may assume $\varphi^{(k+1)}(x) \geq 1$ for all $x \in [a, b]$. This gives $\varphi^{(k)}$ to be a strictly increasing function.

Because $\varphi^{(k)}$ is continuous, we can find c in $[a, b]$ so that $|\varphi^{(k)}(x)|$ assumes its minimum value. To be precise, this means one of the three cases: (i) $\varphi^{(k)}(a) \geq 0$, so $c = a$; (ii) $\varphi^{(k)}(b) \leq 0$, so $c = b$; (iii) $\varphi^{(k)}(a) < 0$ but $\varphi^{(k)}(b) > 0$, so there is $c \in (a, b)$ so that $\varphi^{(k)}(c) = 0$.
Fix $\delta > 0$ to be a constant that'll be chosen later. Because $\varphi^{(k+1)} \geq 1$, we know $|\varphi^{(k)}| \geq \delta$ outside of $(c - \delta, c + \delta)$ (or the corresponding intervals in cases (i) and (ii), but we'll trust the reader to be able to adjust accordingly and not bother with this again).
Write $\lambda\varphi = (\delta\lambda)(\varphi/\delta)$. If we normalize, we know that outside of $(c - \delta, c, \delta)$ we have $|\varphi^{(k)}/\delta| \geq 1$. Thus the induction hypothesis kicks in and we have

$$\left|\int_a^{c-\delta} e^{i\lambda\varphi(x)}\, dx\right| + \left|\int_{c+\delta}^{b} e^{i\lambda\varphi(x)}\, dx\right| \leq 2c_k(\delta\lambda)^{-1/k}$$

On the other hand, we can have the crude bound of

$$\left|\int_{c-\delta}^{c+\delta} e^{i\lambda\varphi(x)}\, dx\right| \leq 2\delta$$

In any case, combining these two estimates give

$$\left|\int_a^{b} e^{i\lambda\varphi(x)}\, dx\right| \leq \frac{2c_k}{(\lambda\delta)^{1/k}} + 2\delta$$

Optimizing δ by taking $\delta = (\lambda\delta)^{-1/k}$, we get $\delta = \lambda^{-1/(k+1)}$. This gives $c_{k+1} = 2c_k + 2$ and completes the proof.

□

If we don't assume ψ to vanish at the end points, we'll get similar estimate with the value of the constant now also depends on ψ. In any case, we can see that the asymptotic behaviour is controlled by the critical points of φ are (or aren't).
Now we'll move to the multi-dimensional case. Here the theory isn't quite as easy, because of the ambiguity involved in defining the critical point. Some things, however, stay the same. For example, we'll prove now a parallel of Proposition 11.1

Proposition 11.3. *Let $\psi \in C_c^\infty(\mathbb{R}^n)$ and $\varphi \in C^\infty(\mathbb{R}^n)$ so that none of its critical points lie in $\operatorname{supp}\psi$. Then for all $N \geq 1$,*

$$|I(\lambda)| \leq C_N \lambda^{-N}$$

Proof. Because $\nabla\varphi \neq 0$ in the support of ψ, we can find for each $x_0 \in \operatorname{supp}\psi$ an unit vector $\xi = \xi(x_0)$ so that $\xi \cdot \nabla\varphi(x) \geq c_0 > 0$ for all $x \in B_\varepsilon(0)$. To be explicit, we can take ξ to be $\nabla\varphi(x_0)/|\nabla\varphi(x_0)|$. Such ξ is an unit vector and $\xi \cdot \nabla\varphi(x_0)$ is non-negative. We know $\nabla\varphi$ is continuous, which then gives such ball $B_\varepsilon(x_0)$.
In fact, because $\operatorname{supp}\psi$ is compact, we can cover it by finite number of such balls $B_{\varepsilon_k}(x_k) =: B_k$ for $1 \leq k \leq M$ (with corresponding ξ_k). Each B_k gives a lower bound c_0, so we can take the smallest to be the overall lower bound.

Consider now a partition of unity adapted to B_k, i.e. consider $\eta_k \in C_c^\infty$ so that $\operatorname{supp} \eta_k \subseteq B_k$ and $\sum_{k=1}^M \eta_k \equiv 1$ on $\operatorname{supp} \psi$. Set $\varphi_k = \psi \, \eta_k$. Then we can write

$$I(\lambda) = \sum_{k=1}^{M} \underbrace{\int_{\mathbb{R}^n} e^{i\lambda\varphi(x)} \psi_k(x)\, dx}_{\mathrm{I}_k}$$

It then suffices to bound individual I_k. Choose a coordinate system such that the first coordinate x_1 is in the direction of ξ_k. We can evaluate I_k as an iterated integral, i.e.

$$\int_{\mathbb{R}^n} e^{i\lambda\varphi(x)} \psi_k(x)\, dx = \int \cdots \int e^{i\lambda\varphi(x_1,\cdots,x_n)} \psi_k(x_1, \cdots, x_n)\, dx_1 \cdots dx_n$$

We know the one-dimensional Proposition 11.1 holds, so applying it to the dx_1 integral (and keeping everything else constant) will give the term λ^{-N}. The rest is finite, because $\operatorname{supp} \psi$ is finite. Bounding similarly and summing over all k will then give the constant C_N.

□

As noted above, the critical point in the multi-dimensional case works somewhat different. For example, we may have a saddle point or inflection point. To rule these out, we'll only consider when the critical point is meaningful:

Definition 11.4. Let $\varphi \in C^2(\mathbb{R}^n)$ and suppose $\nabla\varphi(x_0) = 0$. Then x_0 is a non-degenerate critical point of φ if $\det(H\varphi)(x_0) \neq 0$. Here $H\varphi$ is the Hessian matrix of φ.

In particular, from Taylor expansion we can see that x_0 is an isolated critical point. The reason for considering non-degenerate critical point can be seen if we recall van der Corput's Lemma (Proposition 11.2): we need a way to make sense of the monotonicity of the derivative. Indeed, we can see that the following proposition is a weak analogue of Proposition 11.2.

Proposition 11.5. *Suppose $\varphi(x_0) = 0$ and x_0 is a non-degenerate critical point of φ. If ψ is supported on a sufficiently small neighbourhood of x_0 so that no other critical point of φ is contained in this neighbourhood, then*

$$\left| \int_{\mathbb{R}^n} e^{i\lambda\varphi(x)} \psi(x)\, dx \right| \lesssim \lambda^{-n/2}$$

Note. The assumption of φ vanishing at x_0 is not very important. We can always replace $\varphi(x)$ by $\phi(x) = \varphi(x) - \varphi(x_0)$. Because φ is real-valued, we only change the integral by a constant of modulus 1, which has no effect on the asymptotics.

Proof. After some possible translation, take $x_0 = 0$. We'll first prove a special case: suppose

$$\varphi(x) = Q(x) := \sum_{j=1}^{m} x_j^2 - \sum_{j=m+1}^{n} x_j^2$$

for some $0 \leq m \leq n$. Because Q takes this special form, we know that $\frac{\partial}{\partial x_j} Q = \pm 2x_j$. In particular, this means ∇Q vanishes at the origin so $x_0 = 0$ is a critical point. This critical point is non-degenerate, because HQ is a constant diagonal matrix with 2's and -2's on the diagonal and zero otherwise.
Define

$$\Gamma_j^0 = \left\{ x \in \mathbb{R}^n \middle| |x_j|^2 \geq \frac{|x|^2}{n} \right\} \qquad \Gamma_j = \left\{ x \in \mathbb{R}^n \middle| |x_j|^2 \geq \frac{|x|^2}{2n} \right\}$$

In other words Γ_j^0 (and Γ_j) are double cones with axis in the direction of x_j. Observe that $\Gamma_j^0 \subseteq \Gamma_j$ and $\bigcup_{j=1}^n \Gamma_j^0 = \mathbb{R}^n$. Indeed, we know $|x|^2 \leq \sum_{j=1}^n |x_j|^2$, so by pigeon hole principle there must be at least one j so that $|x_j|^2 \geq |x|^2/n$ (We should be pardoned for the grave abuse of notation here: x_j is both the j-th direction in the coordinate system and the j-th component of a given x).
Now choose a smooth partition of unity $\Omega_1, \cdots, \Omega_n \in C^\infty(S^{n-1})$ with respect to Γ_j so that each Ω_j is homogenous of degree zero. Then $\sum_{j=1}^n \Omega_j(x) = 1$ for all $x \neq 0$. Note here we take the Ω_j's to be adapted to Γ_j instead of Γ_j^0, because for the smooth transition to work, we need some overlap between the partitions of the space.
As in previous proposition (Proposition 11.3), we'll decompose the wanted integral and bound individually. So,

$$\int_{\mathbb{R}^n} e^{i\lambda Q(x)} \psi(x)\, dx = \sum_{j=1}^n \underbrace{\int_{\mathbb{R}^n} e^{i\lambda Q(x)} \psi(x)\, \Omega_j(x)\, dx}_{\mathrm{I}_j}$$

Set $D_j f = (\pm 2i\lambda x_j)^{-1} \frac{\partial f}{\partial x_j}$, where the sign is chosen to match the sign of $\frac{\partial}{\partial x_j} Q$. In this construct, we have $D_j e^{i\lambda Q} = e^{i\lambda Q}$.
In the support of Ω_j (i.e. Γ_j) we have $|x_j| \gtrsim |x|$. Then because Ω_j is homogenous of degree zero, $|\nabla^N \Omega_j| \leq c_N |x|^{-N}$ for $N \geq 1$. Set D_j^T to be the transpose of D_j, e.g.

$$D_j^T f = \frac{\partial}{\partial x_j} \left(\frac{\mp f}{2i\lambda x_j} \right)$$

Again, the sign is chosen as in D_j. Then

$$|(D_j^T)^N \Omega_j(x)| \leq c_N \lambda^{-N} |x|^{-N} |\nabla^N \Omega_j(x)| \leq c_N \lambda^{-N} |x|^{-2N}$$

Here we used the $|x_j|^{-N} \lesssim |x|^{-N}$ and modify c_N slightly. On the other hand,

$$|(D_j^T)^N \psi(x)| \leq c_N \lambda^{-N} |x|^{-N} \left| \frac{\partial^N}{\partial x_j^N} \psi(x) \right| \leq c_N \lambda^{-N} |x|^{-N}$$

where we used the fact that ψ is a smooth function with compact support, so its N-th j-derivative is bounded. Moreover, if $\operatorname{supp} \psi$ is sufficiently small we may replace $|x|^{-N}$ by $|x|^{-2N}$. Combining these observations we have

$$|(D_j^T)^N (\psi(x)\, \Omega_j(x))| \leq c_N \lambda^{-N} |x|^{-2N} \tag{32}$$

We 'know' that x_0 controls the decay behaviour, so we'll try to cut it out. Consider a smooth radial cut-off function α supported on $B_2(0)$ with $\alpha \equiv 1$ on $B_1(0)$. For some $\varepsilon > 0$ to be determined, we can write

$$I_j(\lambda) = \underbrace{\int_{\mathbb{R}^n} e^{i\lambda Q(x)}\, \psi(x)\Omega_j(x)\, \alpha\left(\frac{|x|}{\varepsilon}\right) dx}_{\mathrm{I}'_j} + \underbrace{\int_{\mathbb{R}^n} e^{i\lambda Q(x)}\, \psi(x)\Omega_j(x) \left[1 - \alpha\left(\frac{|x|}{\varepsilon}\right)\right] dx}_{\mathrm{II}'_j}$$

Note that the first integral is really taken over a small ball around the origin. Thus if we just take the uniform bound,

$$|\mathrm{I}'_j| \leq \|\psi\|_\infty \int_{|x|\leq 2\varepsilon} dx = C\varepsilon^n \tag{33}$$

because both α and Ω are chosen to be partition of unity.
We choose the cut-off α precisely so that the integral II_j leaves out a neighbourhood of the origin (where the singularity is). Just like Ω_j, because α is homogenous of degree zero we know $|\nabla^N \alpha| \leq C_N |x|^{-N}$. Combining this fact with (32), we can then bound the integral as follows:

$$\begin{aligned}
|\mathrm{II}'_j| &= \left| \int_{\mathbb{R}^n} \left(D_j^N e^{i\lambda Q(x)}\right) \psi(x)\, \Omega_j(x) \left[1 - \alpha\left(\frac{|x|}{\varepsilon}\right)\right] dx \right| \\
&= \left| \int_{\mathbb{R}^n} e^{i\lambda Q(x)} \cdot (D_j^T)^N \left(\psi(x)\, \Omega_j(x) \left[1 - \alpha\left(\frac{|x|}{\varepsilon}\right)\right]\right) \right| \\
&\leq C_N\, \lambda^{-N} \int_{|x|>\varepsilon} |x|^{-2N}\, dx \\
&\leq C_N\, \lambda^{-N}\, \varepsilon^{n-2N} \qquad \left(N > \frac{n}{2}\right)
\end{aligned} \tag{34}$$

Now we optimize ε: set (33) and (34) to be equal, so

$$\varepsilon^N = \lambda^{-N} \varepsilon^{n-2N} \qquad \Rightarrow \qquad \varepsilon = \lambda^{-1/2}$$

Substituting this value of ε to the bound for $I_j(\lambda)$, we have $|I_j(\lambda)| \lesssim \lambda^{-n/2}$.
Now we move to the general case for 'arbitrary' (smooth etc.) phase φ. Here we don't need to assume the critical point is at the origin. We'll need the following result:

Lemma 11.6 (Morse's Lemma)**.** *Suppose* $\varphi(x_0) = \nabla\varphi(x_0) = 0$ *and* $\det(H\varphi)(x_0) \neq 0$. *Then there exists* $\mathcal{U}$ *a neighbourhood of* x_0*,* V *a neighbourhood of the origin, and* $\varphi : U \to V$ *a diffeomorphism so that*

$$(\varphi \circ \rho^{-1})(y) = \sum_{j=1}^{m} y_j^2 - \sum_{j=m+1}^{n} y_j^2 =: Q(y)$$

for some integer $0 \leq m \leq n$.

The proof of this lemma is elementary but tedious, so we'll skip it. Instead we'll proceed with the proof of the theorem. Let η be a smooth cut-off function on a small neighbourhood $\mathcal{U}$ of x_0 so that Morse's Lemma holds there. Then we can split $I(\lambda)$ as

$$I(\lambda) = \underbrace{\int_{\mathbb{R}^n} e^{i\lambda\varphi(x)}\psi(x)\,\eta(x)\,dx}_{\mathrm{I}_1} + \underbrace{\int_{\mathbb{R}^n} e^{i\lambda\varphi(x)}\psi(x)\,(1-\eta(x))\,dx}_{\mathrm{I}_2}$$

The second integral is easy: since we the integral is away from the critical point x_0, the hypotheses of Proposition 11.3 is satisfied, so we know the rate of decay of I_2 can be made arbitrarily fast. Thus the only non-trivial term is I_1.

Set $\psi_0 = \psi\eta$. By Morse's Lemma there is a smooth bijection ρ betwenn $\mathcal{U}$ to a neighbourhood of the origin and so that $\varphi(\rho^{-1}(y)) = Q(y)$. Thus,

$$\mathrm{I}_1(\lambda) = \int_{\mathbb{R}^n} e^{i\lambda Q(y)}\,\underbrace{\psi_0(\rho^{-1}(y))\,J_{\rho^{-1}}(y)\,dy}_{\widetilde{\psi}(y)}$$

Note tha $\widetilde{\psi}(y)$ is smooth and compactly supported in a neighbourhood of 0 (because $J_{\rho^{-1}}$ is non-vanishing, as x_0 is a non-degenerate critical point). Thus, we have

$$\int_{\mathbb{R}^n} e^{i\lambda Q(y)}\widetilde{\psi}(y)\,dy$$

which reduces to the first case. This completes the proof. □

11.2 Hypersurfaces with Non-vanishing Gaussian Curvature

In this section we'll see an application of the the previous subsection. We begin by describing the set-up: let $S \subseteq \mathbb{R}^{n+1}$ be a smooth surface (so S is a n-dimensional surface). Suppose S has a non-vanishing Gaussian curvature K at every point on S (we haven't defined this yet). Now fix $x_0 \in S$ and choose a coordinate system so that $X_0 = 0$. Moreover, rotate this coordinate system so that the hyperplane $\{X^{n+1} = 0\}$ is tangent to S at 0.

Now parameterize the surface locally as $X = (x, t) \in \mathbb{R}^n \times \mathbb{R}$ so that its tangent plane at the origin is exactly $\{t = 0\}$. We know locally that S is smooth graph, i.e. given $N(0)$ a neighbourhood of zero (or any other point, for that matter), S is given by

$$S \cap N(0) \subseteq \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid t = \varphi(x)\}$$

where $\varphi \in C^\infty(\mathbb{R}^n)$ with $\varphi(0) = \nabla\varphi(0) = 0$. In other words, S is a graph of a smooth function φ. Here it should be noted that such S depends on no particular representation φ.

Let $H\varphi$ be the Hessian of φ. Given any $x_0 \in \mathbb{R}^n$, the eigenvalues of $(H\varphi)(x_0)$, call them $\nu_1 = \nu_1(x_0), \cdots, \nu_n = \nu_n(x_0)$, are called the *principal curvature* of S at the point $(x_0, \varphi(x_0))$ (Here the understanding of the graph enables us to confuse a point on $\mathbb{R}^n$ with a point on the parameterization).

Their product $K(x_0) = \nu_1(x_0) \cdot \cdots \cdot \nu_n(x_0)$ is called the *Gaussian curvature* of S at $(x_0, \varphi(x_0))$. Put in another way, this means the Gaussian curvature at x_0 is equal to $\det(H\varphi)(x_0)$.
It is certainly true that K measures the 'shape' of S. In fact, Gaussian curvature is an intrinsic property of the surface, so even if it can be understood in terms of determinant, it is independent of φ. There is also another reason we're interested in K: if K doesn't vanish, the Hessan $H\varphi$ also doesn't vanish. This enables us to relate K with the result in previous subsection.

Theorem 11.7. *Suppose $S \subseteq \mathbb{R}^{n+1}$ is a smooth surface with nowhere vanishing Gaussian curvature. Let $\psi \in C_c^\infty(\mathbb{R}^n)$, so $\operatorname{supp}\psi \cap S$ is a compact piece of S. Set the measure $d\mu = \psi\, d\sigma$, where σ is the surface measure on $\mathbb{R}^n$. Then $|\widehat{\mu}(\xi)| \leq C|\xi|^{-n/2}$.*

Note. Here we define $\widehat{\mu}$ as the regular $\mathbb{R}^{n+1}$-Fourier transform, i.e.

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}^{n+1}} e^{-2\pi i \xi\cdot x}\, d\mu(x)$$

This is certainly well-defined for finite measure μ (or the next best thing, when $\psi \in L^1$).
If S is a compact surface, e.g. $S = S^n$ (the n-dimensional unit sphere), then we may take $\mu = \sigma$, so the theorem says $|\widehat{\sigma}(\xi)| \leq C|\xi|^{-n/2}$.

Proof. By possibly taking a partition of unity, we may assume $\operatorname{supp}\psi$ is small enough so that $\operatorname{supp}\psi \cap S$ is contained in just one coordinate patch, i.e. we only has to work with one parameterization. This means in $\operatorname{supp}\psi$, S is given by the graph $\{t = \varphi(x)\}$. WLOG we may assume $\varphi(0) = \nabla\varphi(0) = 0$ (as we assume in the set-up). Then $K(x) = \det(H\varphi)(x)$ and $d\sigma(x) = \sqrt{1 + |\nabla\varphi(x)|^2}$. Define now the upward unit normal to S by

$$\vec{N} = -\frac{1}{\sqrt{1 + |\nabla\varphi|^2}}(-\nabla\varphi, 1)$$

Since we assume $\nabla\varphi(0) = 0$, this means at the origin $\vec{N} = e_{n+1}$. Now we'll introduce a terminology: a vector ξ is called a singular point for $\widehat{\mu}(\xi)$ if $\xi/|\xi|$ is 'close' to the unit normal $\vec{N}$. As we will see, this sense of 'closeness' will allow us to approximate the stationary points of this measure.
Our integral of $\widehat{\mu}$ is now somewhat harder to handle because of the surface integral. However if we substitute $d\sigma = \sqrt{1 + |\nabla\varphi|^2}$, we can write

$$\widehat{\mu}(\xi) = \int_S e^{i\xi\cdot P(y)} \underbrace{\psi(y)\sqrt{1 + |\nabla\varphi(y)|^2}}_{\widetilde{\psi}(y)}\, dy \tag{35}$$

where $P(y) = (y, \varphi(y)) = Y$. In particular, $\widetilde{\psi}$ is also of class C_c^∞, so we have not changed the character of the oscillatory integral.
Set $\xi := \lambda\eta$ where $\lambda = |\xi|$ and $\eta = \xi/|\xi| \in S^n$. Write $\eta = (\eta_1, \cdots, \eta_{n+1}) = (\eta', \eta_{n+1})$. We can then rewrite (35) as

$$\int_S e^{i\lambda\Phi(y,\eta)}\widetilde{\psi}(y)\, dy =: I_\eta(\lambda)$$

where $\Phi(y, \eta) = \eta \cdot (y, \varphi(y))$. This I_η is the oscillatory integral that we'll consider with Φ to be the phase function. Indeed when posed in this light, we can see that the strategy for the remaining part should be to show, with Proposition 11.5, that $|I_\eta(\lambda)| \le C\lambda^{-n/2}$ uniformly in η.
We consider now three cases, based on where η lies: (i) η is 'close' to the 'north pole' $\eta_N = e_{n+1}$, (ii) η is 'close' to the 'south pole' $\eta_S = -e_{n+1}$, and (iii) the remaining cases.
We begin the first case with some observations: we have reduced the case to when $\varphi(0) = \nabla\varphi(0) = 0$. Then $\Phi(0, \eta) = 0$ for all η. Also because $\nabla_y \Phi(y, \eta_N) = \nabla\varphi(y)$, we know $\nabla_y \Phi(0, \eta_N) = 0$. Last, because $\Delta\varphi \neq 0$,

$$\det \left[\frac{\partial^2 \Phi}{\partial y_j \partial y_k}(0, \eta_N) \right] = \det(H\varphi)(0) \neq 0$$

We shall now give a quantitative sense for "η being close to η_N". Note that $H\Phi = D_y(\nabla_y \Phi)$. Then $\det(H\Phi) = J_y(\nabla_y \Phi)$ is non-zero if $y = 0$ and $\eta = \eta_N$, by the observation above. Apply now the Implicit Function Theorem to the mapping $\vec{F}(y, \eta) = \nabla_y \Phi(y, \eta)$. Because $D_y \vec{F}(y, \eta_N)|_{y=0}$ has non-zero determinant, we know the system of equations $\vec{F}(y, \eta_N) = \nabla_y \Phi(y, \eta_N) = 0$ has an unique solution $y = y(\eta)$ of class C_c^∞ in a neighbourhood of η_N so that $y(\eta_N) = 0$.
From this discussion, we can conclude that in a neighbourhood $N(\eta_N)$ of the north pole that is small enough (which exists by Implicit Function theorem), then

$$\left| \det \left[\frac{\partial^2 \Phi}{\partial y_j \partial y_k}(y(\eta), \eta) \right] \right| \ge C > 0 \tag{36}$$

for $\eta \in N(\eta_N)$. Here C simply depends on the chosen neighbourhood.
Fix now $\eta \in N(\eta_N)$ and take $x_0 = y(\eta)$, i.e. x_0 is the solution to the equation $\nabla_y \Phi(x_0, \eta_N) = 0$. Observe that by (36), this critical point of Φ is non-degenerate (because the lower bound holds uniformly). At the same time, we know by Mean Value theorem

$$|x_0| \le |y(\eta) - y(\eta_N)| \lesssim |\eta - \eta_N| \le \operatorname{diam}(N(\eta_N))$$

which is small. Consequently $\widetilde{\psi}$ has a small support around x_0 and no other critical point can exist. By Proposition 11.5 this implies

$$|\widehat{\mu}(\xi)| \equiv |I_\eta(\lambda)| \lesssim \eta^{-n/2} = |\xi|^{-n/2} \tag{37}$$

This estimate holds uniformly for $\eta \in N(\eta_N)$, so obtain the claimed rate of decay for the whole neighbourhood. Arguing similarly will give the same rate of decay for $\eta inN(\eta_S)$ (for some neighbourhood of the 'south pole').
It then remains to show the case when η is away from the poles. In this regime, because $\eta_{n+1} \neq \pm 1$ and $\eta \in S^{n-1}$, we must have $|\eta'| \ge c_0 > 0$. Recall that

$$\nabla_y \Phi(y, \eta) = \eta' + \eta_{n+1} \nabla\varphi(y)$$

Because we take $\varphi \in C^\infty$ and $\nabla\varphi(0) = 0$, we have

$$|\nabla\varphi(y)| = |\nabla\varphi(y) - \nabla\varphi(0)| \le C|y| \le \frac{c_0}{2}$$

if $|y|$ is small enough (this comes either because $\operatorname{supp}\widetilde{\psi}$ is small or because we can cover the surface in small local patches). Hence $|\nabla_y \Phi(y,\eta)| \geq c_0/2 > 0$ for all $y \in \operatorname{supp}\psi$. However this means Φ has no stationary phase in $\operatorname{supp}\widetilde{\psi}$. Proposition 11.3 allows us to then make the decay arbitrarily rapid, i.e. for all η in this regime

$$|I_\eta(\lambda)| \leq C_M \lambda^{-M} \qquad M \in \mathbb{N}$$

In practice, this means we can ignore this rapid decay and conclude that on S the decay rate is controlled by (37), as desired. □

11.3 Fourier Restriction

Finally we come to the main object of this section: to estimate the Fourier transform when restricted to a smooth of surface of non-vanishing Gaussian curvature. The set-up is the same: let $S \subseteq \mathbb{R}^{n+1}$ be a smooth surface (of dimension n) with non-vanishing Gaussian curvature. Let $\psi \in C_c^\infty(\mathbb{R}^{n+1})$ and define the finite measure $d\mu = \psi d\sigma$, where σ is the usual surface measure. The Fourier restriction estimate we're interested in will be the following:

$$\left(\int_S |\widehat{f}(\xi)|^2 \, d\mu(\xi) \right)^{1/2} \leq C \, \|f\|_X \tag{38}$$

where X is some Banach function space. This is called the L^2 Fourier restriction.
Let's consider for a bit the possible candidates for X. Certainly L^2 is out of question: by Plancherel's theorem (Theorem 4.5), the Fourier transform is an isometry on L^2. In particular it is surjective, so we may take $\widehat{f}$ to be arbitrary function on L^2. However S is a set of zero $(n+1)$-dimensional Lebesgue measure, so it is insufficient to define a L^2 function in such set.
The opposite is true if we take X to be L^1. Proposition 3.2 gives that $\widehat{f}$ is uniformly continuous. Furthermore $\widehat{f}$ is also bounded (by first part of Proposition 3.2), so we can certainly control behaviour of $\widehat{f}$ on S by $\|f\|_1$. It is then fair to ask what happens in the intermediate L^p spaces. This is indeed the content of the following theorem.

Theorem 11.8 (Stein-Tomas restriction)**.** *Assume the standard set-up. Set* $p_{n+1} = \dfrac{2n+4}{n+4}$*. Then for* $1 \leq p \leq p_{n+1}$*, Equation* (38) *holds with X to be the space* $L^p(\mathbb{R}^{n+1})$*.*

Note. This theorem should indeed take care of all the intermediate cases, because $1 \leq p_{n+1} < 2$ and $p_{n+1} \overset{n\to\infty}{\to} 2$.

Proof. Really it suffices to prove the boundary case when $p = p_{n+1}$, because the other values in between are covered by interpolation via Riesz-Thorin theorem (Theorem 4.7).
Define the Fourier restriction operator on S by $Rf := \widehat{f}|_S$. We want then to show that $R : X \to L^2(S, d\mu)$ is a bounded operator. Here we'll actually consider a bit more than the theorem, as we'll take X from a family of Banach function spaces that includes $L^{p_{n+1}}$.
Recall that L^2 is a Hilbert space, so

$$\int_S |Rf|^2 \, d\mu = \langle Rf, Rf \rangle_\mu = \langle R^* Rf, f \rangle$$

Thus to show that R is bounded, it is equivalent to show that $R^* R : X \to X^*$. We'll make this more precise. By smooth partition of unity and compactness of each patch, we may assume that supp ψ is contained in a single coordinate patch so that S is given by $\tau = \varphi(\xi)$, so $\zeta = (\xi, \tau) \in \mathbb{R}^n \times \mathbb{R}$ (equivalently, $S = \{(\xi, \varphi(\xi))\}$. Thus we may rewrite the LHS of Equation (38) by

$$\int_S |\hat{f}(\zeta)|^2 \, d\mu(\zeta) = \int_{\mathbb{R}^n} \hat{f}(\xi, \varphi(\xi)) \overline{\hat{f}(\xi, \varphi(\xi))} \underbrace{\psi(\xi, \varphi(\xi)) \sqrt{1 + |\nabla \varphi|^2}}_{\widetilde{\psi}(\xi)} \, dx$$

We'll substitute now the definition of $\hat{f}$. For convenience we'll the factor of 2π.

$$\int_{\mathbb{R}^n} \iint_{\mathbb{R}^{n+1}} e^{-i(\xi, \varphi(\xi)) \cdot (x,t)} f(x,t) \, dx \, dt \, \overline{\iint_{\mathbb{R}^{n+1}} e^{-i(\xi, \varphi(\xi)) \cdot (y,s)} f(y,s) \, dy \, ds} \, \widetilde{\psi}(\xi) \, d\xi$$

To save ourselves from some technical headaches, we'll limit our consideration to when f is of class C_c^∞. This class should be dense in X. Then we can apply Fubini to switch the order of integration to get

$$\iint_{\mathbb{R}^{n+1}} f(x,t) \iint_{\mathbb{R}^{n+1}} \Bigg[\underbrace{\int_{\mathbb{R}^n} \exp[-i\xi \cdot (x-y)] \exp[-i\varphi(\xi) \cdot (t-s)] \widetilde{\psi}(\xi) \, d\xi}_{K_{t-s}(x-y)} \Bigg] \overline{f(y,s)} \, dy \, ds \, dx \, dt$$

In other words,

$$\int_S |\hat{f}(\zeta)|^2 \, d\mu(\zeta) = \iint_{\mathbb{R}^{n+1}} f(x,t) \iint_{\mathbb{R}^{n+1}} K_{t-s}(x-y) \, \overline{f(y-s)} \, dy \, ds \, dx \, dt \tag{39}$$

Observe that by definition

$$K_{t-s}(x-y) \equiv K(x-y, t-s) = \hat{\mu}(x-y, t-s)$$

Then by Theorem 11.7, we have that

$$|K_{t-s}(x-y)| \le C|(x-y, t-s)|^{-n/2} \le C|t-s|^{-n/2}$$

uniformly in x, y (the exponent is negative, so we're just removing some terms from the denominator). In particular,

$$\|K_{t-s}(\cdot)\|_\infty \le C|t-s|^{-n/2}$$

This means if we take $g \in L^1$ and consider the $\mathbb{R}^n$ convolution,

$$\|K_{t-s} * g\|_\infty \le C \|K_{t-s}\|_\infty \|g\|_1 \tag{40}$$

Now recall that in the graph coordinate, we can write $K_{t-s}(x) = (e^{-i\varphi(\cdot)(t-s)} \widetilde{\psi}(\cdot))^\wedge(x)$. Then $\check{K}_{t-s}(\xi) = e^{-i\varphi(\xi)(t-s)} \widetilde{\psi}(\xi)$ (we're back in the ξ variable). Notice however that by convolution theorem, $\check{K}_{t-s}(\xi) = \hat{K}_{t-s}(-\xi)$, so we can see that $\left\| \hat{K}_{t-s} \right\|_\infty \le \|\widetilde{\psi}\|_\infty$. Moreover this observation means

$$\|T_{t-s} g\|_2 = \|K_{t-s} * g\|_2 \approx \hat{K}_{t-s} \hat{g} 2 \le \|\widetilde{\psi}\|_\infty \|g\|_2$$

where the approximation accounts for the normalizing constant (because we lose $'2\pi'$ on the FT). This means T_{t-s} is L^2-bounded and its norm is bounded by $\|\widetilde{\psi}\|_\infty \approx O(1)$. Combinining with the L^1-bound (40), Riesz-Thorin (Theorem 4.7) gives that

$$\|T_{t-s}g\|_{r'} \leq C|t-s|^{-\beta} \|g\|_r$$

where $1 \leq r \leq 2$ and where β is determined by $\beta = \theta n/2$ and $\frac{1}{r} = \frac{1-\theta}{2} + \frac{\theta}{1}$ $(0 \leq \theta \leq 1)$.
Observe that for our purpose, the critical index is the case $\beta = 1$, so $\theta = 2/n$ and $r = \dfrac{2n}{n+2} < 2$.
In other words, we have the boundary case as for $0 \leq \theta < 2/n$, we have $\dfrac{2n}{n+2} < r \leq 2$.
In the same regime as above, consider the fractional integral

$$I_{1-\beta}h(t) = C\int_{\mathbb{R}} |t-s|^{-\beta} h(s)\, ds$$

The H-L-S inequality (Theorem 7.5) says that $I_{1-\beta}$ is a bounded map from L^p to L^q, where $\frac{1}{q} = \frac{1}{p} - 1 + \beta$. In particular, if we take $p = \dfrac{2}{2-\beta}$, setting $q = 2/\beta = p'$, and $\left\|I_{1-\beta}h\right\|_{p'} \leq C\|h\|_p$.
So we see that the operator with kernel K_{t-s} is L^p-bounded by $|t-s|^{-\beta}$ for $1 \leq p \leq 2$, from which, if we consider (39), we get a sort of fractional integral. Again, Theorem 7.5 shows this fractional is L^q-bounded. There's an issue here: the L^p-spaces on $\mathbb{R}^n$ integral and t integral might not match, so we might not be able to simply take a single norm on $\mathbb{R}^{n+1}$. To solve this problem, we'll introduce the notion of mixed norm space.

Definition 11.9. The mixed norm space $L^p_t(L^r_x)(\mathbb{R}^{n+1})$ is defined as the space of function $f(x,t)$ on $\mathbb{R}^{n+1}$ so that

$$\|L^p_t(L^r_x)\|_: = \left(\int_{-\infty}^{\infty}\left(\int_{\mathbb{R}^n} |f(x,t)|^r\, dx\right)^{p/r} dt\right)^{1/p}$$

is finite. Its dual is $L^{p'}_t(L^{r'}_x)$.

The result of the theorem will then follow from the more general version, where we take X in Equation (38) to be $L^p_t(L^r_x)$: the operator

$$R : L^p_t(L^r_x) \to L^2$$

as defined above is bounded for all $r = \dfrac{2}{1+\theta}$, $p = \dfrac{4}{4-n\theta} = \dfrac{2}{2-\beta}$. The proof follows the same argument as above: keeping the same notations,

$$\begin{aligned}
\|Rf\|^2 &= \iint_{\mathbb{R}^{n+1}} f(x,t) \iint_{\mathbb{R}^{n+1}} K_{t-s}(x-y)\overline{f(y,s)}\, dy\, ds\, dx\, dt \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathbb{R}^n} f(x,t)\left(T_{t-s}\overline{f(\cdot,s)}\right)(x)\, dx\, ds\, dt \\
&\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \|f(\cdot,t)\|_{L^r_x} \left\|T_{t-s}f(\cdot,s)\right\|_{L^{r'}_x}\, ds\, dt && \text{(Hölder)}
\end{aligned}$$

$$\lesssim \int_{-\infty}^{\infty} \|f(\cdot,t)\|_{L_x^r} \int_{-\infty}^{\infty} |t-s|^{-\beta} \|f(\cdot,s)\|_{L_x^r} \, ds\, dt$$

Now take $\|f(\cdot,t)\|_{L_x^r}$ to be $h(t)$ (as in the fractional integral. Thus

$$\begin{aligned}
&\approx \int_{-\infty}^{\infty} h(t)\,(I_{1-\beta}h)(t)\,dt \\
&\lesssim \|h\|_{L_t^p} \left\| I_{1-\beta} \right\|_{L_t^{p'}} && \text{(Hölder)} \\
&\lesssim \|h\|_{L_t^p}^2 = \|f\|_{L_t^p(L_x^r)}^2
\end{aligned}$$

This completes the proof.

□

References

[1] Stein, Elias; *Singular Integral Operator and Differentiability Properties of Functions*, 1970

[2] Stein, Elias; *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, 1993