APPLICABLE NONLINEAR ANALYSIS
Volume 1 (2024), No. 2, 110–159
https://doi.org/10.69829/apna-024-0102-ta01
Tulipa Opera Scholarum

CAUCHY-SZEGÖ PROJECTIONS AND RELATED TOPICS

DER-CHEN CHANG[1,2], XINFENG WU[3,*], AND JEN-CHIH YAO[4,5]

[1] *Department of Mathematics and Statistics, Georgetown University, Washington DC, 20057, USA*
[2] *Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, New Taipei City 242, Taiwan, ROC*
[3] *Department of Mathematics, China University of Mining and Technology, Beijing 100083, P. R. China*
[4] *China Medical University, Taichung, 406040, Taiwan, ROC*
[5] *Academy of Romanian Scientists, 50044 Bucharest, Romania*

ABSTRACT. In this survey article, we give a comprehensive review of Calderón-Zygmund operators from the point of view of Cauchy-Szegö projections and the sharp estimates of the operators in Hardy spaces. Cauchy-Szegö projections is closed related to the Hilbert transform which is a typical example of the so-called "first generation" of singular integral operators and has been studied by mathematicians for many years, We started with the case in unit disk $\mathbb{D}^1$ in $\mathbb{C}^1$ and then move to the unbounded unit ball $\mathbb{B}^{n+1}$ in $\mathbb{C}^{n+1}$. Analysis on $\mathbb{R}^n$ and $\mathbf{H}_n$ are quite different. We try to explain the idea behind it carefully.

Keywords. Poisson integral, Hilbert transform, Cauchy-Szegö projection, Calderón-Zygmund operators, Heisenberg group, Hardy spaces.

1. INTRODUCTION

The theory of singular integrals operators (CZO) introduced by Calderón and Zygmund [3] as part of the theory of elliptic partial differential equations, has seen many extensions to different settings. Remaining within $\mathbb{R}^n$ as the ambient space, the variations introduced involve the following aspects, possibly also combined together:

(a) replace the standard dilations, *i.e.,* scalar multiplications, with non-isotropic ones, which will be explained in Section 6;

(b) distinguish between a "*global*" theory and a "*local*" one, which will be explained in Section 10;

(c) allow multi-parameter dilations. This will be explained in a forthcoming paper.

The basic property that is common to all these types of singular integral operators is L^p-boundedness for $1 < p < \infty$ and *failure* of L^p-boundedness, in general, for other values of p.

Hardy spaces H^p enter into this picture as the natural substitutes of L^p with $0 < p \leq 1$, allowing positive results about $H^p \to H^p$ and $H^p \to L^p$ boundedness of singular integrals for these values of p. The point is that each of the classes of CZO mentioned above admits its own Hardy spaces, so that, whenever a new class of CZO is introduced, it is natural to ask what are its Hardy spaces.

In this paper, we study the Cauchy-Szegö projections and the sharp estimates of the operators in Hardy spaces. Cauchy-Szegö projections is closed related to the Hilbert transform which is a typical example of the so-called "first generation" of singular integral operators and has been studied by mathematicians for many years, We started with the case in unit disk $\mathbb{D}^1$ in $\mathbb{C}^1$ and then move to the

*Corresponding author.
E-mail addresses: chang@georgetown.edu (D.-C. Chang), wuxf@cumtb.edu.cn (X. Wu), yaojc@math.nsysu.edu.tw (J.-C. Yao)

2020 Mathematics Subject Classification: Primary: 35H05, 35H20, 33C05; Secondary: 58I99, 22E25.
Accepted: September 18, 2024.

unbounded unit ball $\mathbb{B}^{n+1}$ in $\mathbb{C}^{n+1}$. The boundary of $\mathbb{B}^{n+1}$ can be identified as the Heisenberg group $\mathbf{H}_n$ which is the simplest noncommutative nilpotent group. In this case, the dilation structure is different which is sort of "second generation" of singular integral operators. Hence, the geometry on $\mathbf{H}_n$ has significant different with $\mathbb{R}^n$. Hence, analysis on $\mathbf{H}_n$ is much more delicate. Let us start with the upper half plane $\mathbb{R}^2_+$.

2. The Cauchy-Szegö Projection in a Reasonable Domain in $\mathbb{C}^1$

The Cauchy-Szegö projection is one of the canonical integrals arising in several complex variables in $\mathbb{C}^{n+1}$ with reproducing property. This is even happened in the case $n = 0$, the complex plane. Let us give a quick review this case [16]. Let Ω be a bounded smooth domain in the complex plane $\mathbb{C}^1$ and u is an arbitrary smooth function on the boundary $\partial\Omega$. The Cauchy kernel defines a holomorphic functions U on Ω by

$$U(w) \;=\; \frac{1}{2\pi i}\int_{z\in\partial\Omega} \frac{u(z)}{z-w}\,\dot{\xi}(z)\,d\sigma(z), \qquad w\in\Omega, \tag{2.1}$$

or more briefly

$$\mathbf{K}(u)(w) \;=\; U(w) \;=\; \int_{z\in\partial\Omega} K(w,z)u(z)\,dz, \qquad w\in\Omega. \tag{2.2}$$

Here $dz = \dot{\xi}(z)\,d\sigma(z)$, $z = \xi(s)$ is the arc length parametrization of $\partial\Omega$, and $d\sigma$ is the linear measure (arc length) on $\partial\Omega$.

The integral is non-singular so long as $w \in \Omega$, but it ceases to exist as a Lebesgue integral if $w \in \partial\Omega$. However, $\mathbf{K}(u)$ admits a continuous extension up to $\bar{\Omega}$ (also denoted by $\mathbf{K}(u)$). At the same time the following limit exists and is finite.

$$\lim_{\varepsilon\to 0^+}\int_{z\in\partial\Omega, |z-w|>\varepsilon} K(w,z)u(z)\,dz \;=:\; P.V.\int_{z\in\partial\Omega} K(w,z)u(z)\,dz \tag{2.3}$$

for $w \in \partial\Omega$.

Formula (2.2) gets modifies to the Plemelj formula

$$\mathbf{K}(u)(w) \;=\; \frac{1}{2}u(w) + P.V.\int_{z\in\partial\Omega} K(w,z)u(z)\,dz, \qquad w\in\Omega. \tag{2.4}$$

The kernel $K(z,w)$ is the *Cauchy kernel*:

$$K(z,w) = \frac{1}{2\pi i}\,\frac{1}{z-w}\dot{\xi}(z).$$

An important point is that the deleted neighborhood around w in (2.3) is symmetric. If it were, *e.g.,* proportionately much longer to one side than to the other, then the limit might fail to exist or the number $\frac{1}{2}$ in (2.4) might have to be modified.

The restriction of $\mathbf{K}(u)$ to $\partial\Omega$ which we still call it $\mathbf{K}(u)$ satisfies L^p and Lipschitz estimates

$$\|\mathbf{K}(u)\|_{L^p(\partial\Omega)} \;\le C_p\|u\|_{L^p(\partial\Omega)}, \qquad 1<p<\infty. \tag{2.5}$$

Hence, $\mathbf{K}$ can be extended to a bounded operator

$$\mathbf{K}:\; L^p(\partial\Omega) \;\to\; L^p(\partial\Omega),$$

which associates to $u \in L^p(\partial\Omega)$ the boundary values of a function holomorphic in Ω.

When Ω is the unit disc and $p = 2$, then $\mathbf{K}(u)$ is obtained from u by chopping off the negative terms in the Fourier series of u, *i.e.,* $\mathbf{K}(u)$ is the *orthogonal projection* of u on the subspace of $L^2(\partial\Omega)$ consisting of boundary values of functions holomorphic in Ω. This is true only if Ω is a disc. For any other Ω, $\mathbf{K}$ is still a projection, *i.e.,* $\mathbf{K}^2 = \mathbf{K}$ (because the Cauchy kernel reproduces holomorphic functions) but it is no longer orthogonal. The purpose of this series of lectures is to explain some basic idea of the Calderón-Zygmund singular integral operators via the operator $\mathbf{K}$. Once we have the kernel

K of the operator $\mathbf{K}$, then we may study sharp estimates of the operator $\mathbf{K}$ on Hardy spaces H^p for $0 < p < \infty$.

3. Poisson Integrals and the Hilbert Transform

Let us start with the complex plane $\mathbb{C}^1 \cong \mathbb{R}^2$. Suppose that $f \in L^p(\mathbb{R})$, $1 < p < \infty$, and consider the function

$$F(z) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{f(t)}{t-z}dt, \qquad z = x+iy \tag{3.1}$$

in the upper half plane $\mathbb{R}^2_+ =: \{x+iy \in \mathbb{R}^2 : y = \text{Im}(z) > 0\}$. Since $\frac{1}{t-z}$, as a function of t, belongs to $L^q(\mathbb{R}^1)$ for every $q > 1$, the integral is finite by Hölder's inequality and therefore, $F(z)$ is well-defined and is holomorphic in $\mathbb{R}^2_+$. In fact, if the support of f is compact, then

$$F(z) = \frac{1}{2\pi i}\int_{-N}^{+N} \frac{f(t)}{t-z}dt$$

and

$$\begin{aligned}\frac{F(z+w)-F(z)}{w} &= \frac{1}{2\pi i}\int_{-N}^{+N} f(t)\frac{1}{w}\left[\frac{1}{t-z-w} - \frac{1}{t-z}\right]dt \\ &= \frac{1}{2\pi i}\int_{-N}^{+N} \frac{f(t)}{(t-z-w)(t-z)}dt.\end{aligned}$$

But, as $w \to 0$, the kernel $\frac{1}{(t-z-w)(t-z)} \to \frac{1}{(t-z)^2}$ for $|t| \le N$, hence $F'(z) = \frac{1}{2\pi i}\int_{-N}^{+N} \frac{f(t)}{(t-z)^2}dt$ exists and $F(z)$ is holomorphic in $\mathbb{R}^2_+$.

For general L^p function f, we consider the truncations

$$f_N(t) = \begin{cases} f(t) & \text{if} \quad |t| \le N \\ 0 & \text{elsewhere;} \end{cases}$$

then the functions $F_N(z) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{f_N(t)}{t-z}dt$ are all holomorphic and converge to $F(z)$ uniformly.

Now, with $z = x+iy$ and $y > 0$, we decompose the kernel $\frac{1}{i(t-z)}$ into its real and imaginary parts. Then we can rewrite $F(z)$ into the following form:

$$\begin{aligned} F(z) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(t)\frac{y}{(x-t)^2+y^2}dt + \frac{i}{2\pi}\int_{-\infty}^{+\infty} f(t)\frac{x-t}{(x-t)^2+y^2}dt \\ &= \frac{1}{2}(f * \mathcal{P}_y)(x) + \frac{i}{2}(f * \mathcal{Q}_y)(x) \end{aligned} \tag{3.2}$$

where $\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{x^2+y^2}$ is the *Poisson kernel* in $\mathbb{R}^2_+$, and

$$\mathcal{Q}_y(x) = \frac{1}{\pi}\frac{x}{x^2+y^2} \tag{3.3}$$

is called the *conjugate Poisson kernel* in $\mathbb{R}^2_+$.

The integral

$$u(x,y) = (f * \mathcal{P}_y)(x) = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(t)\frac{y}{(x-t)^2+y^2}dt$$

is the *Poisson integral* of f. We may generalize it to $\mathbb{R}^{n+1}$:

$$\mathcal{P}_1(x) = C_n\frac{1}{(1+|x|^2)^{\frac{n+1}{2}}} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}}\frac{1}{(1+|x|^2)^{\frac{n+1}{2}}}.$$

For $y > 0$, we define

$$\mathcal{P}_y(x) = \mathcal{P}(x,y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \frac{y}{(y^2+|x|^2)^{\frac{n+1}{2}}}.$$

It follows that

$$\int_{-\infty}^{+\infty} \mathcal{P}_1(x)\,dx = \frac{2}{\pi}\int_0^\infty \frac{dr}{1+r^2} = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} d\theta = 1.$$

Changing variable $r = \tan\theta$ implies that $1+\tan^2\theta = \frac{1}{\cos^2\theta}$ and $dr = \frac{d\theta}{\cos^2\theta}$.

Theorem 3.1. *Let $f(x)$ be an L^p function on $\mathbb{R}^n$ for some $1 \le p \le \infty$, and let $u(x,y) = f * \mathcal{P}_y(x)$ be its Poisson integral for $(x,y) \in \mathbb{R}^{n+1}_+$. Then we have*

$$\Delta u(x,y) = \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

and $\lim_{y\to 0^+} u(x,y) = f(x)$ for almost all $x \in \mathbb{R}^n$. In addiiton,

$$\sup_{y>0} \left|u(x,y)\right| \le M_{HL}(f)(x) \qquad \text{for every } x \in \mathbb{R}^n.$$

Here $M_{HL}(f)(x)$ is the Hardy-Littlewood maximal function of the function f.

Furthermore, we also have L^p norm convergence of $u(x,y)$ to $f(x)$ if $f \in L^p(\mathbb{R}^n)$ and $1 \le p < \infty$.

It can be shown that $(f * \mathcal{P}_y)(x) \to f(x)$ for almost every x, as $z = x+iy \to x$ non-tangentially, even if $f \in L^\infty$. Under the hypothesis that f is real valued,

$$u(x,y) = \text{Re}\left[F(z)\right].$$

Here $F(z)$ is defined in (3.1). We now consider the integral

$$v(x,y) = (f * \mathcal{Q}_y)(x) = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(t)\frac{x-t}{(x-t)^2+y^2}dt, \tag{3.4}$$

which is called the *conjugate Poisson integral* of f and we have the following theorem.

Theorem 3.2. *For any $f \in L^p(\mathbb{R})$, $1 < p < \infty$, and almost every x, the conjugate Poisson integral of f tends, as $z \to x$ non-tangentially, to a finite limit.*

The *Hilbert transform* is a typical example of the so-called "first generation" Calderón-Zygmund operators. On $\mathbb{R}^1$, set

$$Hf(x) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{f(y)}{x-y}dy. \tag{3.5}$$

Let us look at the definition of the Hilbert transform (3.5). From the very beginning, there is a problem. The above integral need not converge absolutely, so $Hf(x)$ need not be defined. Indeed if f is continuous at x_0 and $f(x_0) \neq 0$, then

$$\int_{-\infty}^{+\infty} \frac{|f(y)|}{|x_0-y|}dy = \infty.$$

However, the kernel $K(x) = \frac{1}{\pi}\frac{1}{x}$ is odd. If we assume that $f \in C_0^1(\mathbb{R})$ and interpret the possibly divergent integral as a principal-valued integral:

$$Hf(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|x-y|>\varepsilon} \frac{f(y)}{x-y}dy = \frac{1}{\pi}\text{p.v.}\int \frac{f(y)}{x-y}dy. \tag{3.6}$$

Fix a $\varepsilon > 0$, we may write the integral as follows:

$$\begin{aligned}\frac{1}{\pi}\int_{|x-y|>\varepsilon}\frac{f(y)}{x-y}dy &= \frac{1}{\pi}\int_{\varepsilon<|x-y|\leq 1}\frac{f(y)}{x-y}dy + \frac{1}{\pi}\int_{|x-y|>1}\frac{f(y)}{x-y}dy \\ &= I_1 + I_2.\end{aligned}$$

For term I_2, we certainly have

$$\left|\int_{|x-y|>1}\frac{f(y)}{x-y}dy\right| < \infty,$$

since the support of f is compact. We may rewrite term I_1 as follows:

$$\int_{\varepsilon<|x-y|\leq 1}\frac{f(y)}{x-y}dy = \int_{\varepsilon<|x-y|\leq 1}\frac{f(y)-f(x)}{x-y}dy$$

since

$$\int_{\varepsilon<|x-y|\leq 1}\frac{1}{x-y}dy = 0.$$

By the smoothness assumption on f,

$$\frac{|f(y)-f(x)|}{|x-y|}$$

is uniformly bounded and hence

$$\int_{|x-y|\leq 1}\frac{|f(y)-f(x)|}{|x-y|}dy < \infty.$$

Therefore,

$$\begin{aligned}\int_{\varepsilon<|x-y|\leq 1}\frac{f(y)-f(x)}{x-y}dy &= \int_{|x-y|\leq 1}\frac{f(y)-f(x)}{x-y}dy - \int_{|x-y|\leq\varepsilon}\frac{f(y)-f(x)}{x-y}dy \\ &= \int_{|x-y|\leq 1}\frac{f(y)-f(x)}{x-y}dy + \mathcal{O}(\varepsilon).\end{aligned}$$

Then the limit exists uniformly as $\varepsilon \to 0$.

If the kernel $\frac{1}{x-y}$ were replaced by its absolute value in (3.5) then the principal-value limit would fail to exist whenever $f(x) \neq 0$. The key point here is that the very definition of the Hilbert transform depends on *cancelation* in the integral. Indeed such cancelation lies at the heart of the entire theory of singular integral operators.

Differentiability of f is not really required for this limiting procedure; it would suffice for instance to have a Hölder condition, *i.e.,*

$$|f(x) - f(y)| \leq C \cdot |x-y|^{\delta} \quad \text{for some} \quad \delta > 0.$$

Then

$$\begin{aligned}|Hf(x)| &= \frac{1}{\pi}\left|\int_0^\infty \frac{f(x+t)-f(x-t)}{t}dt\right| \leq \frac{1}{\pi}\int_0^\infty \frac{|f(x+t)-f(x-t)|}{|t|}dt \\ &= \frac{1}{\pi}\int_0^\varepsilon + \frac{1}{\pi}\int_\varepsilon^\infty = I_1 + I_2\end{aligned}$$

and, by the Hölder condition,

$$I_1 \leq C' \cdot \int_0^\varepsilon |t|^{\delta-1}dt < \infty, \quad \text{since} \quad \delta > 0,$$

while I_2 is finite, by Hölder's inequality, for $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Now we link the relation between the Hilbert transform and conjugate Poisson integral. We have the following important theorem.

Theorem 3.3. *For any $f \in L^p(\mathbb{R})$, $1 < p < \infty$, the Hilbert transform $Hf(x)$ exists and is finite a.e. Moreover, it is equal to the limit of the conjugate Poisson integral of f at every point of the Lebesgue set $\mathcal{L}_f$ of f, i.e,*

$$\lim_{y \to 0^+} \left(\int_{-\infty}^{+\infty} f(t) \frac{x-t}{(x-t)^2 + y^2} dt - \int_{|x-t|>y} \frac{f(t)}{x-t} dt \right) = 0 \tag{3.7}$$

if $x \in \mathcal{L}_f$.

The proof of the above theorem based on the equivalence of non-tangential bounded and the existence of non-tangential limit of harmonic functions on the upper half plane. We refer the readers to the book of Stein and Weiss [21] for a good reference

Remarks.
(1). The above theorem does not apply to L^∞. Let $f(x) \equiv 1$. Then $Hf(x)$ need not exist. However, if we assume further that f is bounded and that

$$\int_{|x|>1} \frac{|f(x)|}{|x|} dx \ < \ +\infty,$$

then it is easily seen that the $Hf(x)$ exists and is finite almost everywhere.
(2). The above theorems also excludes the case $p = 1$ also. The example $f(x) = \chi_{(a,b)}(x)$ shows that Hf is not necessary integrable if $f \in L^1(\mathbb{R})$. In fact, for $x \notin (a, b)$, we have

$$Hf(x) = \frac{1}{\pi} \int_a^b \frac{dt}{x-t} = \frac{1}{\pi} \log \left| \frac{a-x}{b-x} \right|.$$

If $x \in (a, b)$ and $\varepsilon > 0$ is sufficiently small, then we have

$$Hf(x) \ = \ \frac{1}{\pi} \lim_{\varepsilon \to 0} \left\{ \int_a^{x-\varepsilon} \frac{dt}{x-t} + \int_{x+\varepsilon}^b \frac{dt}{x-t} \right\} = \frac{1}{\pi} \log \left| \frac{a-x}{b-x} \right|.$$

We note that as $|x| \to \infty$,

$$\log \left| \frac{a-x}{b-x} \right| = \log \left| 1 + \frac{b-a}{b-x} \right| \sim \frac{|b-a|}{|x-b|} \sim \frac{|b-a|}{|x|}.$$

This shows that $Hf \notin L^1(\mathbb{R})$.
(3). If $f \in L^1(\mathbb{R}^n)$, then for every $\alpha > 0$, then the operator H is of weak type $(1, 1)$, *i.e.,*

$$\Big| \{ x \in \mathbb{R} : \ Hf(x) > \alpha \} \Big| \ \le \ \frac{C_n}{\alpha} \|f\|_{L^1}.$$

4. L^2 Estimate for the Operator H

Indeed, we may define the Hilbert transform operator

$$H : f \to H(f), \ \text{ for } \ f \in L^p(\mathbb{R}^n), \ \ 1 < p < \infty$$

especially for $p = 2$. As usual, denote $\text{sgn}(x) = 1$ for $x > 0$ and -1 for $x < 0$. Then one may show that

$$\widehat{(Hf)}(x) = -i(\text{sgn}(x))\widehat{f}(x) \tag{4.1}$$

for $f \in L^2(\mathbb{R}^1)$. In particular,

$$\|H(f)\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}, \qquad \text{for all} \quad f \in L^2(\mathbb{R}^1). \tag{4.2}$$

From Plancherel's theorem and the density property of $C_0^1(\mathbb{R})$ in $L^2(\mathbb{R})$, it follows that H has a unique extension to a bounded, linear operator from $L^2(\mathbb{R})$ to itself.

In order to prove (4.2), we may consider the "truncate" transform:

$$(Hf)_{\varepsilon,N}(x) = \int_{-\infty}^{+\infty} f(t)K_{\varepsilon,N}(x-t)dt = (f * K_{\varepsilon,N})(x)$$

with the "truncate" kernel:

$$K_{\varepsilon,N}(x) = \begin{cases} \frac{1}{x} & \text{if} \quad \varepsilon < |x| < N \\ 0 & \text{elsewhere.} \end{cases}$$

Then $(Hf)_{\varepsilon,N}(x)$ is the convolution of a square integrable function f with an integrable kernel $K_{\varepsilon,N}$. It follows that $(Hf)_{\varepsilon,N} \in L^2(\mathbb{R}^1)$ for each fixed ε and N. Hence by basic property of the Fourier transform, one has

$$\widehat{(Hf)_{\varepsilon,N}}(x) = \widehat{f}(x) \cdot \widehat{K_{\varepsilon,N}}(x) \qquad \text{for almost} \quad x \in \mathbb{R}^1.$$

But,

$$\begin{aligned} \widehat{K_{\varepsilon,N}}(x) &= \int_{\varepsilon<|x|<N} \frac{1}{t} e^{-2\pi ixt} dt = \int_{\varepsilon}^{N} \frac{e^{-2\pi ixt} - e^{2\pi ixt}}{t} dt \\ &= -2i \int_{\varepsilon}^{N} \frac{\sin 2\pi xt}{t} dt = -\frac{2i}{\pi}(\text{sgn}(x)) \int_{\varepsilon|x|}^{N|x|} \frac{\sin w}{w} dw, \end{aligned}$$

where the last integral is uniformly bounded and converges to

$$\int_0^\infty \frac{\sin w}{w} dw = \frac{\pi}{2}$$

as $\varepsilon \to 0$ and $N \to \infty$. Hence we know that, for each x, ε and N,
(a). $|\widehat{K_{\varepsilon,N}}(x)| \leq C$;
(b). $\widehat{K_{\varepsilon,N}}(x) \to (-i)(\text{sgn}(x))$ as $\varepsilon \to 0$ and $N \to \infty$.

From (a) and (b), it follows that

$$\left\| \widehat{(Hf)_{\varepsilon,N}} \right\|_{L^2(\mathbb{R})} \leq C \cdot \|\widehat{f}\|_{L^2(\mathbb{R})}.$$

Hence, by Plancherel's theorem, we have

$$\|(Hf)_{\varepsilon,N}\|_{L^2(\mathbb{R})} \leq C \cdot \|f\|_{L^2(\mathbb{R})}.$$

By the Lebesgue dominated convergence theorem, one concludes that

$$\widehat{(Hf)_{\varepsilon,N}}(\xi) \to (-i)(\text{sgn}(\xi))\widehat{f}(\xi) \qquad \text{in} \quad L^2(\mathbb{R}),$$

since

$$\left\| \widehat{(Hf)_{\varepsilon,N}}(\xi) + (-i)(\text{sgn}(\xi))\widehat{f}(\xi) \right\|_{L^2(\mathbb{R})} = \left\| \left[\widehat{K_{\varepsilon,N}}(x) + (-i)(\text{sgn}(x)) \right] \widehat{f}(x) \right\|_{L^2(\mathbb{R})} \to 0$$

as $\varepsilon \to 0$ and $N \to \infty$. Therefore,

$$(Hf)_{\varepsilon,N} \to H(f),$$

in L^2 norm, for some square integrable function $H(f)$ such that

$$\widehat{(Hf)}(x) = -i(\text{sgn}(x))\widehat{f}(x).$$

Moreover, from the above identity, we see that $\|\widehat{Hf}\|_{L^2(\mathbb{R})} = \|\widehat{f}\|_{L^2(\mathbb{R})}$. Therefore, by Plancherel's theorem again, we have

$$\|Hf\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}.$$

Thus, (4.2) is true.

This enables us to see that the conjugate Poisson integral of an L^2 function coincides a.e. with the Poisson integral of its Hilbert transform.

Corollary 4.1. *If* $f \in L^2(\mathbb{R})$ *then for* $y > 0$,

$$(f * \mathcal{Q}_y)(x) = (Hf * \mathcal{P}_y)(x) \qquad \text{a.e.} \tag{4.3}$$

Proof. Formula (4.3) means that

$$\frac{1}{\pi}\int_{-\infty}^{+\infty} f(x-t)\frac{t}{t^2+y^2}dt = \frac{1}{\pi}\int_{-\infty}^{+\infty} Hf(x-t)\frac{y}{t^2+y^2}dt \tag{4.4}$$

holds almost everywhere. Since both sides of (4.4) have the same Fourier transform which equals

$$-i\mathrm{sgn}(\xi)e^{-2\pi|y\xi|}\widehat{f}(\xi).$$

The proof of the corollary is therefore complete. □ □

Note that if $H^2 f = H(Hf)$ then, by (4.1), $H^2 = -\mathbf{I}$, where $\mathbf{I}$ is the identity operator in $L^2(\mathbb{R})$. This fact, together with

$$\|Hf\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})},$$

we know that H is a *unitary operator* in $L^2(\mathbb{R})$.

Example 4.2. Riesz transforms.

Since in $\mathbb{R}^1$ the function $\frac{1}{x}$ may be written as

$$\frac{x}{|x|^2} = \frac{x}{|x|^{n+1}},$$

with $n = 1$ which is the dimension of $\mathbb{R}^1$. Let $n \geq 2$ and $1 \leq j \leq n$. For $f \in C_0^1(\mathbb{R}^n)$ set

$$R_j f(x) = C_n \mathrm{p.v.}\int_{\mathbb{R}^n} f(y)\frac{x_j - y_j}{|x-y|^{n+1}}dy$$

where

$$C_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}}.$$

Again the limit exists for all $f \in C_0^1$ because of the cancelation property of the kernel $C_n\frac{x_j}{|x|^{n+1}}$. The operators R_j, $j = 1, \dots, n$ are called *Riesz transforms* of f. □

In fact, the Hilbert transform and the Riesz transforms are special cases for the following singular integral operators:

$$f \mapsto Tf,$$

where

$$\begin{aligned} Tf(x) &= \mathrm{p.v.}\int_{\mathbb{R}^n} K(x-y)f(y)dy \\ &= \lim_{\varepsilon\to 0}\int_{|x-y|>\varepsilon} K(x-y)f(y)dy. \end{aligned} \tag{4.5}$$

Here the kernel K satisfies the following Calderón-Zygmund conditions:

$$|K(x)| \leq \frac{B}{|x|^n}, \qquad \forall\ x \neq 0; \tag{4.6}$$

$$|K(x) - K(y)| \leq \frac{B|x-y|^\gamma}{|x|^{n+\gamma}}, \qquad 0 < \gamma \leq 1, \quad |x-y| < \frac{1}{2}|x|; \tag{4.7}$$

and

$$\int_{r_1\le|x|<r_2} K(x)dx = 0, \qquad \forall\ 0<r_1<r_2<\infty. \tag{4.8}$$

This section deals with the fundamental properties of the singular integral operators given by convolution with kernels with singularities at the origin and at the infinity. The theory for such singular integrals was developed by Calderón and Zygmund in the 1950s [3]. We say that these kinds of operators are first generation singular integrals. Later, L. Hörmander [14] introduced the following condition to replace condition (4.7):

$$\int_{|x|\ge 2|y|} |K(x-y)-K(x)|dx \le B, \qquad \forall\ |y|>0. \tag{4.9}$$

It is easy to see that condition (4.7) implies condition (4.9). By (4.7), we know that

$$|K(x-y)-K(x)| \le \frac{B|x-y-x|^\gamma}{|x-y|^{n+\gamma}} = \frac{B|y|^\gamma}{|x-y|^{n+\gamma}}.$$

But in the region $\{x,y\in\mathbb{R}^n : |x|\ge 2|y|\}$, one has

$$|x-y| \ge |x|-|y| \ge |x|-\frac{1}{2}|x| = \frac{1}{2}|x|.$$

It follows that

$$|K(x-y)-K(x)| \le \frac{B|y|^\gamma}{|x-y|^{n+\gamma}} \le \frac{\widetilde{B}|y|^\gamma}{|x|^{n+\gamma}}.$$

Therefore,

$$\int_{|x|\ge 2|y|} |K(x-y)-K(x)|dx \le \widetilde{B}|y|^\gamma \int_{|x|\ge 2|y|} |x|^{-n-\gamma}dx = B.$$

One of the corner stones to obtain the L^p, $1<p<\infty$ estimates for the operator T is the following weak type $(1,1)$ estimate for T.

Theorem 4.3. *Let T be an operator defined by (4.5) with kernel K satisfying conditions (4.6), (4.8), and (4.9). We assume that T is bounded on $L^2(\mathbb{R}^n)$,*

$$\|Tf\|_{L^2(\mathbb{R}^n)} \le C_2\|f\|_{L^2(\mathbb{R}^n)}.$$

Then T can be extended as an operator of weak type $(1,1)$ and the operator norm C' depends only on the constant B in (4.9) and C_2:

$$\big|\{x\in\mathbb{R}^n : |Tf(x)|>\alpha\}\big| \le \frac{C'}{\alpha}\|f\|_{L^1(\mathbb{R}^n)}, \qquad \text{for all} \quad f\in L^1\cap L^2. \tag{4.10}$$

The proof of the above theorem is highly nontrivial which relies on the famous result is so-called the *Calderón-Zygmund decomposition.* Readers can consult the books by Stein [19] and [20].

Remark.

Calderón [2] raises a question regarding mapping properties of the Cauchy integral

$$C_\Gamma(f)(w) = \int_{C_\Gamma} \frac{f(z)}{z-w}\,dz$$

namely, to determine the rectifiable Jordan curve Γ for which C_Γ gives ride a bounded operator on $L^2(\Gamma)$. This was solved by G. David and J. Journé in 1984 [8] and G. David, J. Journé, S. Semmes in 1985 who showed that C_Γ is bounded on $L^2(\Gamma)$ when Γ satisfies

$$\mu\big(\Gamma\cap B(z_0;r)\big) \le C\,r$$

for every $z_0 \in \mathbb{C}$, $r > 0$ and some constant C. This opened up a large study of (which was called then) Ahlfors regularity by David and Semmes [9]. When $d = 1$, this condition appeared in Ahlfors's famous paper [1].

Let $E \subset \mathbb{R}^n$ be a closed set, not reduced to a point, and μ be a (positive Borel) measure supported on E. The set E is called *Ahlfors-regular* of dimension d if there exists a constant $C \geq 1$ such that

$$C^{-1} r^d \;\leq\; \mu\big(E \cap B(x;r)\big) \;\leq\; C \cdot r^d$$

for all $x \in E$ and $0 < r < \text{diam}(E)$ and there is a constant C such that

$$C^{-1} \mu(A) \;\leq\; H^d(E \cap A) \;\leq\; C\mu(A)$$

for all Borel sets $A \subset \mathbb{R}^n$. Here H^d denote the d-dimensional Hausdorff measure.

5. Return to the Cauchy-Szegö Projection

There is a second kernel naturally associated with Ω, the *Szegö kernel* $S(z,w)$ which serves the same purpose

$$U(w) \;=\; \int_{z \in \partial\Omega} S(z,w) u(z)\, d\sigma(z), \qquad w \in \Omega \tag{5.1}$$

and which arises from considering the orthogonal projection

$$\mathbf{S} :\; L^2(\partial\Omega, d\sigma) \;\to\; \mathcal{H}^2(\partial\Omega). \tag{5.2}$$

Here $\mathcal{H}^2(\partial\Omega)$ is the closed subspace of $L^2(\partial\Omega)$ of boundary values holomorphic functions in Ω For any $u \in L^2(\partial\Omega)$,

$$\mathbf{S}(u)(w) \;=\; \int_{z \in \partial\Omega} S(z,w) u(z)\, d\sigma(z), \qquad w \in \Omega, \tag{5.3}$$

from which (5.3) follows since, in (5.1), $U = \mathbf{S}(u)$.

Now we may ask an important question: *Does* $S(z,w) = K(z,w)$? In general, the answer is "*No*". There is a key reason why $S(z,w)$ must be more involves than $K(z,w)$. If Ω is simply connected, then the Riemann mapping function can be immediately and explicitly obtained from $S(z,w)$.

The orthogonal projection $\mathbf{S} : L^2(\partial\Omega) \;\to\; \mathcal{H}^2(\partial\Omega)$ is represented by the Szegö kernel $S(w,z)$ of Ω:

$$\mathbf{S}(u)(w) \;=\; \int_{z \in \partial\Omega} S(w,z) u(z)\, d\sigma(z), \qquad w \in \Omega.$$

We have identified $\mathbf{S}(u)$, which is a function in $\mathcal{H}^2(\partial\Omega)$, with its unique holomorphic extension to Ω, and we shall repeatedly perform this identifications. Now if $\{\phi_k\}$ is an arbitrary complete orthonormal system in $\mathcal{H}^2(\partial\Omega)$ (*i.e.,* in the measure $d\sigma$ of $\partial\Omega$) we have

$$S(w,z) \;=\; \sum_{k=1}^{\infty} \phi_k(w) \bar{\phi}_k(z), \qquad w \in \Omega,\;\; z \in \partial\Omega$$

and it follows that the unit disc $\mathbb{D}$ has Szegö kernel

$$S_{\mathbb{D}}(w,z) \;=\; \frac{1}{2\pi} \frac{1}{1 - w\bar{z}}, \qquad |w| < 1,\;\; |z| = 1. \tag{5.4}$$

Notice also that $S(w,z) \;=\; \bar{S}(z,w)$ are defined for $z,\; w \in \Omega$.

5.1. **The Szegö kernel $S(z, z_0)$ in terms of the projection S.** If $u \in \mathcal{H}^2(\partial\Omega)$, then on one hand, Cauchy formula gives

$$u(z_0) \;=\; \frac{1}{2\pi i}\int_{z\in\partial\Omega}\frac{u(z)}{z-z_0}\dot{\xi}(z)\,d\sigma(z), \qquad z_0\in\Omega,$$

where $\dot{\xi}(z)\,d\sigma(z) \;=\; dz$, $\dot{\xi}(z)$ is the unit tangent to $\partial\Omega$ at z and ξ is the counterclockwise arc length parametrization of $\partial\Omega$. Setting

$$\psi_{z_0}(z) \;=\; \text{conjugate of } \Big(\frac{1}{2\pi i}\frac{1}{z-z_0}\,\dot{\xi}(z)\Big)$$

and using scalar product notation $\langle\cdot,\cdot\rangle$ on $\mathcal{H}^2(\partial\Omega)$ we get $u(z_0) \;=\; \langle u, \psi_{z_0}\rangle$ or, neglecting the component of ψ_{z_0} orthogonal to $\mathcal{H}^2(\partial\Omega)$

$$u(z_0) \;=\; \langle u, \mathbf{S}(\psi_{z_0})\rangle. \tag{5.5}$$

On the other hand (5.5) gives

$$u(z_0) \;=\; \langle u(\cdot), \bar{S}(z_0,\cdot)\rangle \;=\; \langle u(\cdot), S(\cdot, z_0)\rangle. \tag{5.6}$$

Both (5.5) and (5.6) hold for *all* $u \in \mathcal{H}^2(\partial\Omega)$ and also $\mathbf{S}(\psi_{z_0})$ and $S(\cdot, z_0)$ are in $\mathcal{H}^2(\partial\Omega)$. Uniqueness shows that we have proved $S(z, z_0) \;=\; \mathbf{S}(\psi_{z_0})$ for $z \in \bar{\Omega}$.

6. The Cauchy Projection K

Now we set up another projection $\mathbf{K} : L^2(\partial\Omega) \;\to\; \mathcal{H}^2(\partial\Omega)$ that is *not orthogonal* but which is given by the Cauchy kernel. However, $\mathbf{K}$ is "closed" in orthogonal though, as comparison with its adjoint $\mathbf{K}^*$ will show. For more details, readers can consult a paper by Kerzman and Stein [16].

For any $u \in L^2(\partial\Omega)$ set

$$U(w) \;=\; \frac{1}{2\pi i}\int_{z\in\partial\Omega}\frac{u(z)}{z-w}\,dz, \qquad w\in\Omega. \tag{6.1}$$

Then, the following is known.

Theorem 6.1. *There is a unique function $\mathbf{K}(u) \in \mathcal{H}^2(\partial\Omega)$ of which $U(w)$ is the holomorphic extension. Moreover,*

$$\|\mathbf{K}(u)\|_{L^2(\partial\Omega)} \;\le\; c\cdot\|u\|_{L^2(\partial\Omega)},$$

where $c = c(\Omega)$ is independent of u.

Here "extension" means that $\|\mathbf{K}(u)(z) - U(z-\varepsilon\nu(z))\| \;\to\; 0$ when $\varepsilon \;\to\; 0$. Here $\varepsilon > 0$ and $\nu(z)$ is the outer normal to $\partial\Omega$ at z. Indeed, one allows to consider non-tangential "extension".

Moreover, the following singular integral representation is valid for any $u \in C^\infty(\partial\Omega)$, $w,\ z \in \partial\Omega$

$$\mathbf{K}(u)(w) \;=\; \frac{1}{2}u(w) + \frac{1}{2\pi i}P.V.\int_{z\in\partial\Omega}\frac{u(z)}{z-w}\,\dot{\xi}(z)\,d\sigma(z). \tag{6.2}$$

We rewrite (6.2) as

$$\mathbf{K}(u)(w) \;=\; \frac{1}{2}u(w) + P.V.\int_{z\in\partial\Omega}K(w,z)\,u(z)\,d\sigma(z) \tag{6.3}$$

with $K(w,z) \;=\; \frac{1}{2\pi i}\frac{1}{z-w}\,\dot{\xi}(z)$, the Cauchy kernel.

Being bounded in $L^2(\partial\Omega)$, $\mathbf{K}$ has a bounded adjoint $\mathbf{K}^* : L^2(\partial\Omega) \to L^2(\partial\Omega)$ defined as $\langle\mathbf{K}(u), v\rangle \;=\; \langle u, \mathbf{K}^*(v)\rangle$ and a simple application of Fubini's theorem to the Plemelj formula (6.3) shows that if $u \in C^\infty(\partial\Omega)$

$$\mathbf{K}^*(u)(w) \;=\; \frac{1}{2}u(w) + P.V.\int_{z\in\partial\Omega}\bar{K}(z,w)\,u(z)\,d\sigma(z), \quad w\in\partial\Omega \tag{6.4}$$

Comparison of $\langle \mathbf{K}(u), v\rangle = \langle u, \mathbf{K}^*(v)\rangle$ and (6.4) brings in the following theorem.

Theorem 6.2. *The kernel* $E(w,z) =: \bar{K}(z,w) - K(w,z)$ *is in* $C^\infty(\partial\Omega\times\partial\Omega)$ *if it is defined on the diagonal as*

$$E(z,z) \;=\; -\frac{1}{2\pi i}\;\text{Real part of}\;\left(\frac{\ddot{\xi}(z)}{\dot{\xi}(z)}\right).$$

Remark. The singularities of $\bar{K}(z,w)$ and $K(w,z)$ exactly cancel out, hence $E(z,z) = 0$ since $\ddot{\xi}(z)$ is orthogonal to $\dot{\xi}(z)$ (the parameter being arc length). Hence the Cauchy kernel $K(w,z)$ is closed to self-adjoint for z near w on $\partial\Omega$.

Proof. It is obvious that $E(z,w) \in C^\infty(\partial\Omega\times\partial\Omega\setminus\Sigma)$ where Σ is the diagonal of $\partial\Omega\times\partial\Omega$. By the above remark, we know that $E(z,w)$ is well-defined on Σ. Let $z = \xi(s)$ and $w = \xi(t)$ where s and t are arc length parameters, z and w are "close" with $z \neq w$ and s, t lie in a tiny interval $[a,b]$. By Taylor's Theorem, one has

$$\begin{aligned} z(s) &= \xi(s) \;=\; \xi(t) + \dot{\xi}(t)(s-t) + \frac{1}{2}\ddot{\xi}(t)(s-t)^2 + (s-t)^3\varphi(\theta) \\ w(t) &= \xi(t) \end{aligned} \tag{6.5}$$

where $a < \theta < b$. Then we have

$$z(s) - w(t) \;=\; \dot{\xi}(t)(s-t)\Big\{1 + \frac{1}{2}\frac{\ddot{\xi}(t)}{\dot{\xi}(t)}(s-t) + (s-t)^2\cdot\varphi(\theta)\Big\}. \tag{6.6}$$

Now we apply the identity $\frac{1}{1+\eta} = 1 - \eta + R(\eta)$ with $R(\eta) = \eta^2\,\varphi(\eta)$ with $\varphi \in C^\infty$ near $\eta = 0$ to obtain

$$\frac{1}{z(s)-w(t)} \;=\; \frac{1}{\dot{\xi}(t)(s-t)}\Big\{1 - \frac{1}{2}\frac{\ddot{\xi}(t)}{\dot{\xi}(t)}(s-t) + (s-t)^2\cdot\varphi(\theta)\Big\}. \tag{6.7}$$

Next we differentiate (6.5) with respect to the variable s and obtain

$$\dot{\xi}(s) \;=\; \dot{\xi}(t) + \ddot{\xi}(t)(s-t) + (s-t)^2\varphi(\theta)$$

Combining the above identity with (6.7) to get

$$\frac{\dot{\xi}(s)}{z(s)-w(t)} \;=\; \frac{1}{s-t} + \frac{1}{2}\frac{\ddot{\xi}(t)}{\dot{\xi}(t)} + (s-t)\cdot\varphi(\theta). \tag{6.8}$$

Similarly, we have

$$\frac{\dot{\xi}(t)}{z(s)-w(t)} \;=\; \frac{1}{s-t} - \frac{1}{2}\frac{\ddot{\xi}(t)}{\dot{\xi}(t)} + (s-t)\cdot\varphi(\theta). \tag{6.9}$$

Substracting (6.9) from (6.8) and (6.7) show that the singularities of $\frac{1}{s-t}$ cancel out and

$$E(w,z) \;=:\; \bar{K}(z,w) - K(w,z) \;=\; -\frac{1}{2\pi i}\left\{\text{Re}\Big(\frac{\ddot{\xi}(t)}{\dot{\xi}(t)}\Big) + (s-t)\cdot t\varphi(\theta)\right\}$$

which holds for $s,\, t \in [a,b]$, $s \neq t$. As we discussed at the beginning of the proof, $E(w,z)$ is also defined for $s = t$. Hence the proof of this theorem is therefore complete. □

Theorem 6.3. *The integral operator* $\mathbf{E} : L^2(\partial\Omega) \to L^2(\partial\Omega)$ *defined by*

$$\mathbf{E}(u)(w) \;=\; \int_{z\in\partial\Omega} E(w,z)\,u(z)\,d\sigma(z), \quad w \in \partial\Omega$$

is compact and $i\mathbf{E}$ *is in addition self adjoint. The operator* $1 - \mathbf{E}$ *is one-to-one onto* $L^2(\partial\Omega) \to L^2(\partial\Omega)$ *and hence it has a bounded inverse* $(1-\mathbf{E})^{-1} : L^2(\partial\Omega) \to L^2(\partial\Omega)$.

Now, both $\mathbf{S}$ and $\mathbf{K}$ reproduce holomorphic functions so that, as operators on $L^2(\partial\Omega)$, $\mathbf{SK} = \mathbf{K}$ and $\mathbf{KS} = \mathbf{S}$ and taking adjoints

$$\mathbf{K}^*\mathbf{S} = \mathbf{K}^* \quad \text{and} \quad \mathbf{SK}^* = \mathbf{S}$$

since $\mathbf{S}^* = \mathbf{S}$. Subtracting off

$$\mathbf{S}\big(\mathbf{K}^* - \mathbf{K}\big) = \mathbf{S} - \mathbf{K} \;\; \Rightarrow \;\; \mathbf{S} = \mathbf{K}\big(1 - \mathbf{E}\big)^{-1}. \tag{6.10}$$

This is our basic desired formula. Notice that in the process of subtracting $\mathbf{K}^* - \mathbf{K}$, the principal value signs in (6.2) and (6.4) have disappeared.

Let $\widetilde{\dot{\xi}}(w)$ be the vector (*i.e.,* complex number) which results from reflecting $\dot{\xi}(w)$ in the chord determined by $z \in \partial\Omega$ and $w \in \partial\Omega$. Then

$$E(w, z) = \frac{1}{2\pi i}\frac{1}{z - w}\big[\widetilde{\dot{\xi}}(w) - \dot{\xi}(z)\big] \tag{6.11}$$

In order to show (6.11), we may apply a rotation so that $z - w$ is horizontal *i.e.,* $z - w \in \mathbb{R}$. Now the circle is that only plane curve such that the chord determined by any two of its points meets the curve with the same angle at both points. Hence $E(w, z) \equiv 0$ for all z, $w \in \partial\Omega$ implies that Ω is a circle. In fact, we just proved the following.

Theorem 6.4. *The only bounded, smooth, simply connected plane region Ω whose Szegö kernel $S(w, z)$ coincides with the Cauchy kernel $K(w, z)$ for all $w \in \Omega$ and $z \in \partial\Omega$ is the disc.*

We are interested in the special case in which the $L^2(\partial\Omega)$ operator norm of $\mathbf{E}$ is less than 1 because then (6.10) can be rewritten as a geometric series. Assume then that Ω is "*nearly-circular*" in the sense that

$$\frac{1}{4\pi^2}\int_{z\in\partial\Omega}\int_{w\in\partial\Omega}\frac{1}{|z - w|^2}\big|\widetilde{\dot{\xi}}(w) - \dot{\xi}(z)\big]^2\, d\sigma(z)\, d\sigma(w) < 1 \tag{6.12}$$

which implies $\|\mathbf{E}\|_{op} < 1$. In this case we can summarize our results in

Theorem 6.5. *Assume* (6.12) *holds. Then*

(1). $\mathbf{S} = \sum_{j=0}^{\infty}\mathbf{K}\,\mathbf{E}^j$*, where the series converges in the $L^2(\partial\Omega)$ operator norm.*

(2). *For any $u \in L^2(\partial\Omega)$ the remainder $\mathcal{R}_N(u) =: \mathbf{S}(u) - \sum_{j=0}^{N}\mathbf{K}\,\mathbf{E}^j(u) \to 0$ uniformly on $\partial\Omega$. The same holds for any derivative with respect to arc length $D^\alpha\mathcal{R}_N(u)$ which exists if $N \geq 1$. By the maximum principle the convergence are also uniform on $\bar{\Omega}$.*

7. Analysis on the Unbounded Unit Ball in $\mathbb{C}^{n+1}$

7.1. **Action on the Siegel upper half space.** Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ be the unit dick in the complex plane and let $\Omega_1 = \{w \in \mathbb{C} : \mathrm{Im}(w) > 0\}$ be the upper half plane in $\mathbb{C}$. Consider the Cayley transform

$$\mathfrak{C} : \mathbb{D} \to \Omega_1$$

is given explicitly by

$$\mathfrak{C}(z) = \frac{i(1 - z)}{1 + z}.$$

Notice that $\mathfrak{C}$ is a bijective may and its inverse is given by

$$\mathfrak{C}^{-1}(w) = \frac{i - w}{i + w}.$$

Now we would like to generalize this transform to high dimensional cases. In $\mathbb{C}^{n+1}$, the unit ball can be written as $\mathbb{B} = \{(z_1, \dots, z_{n+1}) \in \mathbb{C}^{n+1} : \sum_{k=1}^{n+1} |z_k|^2 < 1\}$. It turns out that the unbounded realization of the ball $\mathbb{B}$ is given by

$$\Omega_{n+1} = \big\{(w_1, \dots, w_{n+1}) = (w', w_{n+1} \in \mathbb{C}^{n+1} : \text{Im}(w_{n+1}) > \sum_{k=1}^{n} |w_k|^2\big\}.$$

Now the mapping that shows $\mathbb{B}$ and Ω_{n+1} to be biholomorphically equivalent is given by

$$\begin{aligned} \Phi \,:\mathbb{B} \,&\to\, \Omega_{n+1} \\ (z_1, \dots, z_n, z_{n+1}) \,&\mapsto\, \Big(\frac{z_1}{1+z_{n+1}}, \dots, \frac{z_n}{1+z_{n+1}}, \frac{i(1-z_{n+1})}{1+z_{n+1}}\Big). \end{aligned}$$

Denote $Aut(\Omega_{n+1})$ the collection of all biholomorphic self-mappings of Ω_{n+1}. This set forms a group when equipped with the binary operation of composition of mappings. In fact it is a topological group with the topology of uniform convergence on compact sets. There is a natural isomorphism between $Aut(\mathbb{B})$ and $Aut(\Omega_{n+1})$ which is given by

$$Aut(\mathbb{B}) \,\ni\, \psi \,\mapsto\, \Phi \circ \psi \circ \Phi^{-1} \,\in\, Aut(\Omega_{n+1}).$$

It turns out that we can understand the automorphism group of $\mathbb{B}$ more completely by passing to the automorphism group of Ω_{n+1}. We shall use the idea of the Iwasawa decomposition $G = KAN$ where K is compact, A is Abelian, and N is nilpotent.

The compact part of $Aut(\mathbb{B})$ is the collection of all automorphisms that fix the origin. Using Schwartz lemma, it is known that any such automorphism is a unitary rotation. This is an $(n+1) \times (n+1)$ complex matrix whose rows (or columns) form a Hermitian orthonormal basis of $\mathbb{C}^{n+1}$. Let us denote this subgroup by K. We see that the group is compact just using a normal families argument: if $\{\phi_j\}$ is a sequence in K then Montel's theorem guarantees that there will be a subsequence converging uniformly on compact sets. It is easy to show that the limit function will be a biholomorphic mapping that fixes the origin.

Now let us look at the Abelian part. For this part, it is more convenient to begin our analysis on Ω_{n+1}. Let us consider the group of dilations, which consists of the nonisotropic mappings

$$\widetilde{\delta}_\varepsilon \,:\, \Omega_{n+1} \,\to\, \Omega_{n+1}$$

given by

$$\widetilde{\delta}_\varepsilon(w_1, \dots, w_n, w_{n+1}) \,=\, (\varepsilon w_1, \dots, \varepsilon w_n, \varepsilon^2 w_{n+1})$$

for any $\varepsilon > 0$. This group is clearly Abelian. It corresponds, under the mapping Φ, to the group of mappings on $\mathbb{B}$ given by

$$\delta_\varepsilon(z_1, \dots, z_n, z_{n+1}) \,=\, \Phi^{-1} \circ \widetilde{\delta}_\varepsilon \circ \Phi(z).$$

Now it is immediate to calculate that

$$\Phi^{-1}(w) \,=\, \left(\frac{2iw_1}{i+w_{n+1}}, \dots, \frac{2iw_n}{i+w_{n+1}}, \frac{i-w_{n+1}}{i+w_{n+1}}\right).$$

After long computation, one has

$$\delta_\varepsilon(z) \,=\, \left(\frac{2\varepsilon z_1}{(1+\varepsilon^2)+z_{n+1}(1-\varepsilon^2)}, \dots, \frac{2\varepsilon z_n}{(1+\varepsilon^2)+z_{n+1}(1-\varepsilon^2)}, \frac{(1-\varepsilon^2)+z_{n+1}(1+\varepsilon^2)}{(1+\varepsilon^2)+z_{n+1}(1-\varepsilon^2)}\right).$$

One may verify directly that $z \in \mathbb{B} \Leftrightarrow \delta_\varepsilon(z) \in \mathbb{B}$. We shall find the "nilpotent" part. Again, it is much easier to work on the unbounded domain Ω_{n+1}.

Denote by $\mathbf{H}^n$ the Heisenberg group:

$$\mathbf{H}^n \approx \mathbb{C}^n \times \mathbb{R} = \{(z,t) = (z_1, \dots, z_n, t) : z \in \mathbb{C}^n, t \in \mathbb{R}\}$$

with group law

$$(z,t)\cdot(w,s)=\left(z+w,t+s+2\mathrm{Im}\sum_{j=1}^{n}z_j\overline{w}_j\right). \tag{7.1}$$

It is clear, because of the Hermitian inner product $z\cdot\bar{w}=\sum_{j=1}^{n}z_j\overline{w}_j$, that the group operation (7.1) is non-abelian. Now an element of $\partial\Omega_{n+1}$ has the form

$$(w_1,\dots,w_n,\mathrm{Re}(w_{n+1})+i|w'|^2),\qquad \text{where } w'=(w_1,\dots,w_n).$$

We identify this boundary point with the Heisenberg group element $(w',\mathrm{Re}(w_{n+1}))$, and call the corresponding mapping Ψ. Now we can specify how the Heisenberg group acts on $\partial\Omega_{n+1}$. If $w=(w',w_{n+1})\in\partial\Omega_{n+1}$ and $g=(z',t)\in\mathbf{H}^n$ then we have the action

$$g(w)=\Psi^{-1}\big(g\cdot\Psi(w)\big)=\Psi^{-1}\big[g\cdot(w',\mathrm{Re}(w_{n+1}))\big]=\Psi^{-1}\big[(z',t)\cdot(w',\mathrm{Re}(w_{n+1}))\big].$$

More generally, if $w\in\Omega_{n+1}$ is an arbitrary element then we write

$$\begin{aligned}
w&=(w_1,\dots,w_n,w_{n+1})=(w',w_{n+1})\\
&=\big(w_1,\dots,w_n,\mathrm{Re}(w_{n+1})+i|w'|^2+i(\mathrm{Im}(w_{n+1})-|w'|^2)\big)\\
&=\big(w_1,\dots,w_n,\mathrm{Re}(w_{n+1})+i|w'|^2\big)+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w'|^2)\big).
\end{aligned}$$

Now we may introduce the "height function" $\rho=\mathrm{Im}(w_{n+1})-|w'|^2$ on Ω_{n+1}. Then we may let g acts on w as follows:

$$\begin{aligned}
g[w]&=g\Big[\big(w_1,\dots,w_n,\mathrm{Re}(w_{n+1})+i|w'|^2\big)+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w'|^2)\big)\Big]\\
&\equiv g\Big[\big(w_1,\dots,w_n,\mathrm{Re}(w_{n+1})+i|w'|^2\big)\Big]+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w'|^2)\big).
\end{aligned} \tag{7.2}$$

In other words, we let g act on level sets of the height function. From now on, let us drop the prime in the variable. For $g=(z,t)$, one has

$$\begin{aligned}
g[w]&=g\Big[\big(w_1,\dots,w_n,\mathrm{Re}(w_{n+1})+i|w|^2\big)\Big]+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w|^2)\big)\\
&=\Psi^{-1}\big[g\cdot(w,\mathrm{Re}(w_{n+1})\big]+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w|^2)\big)\\
&=\Psi^{-1}\Big[(z+w,t+\mathrm{Re}(w_{n+1})+2\mathrm{Im}(z\cdot\bar{w}))\Big]+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w|^2)\big)\\
&\equiv\big(z+w,t+\mathrm{Re}(w_{n+1})+2\mathrm{Im}(z\cdot\bar{w})+i|z+w|^2\big)+\big(0,\dots,0,i(\mathrm{Im}(w_{n+1})-|w|^2)\big)\\
&=\big(z+w,t+i|z|^2+w_{n+1}+2i\bar{z}\cdot w\big).
\end{aligned}$$

This mapping is plainly holomorphic in w (but not is z). Hence we see explicitly that the action of the Heisenberg group on Ω_{n+1} is a biholomorphic mapping.

As we have mentioned before, the group $\mathbf{H}_n$ acts simply transitively on $\partial\Omega_{n+1}$. It follows that the group may be identified with $\partial\Omega_{n+1}$ in a natrual way. Let us now make this identification explicit. First observe that $\mathbf{0}\in\partial\Omega_{n+1}$. If $g=(z,t)\in\mathbf{H}_n$, then

$$g(\mathbf{0})=\Psi^{-1}\big((z,t)\cdot(0,0)\big)=\Psi^{-1}(z,t)=(z,t+i|z|^2)\ \in\ \partial\Omega_{n+1}.$$

Conversely, if $(w,\mathrm{Re}(w_{n+1})+i|w|^2)\in\partial\Omega_{n+1}$, then let $g=(w,\mathrm{Re}(w_{n+1}))$. Hence

$$g(\mathbf{0})=\Psi^{-1}(w,\mathrm{Re}(w_{n+1})=\big(w,\mathrm{Re}(w_{n+1})+i|w|^2\big)\ \in\ \partial\Omega_{n+1}.$$

Comparing this result with the similar but much simpler situation for the upper half plane $\mathbb{R}^2_+$, we may conclude that $\mathbf{H}_n\cong\partial\Omega_{n+1}$.

7.2. **The Lie group structure of the Heisenberg group.** The Heisenberg group $\mathbf{H}_n$ has $2n+1$ real dimensions and we can define the differentiation of a function in each direction consistent with the group structure by considering 1-parameter subgroup in each direction.

Let $g=(z,t)$ be an element in $\mathbf{H}_n$, where

$$z = (z_1,\dots,z_n) = (x_1+iy_1,\dots,x_n+iy_n)$$

and $t\in\mathbb{R}$. If we let

$$\gamma_{2k-1}(s)=(0,0,\dots,s+i0,\dots,0), \qquad \gamma_{2k}(s)=(0,0,\dots,0+is,\dots,0)$$

for $1\le k\le n$ and the s term in the kth slot, and if we let

$$\gamma_{2n+1}(s)=\gamma_t(s)=(0,s)$$

(with $2n$ zeros and one s), then each forms a one-parameter subgroup of $\mathbf{H}_n$.

WE define the differentiation of f at $g=(z,t)$ in each one-parameter subgroup as follows:

$$\begin{aligned}
X_kf(g)&=\frac{d}{ds}f\big(g\cdot\gamma_{2k-1}(s)\big)\Big|_{s=0}\\
&=\frac{d}{ds}f\big(x_1+iy_1,\dots,x_k+s+iy_k,\dots,x_n+iy_n,t+2y_ks\big)\Big|_{s=0}\\
&=\left(\frac{\partial f}{\partial x_k}+2y_k\frac{\partial f}{\partial t}\right)(z,t),\quad 1\le k\le n,\\
Y_kf(g)&=\frac{d}{ds}f\big(g\cdot\gamma_{2k}(s)\big)\Big|_{s=0}\\
&=\frac{d}{ds}f\big(x_1+iy_1,\dots,x_k+i(y_k+s),\dots,x_n+iy_n,t-2x_ks\big)\Big|_{s=0}\\
&=\left(\frac{\partial f}{\partial y_k}-2x_k\frac{\partial f}{\partial t}\right)(z,t),\quad 1\le k\le n,\\
Tf(g)&=\frac{d}{ds}f\big(g\cdot\gamma_t(s)\big)\Big|_{s=0}\\
&=\frac{d}{ds}f\big(x_1+iy_1,\dots,x_n+iy_n,t+s\big)\Big|_{s=0}\\
&=\frac{\partial f}{\partial t}(z,t).
\end{aligned}$$

Note that

$$[X_j,X_k]=[Y_j,Y_k]=[X_k,T]=[Y_k,T]=0,\qquad \forall\ j,k=1,\dots,n$$

and

$$[X_j,Y_k]=0\qquad\text{if}\quad j\ne k.$$

The only nonzero commutator in the Heisenberg group is $[X_k,Y_k]$:

$$\begin{aligned}
[X_k,Y_k]&=\left(\frac{\partial f}{\partial x_k}+2y_k\frac{\partial f}{\partial t}\right)\left(\frac{\partial}{\partial y_k}-2x_k\frac{\partial}{\partial t}\right)-\left(\frac{\partial}{\partial y_k}-2x_k\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x_k}+2y_k\frac{\partial}{\partial t}\right)\\
&=-2\left(\frac{\partial}{\partial x_k}x_k\right)\frac{\partial}{\partial t}-2\left(\frac{\partial}{\partial y_k}y_k\right)\frac{\partial}{\partial t}=-4\frac{\partial}{\partial t}=-4T.
\end{aligned}$$

In summary, all commutators $[X_j,X_k]$, $[Y_j,Y_k]$ for $j\ne k$ and $[X_k,T]$, $[Y_j,T]$ equal zero. The only nonzero commutator is $[X_j,Y_j]=-4T$. All second-order commutators $[[A,B],C]$ will be zero, just because $[A,b]$ will be either zero or $-4T$. Thus the vector fields on the Heisenberg form a nilpotent Lie algebra of step 2.

7.3. **Analysis on the Heisenberg group.** We define the "homogeneous norm" $|\cdot|_h$ on $\mathbf{H}_n$ to be

$$|g|_h = \left(|z|^4 + t^2\right)^{\frac{1}{4}}. \tag{7.3}$$

Then $|\cdot|_h$ satisfies the following properties

- $|g|_h \geq 0$ and $|g|_h = 0 \Leftrightarrow g = \mathbf{0}$;
- $g \mapsto |g|_h$ is a continuous function from $\mathbf{H}_n$ to $\mathbb{R}^+$ and is smooth on $\mathbf{H}_n \setminus \{\mathbf{0}\}$;
- $|\delta_\varepsilon(g)|_h = \varepsilon |g|_h$ for all $\varepsilon > 0$.

These three properties do not uniquely determine the norm. If ϕ is positive, smooth away from the origin and homogeneous of degree 0 in the group dilation structure, then $\phi(g)|g|_h$ is another "norm". The group $\mathbf{H}_n$ is also equipped with the Euclidean norm in $\mathbb{R}^{2n+1}$. Let us denote it as $|\cdot|_e$:

$$|g|_e = \left(|z|^2 + t^2\right)^{\frac{1}{2}}.$$

Lemma 7.1. *For $|g|_e^2 \leq \frac{1}{2}$, we have*

$$|g|_e \leq |g|_h \leq \sqrt{|g|_e}.$$

Proof. We see that

$$|g|_e = \sqrt{|z|^2 + t^2} \leq \left(|z|^4 + t^2\right)^{\frac{1}{4}} = |g|_h$$

reduces to

$$\left(|z|^2 + t^2\right)^2 \leq |z|^4 + t^2 \quad \text{or} \quad 2t^2|z|^2 + t^4 \leq t^2, \quad \text{or} \quad 2|z|^2 + t^2 \leq 1.$$

Since we assumed that $|g|_e^2 = |z|^2 + t^2 \leq \frac{1}{2}$, we have $2|z|^2 + t^2 \leq 1$. Furthermore

$$|g|_h = \left(|z|^4 + t^2\right)^{\frac{1}{4}} \leq \left(|z|^2 + t^2\right)^{\frac{1}{4}} = \sqrt{|g|_e}.$$

That completes the proof. □

Lemma 7.2. *We have*

$$dV(g) = dx_1 dy_1 \cdots dx_n dy_n dt = r^{2n+1} dr d\sigma(\xi),$$

where $d\sigma$ is a smooth, positive measure on the Heisenberg unit sphere $\{\xi \in \mathbf{H}_n : |\xi|_h = 1\}$.

Proof. Let $x = (x_1, \ldots, x_{2n+1}) \in \mathbf{H}_n$. If we let $r = |x|_h$, then

$$x = (x_1, \ldots, x_{2n+1}) = r(\xi_1, \ldots, \xi_{2n+1}) = (r\xi_1, \ldots, r\xi_{2n}, r^2\xi_{2n+1})$$

where $|\xi|_h = 1$. Then we have

$$\xi_{2n+1}^2 = 1 - \left(\sum_{k=1}^{2n} \xi_j^2\right)^2.$$

Therefore we may consider the coordinate transform

$$(x_1, \ldots, x_{2n+1}) \mapsto (\xi_1, \ldots, \xi_{2n}, r).$$

Calculating the Jacobian matrix, one obtains

$$J = \begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_2} & \cdots & \frac{\partial x_1}{\partial r} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_{2n+1}}{\partial \xi_1} & \frac{\partial x_{2n+1}}{\partial \xi_2} & \cdots & \frac{\partial x_{2n+1}}{\partial r} \end{bmatrix} = \begin{bmatrix} r & 0 & \cdots & 0 & \xi_1 \\ 0 & r & \cdots & 0 & \xi_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & r & \xi_{2n} \\ r^2 \frac{\partial \xi_{2n+1}}{\partial \xi_1} & r^2 \frac{\partial \xi_{2n+1}}{\partial \xi_2} & \cdots & r^2 \frac{\partial \xi_{2n+1}}{\partial \xi_{2n}} & 2r\xi_{2n+1} \end{bmatrix}$$

Therefore,

$$|\det J| = r^{2n+1}\Big(2\xi_{2n+1} - \sum_{k=1}^{2n} \xi_k \frac{\partial \xi_{2n+1}}{\partial \xi_k}\Big).$$

Hence

$$dV(x) = dx_1 \cdots dx_{2n} dx_{2n+1} = r^{2n+1} dr\, d\sigma(x),$$

where

$$d\sigma = \Big(2\xi_{2n+1} - \sum_{k=1}^{2n} \xi_k \frac{\partial \xi_{2n+1}}{\partial \xi_k}\Big) d\xi_1 \cdots d\xi_{2n}.$$

□

Now we can calculate the volume of the ball in $\mathbf{H}_n$ using polar coordinates. Let Σ be the surface area of the unit sphere in $\mathbf{H}_n$:

$$\mathcal{C} \;=\; \int_{|\xi|_h=1} d\sigma(\xi) \;=\; \int_\Sigma d\sigma(\xi).$$

Then the volume of the unit ball in $\mathbf{H}_n$ is

$$|B| \;=\; \int_{|x|_h \le 1} dV(x) = \int_\Sigma \int_0^1 r^{2n+1} dr\, d\sigma \;=\; \frac{\mathcal{C}}{2n+2}.$$

Hence the volume of a ball of radius R will be

$$|B(\mathbf{0};r)| \;=\; \int_{|x|_h \le R} dV(x) \;=\; \int_{|x|_h \le 1} dV(Rx) \;=\; R^{2n+2} \int_{|x|_h \le 1} dV(x) \;=\; R^{2n+2}|B|.$$

Now the integration of characteristic function of balls is well defined. We call $2n+2$ the homogeneous dimension of $\mathbf{H}_n$ even though the topological dimension of $\mathbf{H}_n$ is $2n+1$. The critical index m for a singular integral is such that

$$\int_{B(\mathbf{0};1)} \frac{1}{|x|^\alpha} dV(x) \;=\; \begin{cases} +\infty \ \text{ if } \alpha \ge m \\ < \infty \ \text{ if } 0 < \alpha < m \end{cases}$$

and the critical index coincides with the homogeneous dimension. Thus the critical index for a singular integral in $\mathbf{H}_n$ is $2n+2$, which is different from the topological dimension.

7.4. **$\mathbf{H}_n$ is a space of homogeneous type.** For $x, y \in \mathbf{H}_n$, we define the distance $d(x,y)$ as follows:

$$d(x,y) \;\equiv\; |x^{-1} \cdot y|_h.$$

Then $d(x,y)$ satisfies the following properties:

- $d(x,y) = 0 \Leftrightarrow x = y$;
- $d(x,y) \;=\; d(y,x)$;
- There exists $\gamma_0 > 0$ such that $d(x,y) \le \gamma_0\big(d(x,u) + d(u,y)\big)$.

Proof. The first assertion is obvious. Now we turn to the second assertion. It is easy to check that $x^{-1} = -x$. Thus

$$d(x,y) \;=\; |x^{-1} \cdot y|_h \;=\; |(-x) \cdot y|_h = |x \cdot (-y)|_h \;=\; d(y,x).$$

It remains to show the three assertion. Let

$$\sup_{|x|_h, |y\_h \le 1} d(x,y) \;=\; C, \qquad \inf_{|x|_h, |y|_h, |u|_h \le 1} \big[d(x,u) + d(u,y)\big] \;=\; D.$$

Then $C \geq 1$ and $D > 0$. Therefore we get

$$d(x,y) \;\leq\; C \;\leq\; \frac{C}{D}\big[d(x,u)+d(u,y)\big], \quad \text{if } \;|x|_h, |y|_h, |u|_h \leq 1.$$

Now, for general x, y and u, let $r = \max\{|x|_h, |y|_h, |u|_h\}$. Then $x = rx'$, $y = ry'$ and $u = ru'$ where $|x'|_h, |y'|_h, |u'|_h \leq 1$. Then we have

$$d(x,y) \;=\; d(rx', ry') \;=\; r\, d(x', y')$$

and

$$d(x,u)+d(u,y) \;=\; r\big[d(x',u')+d(u',y')\big].$$

Hence

$$d(x,y) \;\leq\; \frac{C}{D}\big[d(x,u)+d(u,y)\big], \quad \text{for all } \; x, y\, u.$$

□

Define balls in $\mathbf{H}_n$ by $B(x;r) = \{y \in \mathbf{H}_n : d(x,y) < r\}$. Then, equipped with the Lebesgue measure μ on $\mathbb{R}^{2n+1}$, $\mathbf{H}_n$ is a space of homogeneous type. We need to check the following three conditions:

- *The Positivity Property*: $0 < \mu(B(x;r)) < \infty$ for all $x \in \mathbf{H}_n$ and $r > 0$;
- *The Doubling Property*: there exists a constant $C_1 > 0$ such that

$$\mu(B(x;2r)) \leq C_1 \mu(B(x;r));$$

- *The Enveloping Property*: there exists a constant $C_2 > 0$ such that if $B(x;r) \cap B(y;\rho) \neq \varnothing$ and $\rho \geq r$, then $B(y; C_2\rho) \;\supseteq\; B(x;r)$.

Proof. The first assertion is obvious since $\mu(B(x;r)) = r^{2n+2}|B|$ where $|B|$ is the volume of the unit ball. Moreover

$$\mu(B(x;2r)) = 2^{2n+2}\mu(B(x;r)) \quad \Rightarrow \quad C_1 = 2^{2n+2}.$$

Now we turn to the third assertion. The result follows because, equipped with the distance d, the Heisenberg group is a quasi-metric space. In detail, let $v \in B(x;r) \cap B(y,\rho)$. Then $d(x,v) < r$ and $d(v,y) < \rho$. If $u \in B(x;r)$, then we obtain

$$\begin{aligned} d(y,u) &\leq \gamma_0\big[d(y,v)+d(v,u)\big] \leq\; \gamma_0\big[\rho + \gamma_0 d(v,x) + \gamma_0 d(x,u)\big] \\ &\leq \gamma_0\big[\rho + 2\gamma_0 r\big] \leq\; \gamma_0(1+2\gamma_0)\rho. \end{aligned}$$

Thus we may let $C_2 = \gamma_0 + 2\gamma_0^2$. □

Hence, $(\mathbf{H}_n, \mu)$ is a space of homogeneous type in the sense of Coifman and Weiss [6].

We say that a function $f : \mathbf{H}_n \to \mathbb{C}$ is *homogeneous of degree* $m \in \mathbb{R} \Leftrightarrow f(\varepsilon x) = \varepsilon^m f(x)$.

The Schwartz space $\mathcal{S}$ of $\mathbf{H}_n$ is the Schwartz space of $\mathbb{R}^{2n+1}$:

$$\mathcal{S}(\mathbf{H}_n) = \left\{ f : \mathbf{H}_n \to \mathbb{C} : \; \|f\|_{\alpha,\beta} \;\equiv\; \sup_{x\in\mathbf{H}_n} \left| x^\alpha \Big(\frac{\partial}{\partial x}\Big)^\beta f(x) \right| < \infty \right\}.$$

The norm $\|\cdot\|_{\alpha,\beta}$ is a semi-norm and $\mathcal{S}$ is a Frechet space. The dual space of $\mathcal{S}$ is the space of tempered distributions. For $\varphi \in \mathcal{S}$ and $\delta > 0$, set

$$\varphi_\delta(x) = \varphi(\delta x), \qquad \varphi^\delta(x) = \delta^{-2n-2}\varphi\Big(\frac{x}{\delta}\Big).$$

Note the homogeneous dimension playing a role in the definition of φ^δ.

A tempered distribution τ is said to be *homogeneous of degree* m provided that

$$\tau\big(\varphi^\delta\big) \;=\; \delta^m \tau(\varphi).$$

If it happens that the distribution τ is given by integration against a function K which is homogeneous of degree m, then the resulting distribution is homogeneous of degree m:

$$\begin{aligned}\tau\big(\varphi^\delta\big) &= \int K(x)\varphi^\delta(x)dx = \int K(\delta x)\varphi(x)dx \\ &= \int \delta^m K(x)\varphi(x)dx = \delta^m \tau(\varphi).\end{aligned}$$

Proposition 7.3. *Let f be a homogeneous function of degree $\lambda \in \mathbb{R}$. Assume that f is C^1 away from the origin. Then there exists a constant $C > 0$ such that*

$$|f(x) - f(y)| \le C|x-y|_h \cdot |x|_h^{\lambda-1}, \qquad \textit{whenever } |x-y|_h \le \frac{1}{\gamma_0}|x|_h;$$

$$|f(x\cdot y) - f(x)| \le C|y|_h \cdot |x|_h^{\lambda-1}, \qquad \textit{whenever } |y|_h \le \frac{1}{\gamma_0}|x|_h;$$

Proof. Let us look at the first inequality. If we dilate x, y by $\rho > 0$, then

$$\begin{aligned}LHS &= |f(\rho x) - f(\rho y)| = \rho^\lambda |f(x) - f(y)| \\ RHS &= C|\rho x - \rho y|_h \cdot |\rho x|_h^{\lambda-1} = C\rho^\lambda |x-y|_h \cdot |x|_h^{\lambda-1}.\end{aligned}$$

Thus the inequality is invariant under dilation. So it is enough to prove the inequality when $|x|_h = 1$ and $|x-y|_h \le \frac{1}{4\gamma_0}$. Then y is bounded from 0:

$$\begin{aligned}d(x,0) &\le \gamma_0\big(d(x,y) + d(y,0)\big) \\ d(y,0) &\ge \frac{1}{\gamma_0} - d(x,y) \ge \frac{1}{\gamma_0} - \frac{1}{4\gamma_0} = \frac{3}{4\gamma_0} > 0.\end{aligned}$$

Applying the classical Euclidean mean value theorem to $f(x)$:

$$|f(x) - f(y)| \le \sup|\nabla f| \cdot |x-y|_e.$$

Note that the supremum is taken on the segment connecting x and y. Since $|x|_h = 1$ and y is bounded from 0, we have

$$|f(x) - f(y)| \le C|x-y|_e \le C|x-y|_h.$$

The last inequality is by Lemma 7.1. We may show the second inequality by similar method and conclude the proof of the proposition. □

Let (X, μ) be a measure space and $f : X \to \mathbb{C}$ a measurable function. We say f is *weak type* p, $0 < p < infty$ if there exists a constant $C > 0$ such that

$$\mu\big\{x \in X : |f(x)| > \lambda\big\} \le \frac{C}{\lambda^p}, \qquad \forall, \lambda > 0.$$

Remark. If $f \in L^p$, then f is weak type p. For suppose that $f \in L^p(X)$, then

$$C \ge \int_X |f(x)|^p \, d\mu(x) \ge \int_{|f|>\lambda} |f(x)|^p \, d\mu(x) \ge \lambda^p \cdot \mu\big\{x \in X : |f(x)| > \lambda\big\},$$

hence f is of weak type p. But the converse is not true. Let $X = \mathbb{R}^+$ and let $f(x) = \frac{1}{x^{1/p}}$ is weak type but not pth power integrable. The following two lemmas are very useful in our discussion which can be found in Folland and Stein [11]

Theorem 7.4. *Let (X,μ), (Y,ν) be measurable spaces. Let*

$$K:\, X\times Y \to \mathbb{C}$$

satisfy

$$\mu\big\{x:\, |K(x,y)|>\lambda\big\} \le \frac{C_1}{\lambda^r}, \qquad (\text{for fixed } y)$$
$$\nu\big\{y:\, |K(x,y)|>\lambda\big\} \le \frac{C_2}{\lambda^r}, \qquad (\text{for fixed } x)$$

where C_1 and C_2 are independent of y and x respectively and $r>1$. Then

$$f \mapsto \int_Y f(y)K(x,y)d\nu(y)$$

maps L^p to L^q where $\frac{1}{q}=\frac{1}{p}+\frac{1}{r}-1$, for $1<p<\frac{r}{r-1}$.

The main idea in the proof of Theorem 7.4 is the following result of Isaiah Schur:

Lemma 7.5. (Schur's Lemma) *Let $1\le r\le\infty$. Let (X,μ), (Y,ν) be measurable spaces and let $K: X\times Y\to\mathbb{C}$ satisy*

$$\left(\int_X |K(x,y)|^r d\mu(x)\right)^{\frac{1}{r}} \le C_1$$
$$\left(\int_Y |K(x,y)|^r d\nu(y)\right)^{\frac{1}{r}} \le C_2,$$

where C_1 and C_2 are independent of y and x respectively. Then

$$f \mapsto \int_Y f(y)K(x,y)\, d\nu(y)$$

maps $L^p(X)$ to $L^q(X)$ where $\frac{1}{q}=\frac{1}{p}+\frac{1}{r}-1$, for $1<p<\frac{r}{r-1}$.

Proof of Theorem 7.4: By the Marcinkiewicz interpolation theorem, it is enough to show that $f\mapsto T(f)$ is weak type (p,q). Fix $\lambda>0$. Let $\rho>0$ be a constant to be specified later. Let us define

$$K_1(x,y) = \begin{cases} K(x,y) & \text{if} \quad |K(x,y)|\ge\rho \\ 0 & \text{if} \quad |K(x,y)|<\rho \end{cases}$$
$$K_2(x,y) = K(x,y)-K_1(x,y).$$

Obvious, $K_2(x,y)$ is bounded. Define

$$T_1(f)(x) = \int_Y K_1(x,y)f(y)\, d\nu(y)$$
$$T_2(f)(x) = \int_Y K_2(x,y)f(y)\, d\nu(y).$$

Then $T(f)=T_1(f)+T_2(f)$. Hence,

$$\begin{aligned}\mu\big\{x\in X:\, |T(f)(x)|>2\lambda\big\} &= \mu\big\{x\in X:\, |T_1(f)(x)+T_2(f)(x)|>2\lambda\big\}\\ &\le \mu\big\{x\in X:\, |T_1(f)(x)|+|T_2(f)(x)|>2\lambda\big\}\\ &\le \mu\big\{x\in X:\, |T_1(f)(x)|>\lambda\big\}+\mu\big\{x\in X:\, |T_2(f)(x)|>\lambda\big\}.\end{aligned} \tag{7.4}$$

Let $f \in L^p(Y)$ and assume $\|f\|_{L^p} = 1$. Choose p' such that $\frac{1}{p} + \frac{1}{p'} = 1$. Then we get

$$\left|T_2(f)(x)\right| = \left|\int_Y K_2(x,y)f(y)d\nu(y)\right| \leq \left(\int_Y |K_2(x,y)|^{p'}d\nu(y)\right)^{\frac{1}{p'}} \left(\int_Y |f(y)|^p d\nu(y)\right)^{\frac{1}{p}}$$

and from the definition of distribution function,

$$\begin{aligned} \int_Y |K_2(x,y)|^{p'}d\nu(y) \;&=\; \int_0^\rho p's^{p'-1}\alpha_{K_2(x,\cdot)}(s)ds \\ \leq \int_0^\rho p's^{p'-1}\frac{C}{s^r}ds \;&=\; Cp'\int_0^\rho s^{p'-1-r}ds \;=\; C'\rho^{p'-r}. \end{aligned}$$

The last inequality holds since

$$p' - 1 - r \;=\; \frac{1}{1-\frac{1}{p}} - 1 - r = \frac{p}{p-1} - 1 - r > -1.$$

Thus we get

$$\left|T_2(f)(x)\right| \;\leq\; (C'\rho^{p'-1})^{\frac{1}{p'}}\|f\|_{L^p} = C^\star\rho^{1-\frac{r}{p'}}.$$

Let $\rho = \left(\frac{s}{C^\star}\right)^{\frac{q}{r}}$. Then

$$\left|T_2(f)(x)\right| \;\leq\; C^\star\left(\frac{\lambda}{C^\star}\right)^{\frac{q}{r}(1-\frac{r}{p'})} = \lambda.$$

Therefore one obtains the distribution function $\alpha_{T_2(f)}(s) = \mu\{x \in X : |T_2(f)(x)| > s\}$ at the "height" $s = \lambda$ is 0. From (7.4), it follows that

$$\alpha_{T(f)}(2\lambda) \;\leq\; \alpha_{T_1(f)}(\lambda).$$

Since $|K_1(x,y)| \geq \rho$, we have $\alpha_{K_1(x,\cdot)}(\lambda) = \alpha_{K_1(x,\cdot)}(\rho)$ if $\lambda \leq \rho$. Thus

$$\begin{aligned} \int_Y |K_1(x,y)|d\nu(y) \;&=\; \int_0^\infty \alpha_{K_1(x,\cdot)}(\lambda)d\lambda \\ &= \int_0^\rho \alpha_{K_1(x,\cdot)}(\lambda)d\lambda + \int_\rho^\infty \alpha_{K_1(x,\cdot)}(\lambda)d\lambda \\ &\leq \rho\cdot\alpha_{K_1(x,\cdot)}(\rho) + \int_\rho^\infty \frac{C}{s^r}ds \\ &\leq \rho\frac{C}{\rho^r} + \frac{C}{1-r}\rho^{1-r} \;=\; C^\star\rho^{1-r}. \end{aligned}$$

Similarly, we get

$$\int_X |K_1(x,y)|d\mu(x) \;\leq\; C\rho^{1-r}.$$

Recall that is $L(x,y)$ is a kernel and

$$\int_X |L(x,y)|d\mu(x) \;\leq\; C \qquad \text{and} \qquad \int_Y |L(x,y)|d\nu(y) \;\leq\; C$$

then by Schur's lemma, $f \mapsto \int_Y f(y)L(x,y)d\nu(y)$ is bounded on $L^p(Y)$, $1 \leq p \leq \infty$. Thus T_1 is bounded on $L^p(Y)$ and

$$\|T_1(f)\|_{L^p(X)} \;\leq\; C\rho^{1-r}\|f\|_{L^p(Y)} \;=\; C\rho^{1-r}.$$

Denote $\Lambda = \{x \in X : T_1(f)(x)\lambda\}$. By Tchebycheff's inequality, we have

$$\begin{aligned}\alpha_{T_1(f)}(\lambda) &= \int_\Lambda d\mu(x) \le \int_\Lambda \frac{|T_1(f)(x)|^p}{\lambda^p} d\mu(x)\\ &\le \frac{\|T_1(f)\|^p_{L^p(X)}}{\lambda^p} \le \frac{(C\rho^{1-r})^p}{\lambda^p}\\ &= C' \frac{\left(\frac{\lambda}{C^\star}\right)^{\frac{q}{r}(1-r)p}}{\lambda^p} = \widetilde{C} \cdot \lambda^{\frac{pq(1-r)}{r}-p} = \frac{C}{\lambda^q}.\end{aligned}$$

Therefore,

$$\alpha_{T(f)}(2\lambda) \le \frac{C}{\lambda^q}.$$

□

Now we may apply Theorem 7.4 to L^p estimate for the fractional integral on the Heisenberg group. Let

$$K_\alpha(x) = |x|_h^{-2n-2+\alpha}, \qquad 0 < \alpha < 2n+2,$$

be a kernel. Consider the operator

$$\digamma_\alpha : f \mapsto f * K_\alpha$$

on $\mathbf{H}_n$. A natural way to proceed now is to calculate the weak type of K_α. Then Theorem 7.4 can be applied to obtain the mapping properties of the operator $\digamma_\alpha$. Now

$$m\Big\{x \in \mathbf{H}_n : |K_\alpha(x)| > \lambda\Big\} = m\Big\{x : |x|_h \le \big(\frac{1}{\lambda}\big)^{\frac{1}{2n+2-\alpha}}\Big\} \le C \cdot \big(\frac{1}{\lambda}\big)^{\frac{2n+2}{2n+2-\alpha}}.$$

We see immediately that K_α is of weak type $\frac{2n+2}{2n+2-\alpha}$. Thus the hypotheses of Theorem 7.4 are satisfied with $r = \frac{2n+2}{2n+2-\alpha}$. We conclude that $\digamma_\alpha$ maps $L^p(\mathbf{H}_n)$ to $L^q(\mathbf{H}_n)$ with

$$\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{2n+2}, \qquad 1 < p < \frac{2n+2}{\alpha}.$$

8. L^2 Estimates for Singular Integrals on the Heisenberg Group

We now turn our attention to singular integral operators. First, let us mention the following celebrating result which was obtained by Folland and Stein [11] in 1974.

Theorem 8.1. *Let K be a function non $\mathbf{H}_n$ that is smooth away from the origin and homogeneous of degree $-2n-2$. Assume that*

$$\int_{|x|_h=1} K(x)d\sigma(x) = 0,$$

where $d\sigma$ is the area measure on the unit sphere in the Heisenberg group. Define

$$Tf(x) = P.V.(K * f) = \lim_{\varepsilon\to 0}\int_{|y|_h>\varepsilon} K(y)f(y^{-1}\cdot x)dV(y).$$

Then the limit exists pointwisely and in norm and

$$\|T(f)\|_{L^2} \le C\|f\|_{L^2}.$$

This theorem is an analogue of L^2 estimate of the Hilbert transform and Riesz transforms, where the underlying manifold is $\mathbb{R}^n$ and use isotropic dilations, *i.e.,*

$$\delta(x_1, \dots, x_n) = (\delta x_1, \dots, \delta x_n), \qquad \forall\ \delta > 0.$$

The kernel K is a Calderón-Zygmund kernel, *i.e.,* $K(x)$ satisfying properties (4.6), (4.8) and (4.9) in Section 3. In other words, $K \in C^\infty(\mathbb{R}^n \setminus \Sigma)$ satisfies the following properties:

- K is homogeneous of degree $-n$;
- $\int_{|x|_e=1} K(x)dx = 0$ where $|x|_e$ is the Euclidean norm of the vector x;
- $\int_{|x|_e \geq |y|_e} |K(x-y) - K(x)|dx \leq B, \qquad \forall\ |y|_e > 0.$

Then

$$T(f)(x) = \text{p.v.} \int_{\mathbb{R}^n} f(x-y)K(y)dy$$

is bounded on $L^2(\mathbb{R}^n)$. The main idea in proving L^2 estimate of the Hilbert transform and Riesz transforms is to show that $\widehat{\text{p.v.}K}$ is bounded. Then the L^2 boundedness of T follows immediately from the Plancherel theorem. This Fourier analysis approach works on $\mathbf{H}^n$ in principle but not in practice, since the Plancherel formula on $\mathbf{H}^n$ requires consideration of irreducible unitary representations (of $\mathbf{H}^n$) and the Hilbert-Schmidt norms of operator-valued functions. We will use the method, of "so-called" has a pejorative connotation, *almost orthogonality* which is a very powerful tool in dealing with noncommutative singular integrals.

The problem with our kernel K is that it is integrable neither at 0 not at ∞. Consider $T(f)$ as

$$T(f)(x) = \sum_{k=-\infty}^{\infty} \int_{2^k \leq |y|_h \leq 2^{k+1}} f(x \cdot y^{-1})K(y)dy.$$

Each of the operators in the sum is easy to handle and has operator norm of order of magnitude 1. The triangle inequality then yields that

$$\left\| \sum_{k=-N}^{N} T_k \right\|_{\text{op}} \leq C \cdot N.$$

Such a crude estimate is of no use. The insight of the following lemma is that when the operators being summed act on different (nearly orthogonal) parts of the Hilbert space then the norm of the sum is actually independent of the number of terms. This lemma was discovered by Cotlar [7] in connection with ergodic theory, and in the present form was introduced by Knapp and Stein [17] in order to establish the boundedness of singular integral operators on nilpotent Lie groups. Now we make this precise with the following Cotlar-Knapp-Stein Lemma:

Theorem 8.2. *Let $\mathfrak{H}$ be a Hilbert space and a set of bounded (on $\mathfrak{H}$) operators T_j. Suppose that there exists a bi-infinite sequence of positive numbers a_j with $A = \sum_{j=-\infty}^{\infty} a_j < \infty$ such that*

$$\|T_j T_k^*\|_{op} < a_{j-k}^2, \quad \text{and} \quad \|T_j^* T_k\|_{op} < a_{j-k}^2. \tag{8.1}$$

Then for any finite collection of indices Λ,

$$\left\| \sum_{j \in \Lambda} T_j \right\|_{op} \leq A.$$

Proof. Recall the elementary Hilbert space fact that

$$\|TT^*\| = \|T\|^2 = \|T^*\|^2.$$

But TT^* is self-adjoint, and for a self-adjoint operator B we have

$$\|B^k\|_{\text{op}} = \|B\|_{\text{op}}^k, \quad \forall\ k \in \mathbb{N}.$$

Thus, in our case

$$\|(TT^*)^m\|_{\text{op}} = \|TT^*\|_{\text{op}}^m = \|T\|_{\text{op}}^{2m}.$$

In particular, the hypothesis (8.1) implies that

$$\|T\|_{\text{op}} \leq a_0 \qquad \forall \ j.$$

Let $T = \sum_{j\in\Lambda} T_j$. To make use of our condition on products, we consider

$$(TT^*)^m = \sum_{J\in\Lambda^{2m}, J=(j_1,\ldots,j_{2m})} T_{j_1}T^*_{j_2}T_{j_3}T^*_{j_4}\ldots T_{j_{2m-1}}T^*_{j_{2m}}.$$

Our first estimate is

$$\begin{aligned}\|T_{j_1}T^*_{j_2}\ldots T_{j_{2m-1}}T^*_{j_{2m}}\|_{\text{op}} &= \|(T_{j_1}T^*_{j_2})\ldots(T_{j_{2m-1}}T^*_{j_{2m}})\|_{\text{op}} \\ &\leq a^2_{j_1-j_2}\cdot a^2_{j_3-j_4}\cdots a^2_{j_{2m-1}-j_{2m}}\end{aligned} \tag{8.2}$$

Since some of these pairs may be close, we group them differently:

$$\|T_{j_1}T^*_{j_2}\ldots T_{j_{2m-1}}T^*_{j_{2m}}\|_{\text{op}} = \|T_{j_1}(T^*_{j_2}T_{j_3})\ldots(T^*_{j_{2m-2}}T_{j_{2m-1}})T^*_{j_{2m}}\|_{\text{op}}.$$

Thus

$$\|T_{j_1}T^*_{j_2}\ldots T_{j_{2m-1}}T^*_{j_{2m}}\|_{\text{op}} \leq A^2\cdot a^2_{j_2-j_3}\cdots a^2_{j_{2m-2}-j_{2m-1}}. \tag{8.3}$$

We take the geometric mean of the equations (8.2) and (8.3) and see that

$$\|(TT^*)^m\|_{\text{op}} \leq \sum_{J\in\Lambda^{2m}} A\cdot a_{j_1-j_2}\cdot a_{j_2-j_3}\cdots a_{j_{2m-2}-j_{2m-1}}\cdot a_{j_{2m-1}-j_{2m}}.$$

Now we sum over the index j_{2m}, then over j_{2m-1}, etc., each time yielding a contribution of at most A (by our hypothesis). So

$$\|(TT^*)^m\|_{\text{op}} \leq N\cdot A^{2m}, \quad \text{where } N = |\Lambda|.$$

Now when take the m^{th} root, we see that

$$\|(TT^*)\|_{\text{op}} \leq \sqrt[m]{N}\cdot A^2.$$

Letting $m\to\infty$, we have

$$\|TT^*\|_{\text{op}} \leq A^2$$

so that $\|T\|_{\text{op}} \leq A$. This completes the proof of the lemma. □

This lemma is so fundamental that it bears some discussion. In general, if one is summing N operators T_j, each having norm 1, then one cannot expect the sum to have norm less than N. For example, when all the operators T_j are the same operator, then the operator norm of the sum should be N. However if the operators T_j each operate on a "different part" of the Hilbert space, then one might hope for some improvement.

As an example, let the Hilbert space be $L^2(\mathbb{T})$ and let the j^{th} operator T_j be convolution with the j^{th} character, $e^{2\pi ij\theta}$, $j\in\mathbb{Z}$. Then each T_j has operator norm 1. But, by the Riesz-Fischer theory, also $\sum_{j=-N}^{N} T_j$ has norm 1, for any value of N. Of course this works because the operator T_j operates precisely on the one-dimensional space spanned by $e^{2\pi ij\theta}$. From a Hilbert space point of view, operators T_j and T_k, $j\neq k$, live in different worlds.

To impose a hypothesis analogous to what is true for the operators T_j in the last paragraph would be too restrictive. The Cotlar-Knapp-Stein lemma tells us that if the operators T_j are "almost orthogonal", then the result still holds.

Proof of Theorem 8.1. We start with an auxiliary function $\phi(x) = \phi_0(|x|_h)$, where $\phi_0\in C_0^\infty(\mathbb{R}^1_+)$ and

$$\phi_0(\zeta) = \begin{cases} 1 & \text{if} \quad 0\leq\zeta\leq 1 \\ 0 & \text{if} \quad 2\leq\zeta<\infty \end{cases}$$

Let

$$\psi_j(x) = \phi(2^{-j}x) - \phi(2^{-j+1}x).$$

We stress here that the dilations are taking place in the Heisenberg group structure (the action of the Iwasawa subgroup A).

Note that

$$\phi(2^{-j}x) = \begin{cases} 1 & \text{if} \quad |x|_h \le 2^j \\ 0 & \text{if} \quad |x|_h \le 2^{j+1} \end{cases}$$

and

$$\phi(2^{-j+1}x) = \begin{cases} 1 & \text{if} \quad |x|_h \le 2^{j-1} \\ 0 & \text{if} \quad |x|_h \le 2^j \end{cases}$$

Therefore

$$\psi(x) \,=\, \phi(2^{-j}x) - \phi(2^{-j+1}x) = 0, \quad \text{if} \quad |x|_h < 2^{j-1}, \quad \text{or} \quad |x|_h > 2^{j+1}.$$

In other words,

$$\text{supp}(\psi_j) \,\subset\, \{2^{j-1} \le |x|_h \le 2^{j+1}\}.$$

It follows that for arbitrary x, there exists at most two ψ_j's such that $x \in \text{supp}(\psi_j)$. Now we have

$$\begin{aligned} \sum_{j=-N}^{N} \psi_j(x) \,&= \big[\phi(2^N x) - \phi(2^{N+1}x)\big] + \big[\phi(2^{N-1}x) - \phi(2^N x)\big] + \cdots \\ &\quad + \big[\phi(2^{-N+1}x) - \phi(2^{-N+2}x)\big] + \big[\phi(2^{-N}x) - \phi(2^{-N+1}x)\big] \\ &= -\,\phi(2^{N+1}x) + \phi(2^{-N}x) = 1 \end{aligned}$$

if $2^{-N} \le |x|_h \le 2^N$. Therefore

$$\sum_{j=-\infty}^{\infty} \psi_j(x) \,\equiv\, 1.$$

We let

$$K_j(x) \,=\, \phi_j(x)K(x)$$

and

$$T_j(f)(x) \,=\, f * K_j(x).$$

Then

$$T(f)(x) \,=\, f * P.V.K \,=\, f * \sum_{j=-\infty}^{\infty} K_j(x) \,=\, \sum_{j=-\infty}^{\infty} T_j(f)(x).$$

If we can show that

- $\|T_j\| \,\le\, C$ where the constant C is independent of j;
- $\|T_j T_k^*\| \,\le\, C \cdot 2^{-|j-k|}$;
- $\|T_j^* T_k\| \,\le\, C \cdot 2^{-|j-k|}$.

Suppose the above three properties have been proved. If we let $a_j \,=\, \sqrt{2^{-|j|}}$, then the hypothesis of Cotlar-Knapp-Stein's theorem is satisfied. We may conclude then that finite sums of the T_j have norm that is bounded by C. An additional argument will be provided below to show that the same estimate holds for infinite sums.

Fact 1. If $T(g) = g * L$, then $\|T(g)\|_{L^2} \le \|L\|_{L^1} \cdot \|g\|_{L^2}$ by the generalized Minkowski's inequality.

In order to finish the proof of the theorem, it reduces to prove the claims. We first need to show that $\|K_j\|_{L^1} \le C$. Since K is homogeneous of degree $-2n-2$, we have

$$K(2^{-j}x) = (2^{-j})^{-2n-2}K(x) = 2^{j(2n+2)}K(x).$$

Hence we get

$$\begin{aligned} K_j(x) &= \psi_j(x)K(x) \;=\; 2^{-j(2n+2)}K(2^{-j}x)\psi_j(x) \\ &= 2^{-j(2n+2)}K(2^{-j}x)\big[\phi(2^{-j}x) - \phi(2^{-j+1}x)\big]. \end{aligned}$$

Thus,

$$\begin{aligned} \|K_j\|_{L^1} &= \int_{\mathbf{H}_n} |K_j(x)|dV(x) \\ &= \int_{\mathbf{H}_n} 2^{-j(2n+2)}|K(2^{-j}x)|\cdot\big|\phi(2^{-j}x) - \phi(2^{-j+1}x)\big|dV(x) \\ &= \int_{\mathbf{H}_n} 2^{j(2n+2)}2^{-j(2n+2)}|K(x)|\big|\phi(x) - \phi(2x)\big|dV(x) \\ &= \int_{\mathbf{H}_n} |K(x)|\big|\phi(x) - \phi(2x)\big|dV(x) \;=\; \int_{\mathbf{H}_n} |K_0(x)|dV(x) \;=\; C. \end{aligned}$$

This is exactly the first assertion. Before proving the second assertion, let us note the following:

Fact 2. If $T_1(g) = g * L_1$ and $T_2(g) = g * L_2$, then

$$T_i \circ T_2(g) \;=\; T_1(g * L_2) \;=\; (g * L_2) * L_1 \;=\; g * (L_2 * L_1).$$

Also, if $T(g) = g * L$, then

$$T^*(g) \;=\; g * L^*,$$

where $L^*(x) = \overline{L(x^{-1})}$. Here x^{-1} is the inverse of x in $\mathbf{H}_n$. We see this by calculating

$$\begin{aligned} \langle T^*(g), f\rangle &= \langle g, T(f)\rangle \\ &= \int_{\mathbf{H}_n} g(y)\overline{\left(\int_{\mathbf{H}_n} f(x)L(x^{-1}\cdot y)dV(x)\right)}dV(y) \\ &= \int_{\mathbf{H}_n} \overline{f(x)} \int_{\mathbf{H}_n} g(y)\overline{L(x^{-1}\cdot y)}dV(y)\, dV(x) \\ &= \langle g * L^*, f\rangle. \end{aligned}$$

Now let us turn to the second assertion. Without loss of generality, we may assume that $j \geq k$. Hence we need to show

$$\|T_jT_k^*\| \;\leq\; C\cdot 2^{k-j}.$$

From **Fact 2**, we know that $T_jT_k^*(f) = f * (K_k^* * K_j)$. Therefore, by the generalized Minkowski inequality, it is enough to show that

$$\|K_k^* * K_j\|_{L^1} \;\leq\; C\cdot 2^{k-j}.$$

We can write $K_k^* * K_j$ as follows:

$$\begin{aligned} \int_{\mathbf{H}_n} K_j(y)K_k^*(y^{-1}\cdot x)dV(y) &= \int_{\mathbf{H}_n} (-1)^{2n+2}K_j(x\cdot y^{-1})K_k^*(y)dV(y) \\ &= \int_{\mathbf{H}_n} K_j(x\cdot y^{-1})K_k^*(y)dV(y). \end{aligned} \tag{8.4}$$

Claim 1.

$$\int_{\mathbf{H}_n} K_j(x)dV(x) \;=\; \int_{\mathbf{H}_n} K_k^*(x)dV(x) \;=\; 0. \tag{8.5}$$

To see this, we calculate that

$$\begin{aligned}\int_{\mathbf{H}_n} K_j(x)dV(x) &= \int_{\mathbf{H}_n} K(x)\big[\phi(2^{-j}x)-\phi(2^{-j+1}x)\big]dV(x)\\ &= \int_\Sigma\int_0^\infty K(r\xi)\big[\phi_0(2^{-j}r)-\phi_0(2^{-j+1}r)\big]r^{2n+1}dr\,d\sigma(\xi)\\ &= \int_0^\infty r^{-2n-2}r^{2n+1}\big[\phi_0(2^{-j}r)-\phi_0(2^{-j+1}r)\big]\int_\Sigma K(\xi)d\sigma(\xi)\,dr \;=\; 0.\end{aligned}$$

As a result,

$$\begin{aligned}\int_{\mathbf{H}_n} K_k^*(x)dV(x) = \int_{\mathbf{H}_n}\overline{K_k(-x)}dV(x) \;&=\; \int_{\mathbf{H}_n}(-1)^{2n+2}\overline{K_k(x)}dV(x)\\ = \overline{\int_{\mathbf{H}_n} K_k(x)dV(x)} &= 0.\end{aligned}$$

Thus, from (8.4) and (8.5), we can rewrite $K_j * K_k^*$ as follows:

$$K_j * K_k^* \;=\; \int_{\mathbf{H}_n} K_j(x\cdot y^{-1})K_k^*(y)dV(y) \;=\; \int_{\mathbf{H}_n}\big[K_j(x\cdot y^{-1})-K_j(x)\big]K_k^*(y)dV(y).$$

Claim 2.

$$\int_{\mathbf{H}_n}\big|K_j(x\cdot y^{-1})-K_j(x)\big|dV(x) \;=\; C\cdot 2^{-j}|y|_h. \tag{8.6}$$

To see this, recall that

$$K_j(x) \;=\; K(x)\psi_j(x) \;=\; K(x)\big[\phi(2^{-j}x)-\phi(2^{-j+1}x)\big]$$

and

$$K_0(x) = K(x)\big[\phi(x)-\phi(2x)\big].$$

Therefore,

$$K_j(2^jx) \;=\; K(2^jx)\psi_j(2^jx) \;=\; K(2^jx)\big[\phi(x)-\phi(2x)\big] \;=\; 2^{j(-2n-2)}K_0(x).$$

Hence we obtain

$$\begin{aligned}&\int_{\mathbf{H}_n}\big|K_0(x\cdot y^{-1})-K_0(x)\big|dV(x) \;\le\; C\cdot|y|_h\\ \Leftrightarrow &\int_{\mathbf{H}_n}\big|K_j(2^j(x\cdot y^{-1}))-K_j(2^jx)\big|dV(x) \;\le\; C\cdot 2^{-j(2n+2)}|y|_h\\ \Leftrightarrow &\int_{\mathbf{H}_n} 2^{-j(2n+2)}\big|K_j(x\cdot(2^jy)^{-1})-K_j(x)\big|dV(x) \;\le\; C\cdot 2^{-j(2n+2)}|y|_h\\ \Leftrightarrow &\int_{\mathbf{H}_n}\big|K_j(x\cdot y^{-1})-K_j(x)\big|dV(x) \;\le\; C\cdot\Big|\frac{y}{2^j}\Big|_h \;=\; C\cdot 2^{-j}|y|_h.\end{aligned}$$

Therefore, to prove (8.6), we only need to show that

$$\int_{\mathbf{H}_n}\big|K_0(x\cdot y^{-1})-K_0(x)\big|dV(x) \;\le\; C\cdot|y|_h.$$

First, suppose that $|y|_h \geq 1$. Then

$$\begin{aligned}
&\int_{\mathbf{H}_n} \big|K_0(x\cdot y^{-1}) - K_0(x)\big| dV(x) \;\leq\; \int_{\mathbf{H}_n} \big|K_0(x\cdot y^{-1})\big| dV(x) + \int_{\mathbf{H}_n} \big|K_0(x)\big| dV(x)\\
&= 2\int_{\mathbf{H}_n} \big|K_0(x)\big| dV(x) \;=\; 2\int_\Sigma \int_0^\infty \big|K_0(r\xi)\big| r^{2n+1} dr\, d\sigma(\xi)\\
&= 2\int_\Sigma \int_0^\infty \big|K(r\xi)\big| \cdot \big|\phi_0(r) - \phi_0(2r)\big| r^{2n+1} dr\, d\sigma(\xi)\\
&= 2\int_\Sigma \int_0^\infty \big|K(\xi)\big| \cdot \big|\phi_0(r) - \phi_0(2r)\big| r^{-2n-2} r^{2n+1} dr\, d\sigma(\xi)\\
&\leq C \cdot \int_\Sigma \big|K(\xi)\big| \int_{2^{-1}}^2 \frac{1}{r} dr\, d\sigma(\xi) \;\leq\; C \log 4 \;\leq\; C'.
\end{aligned}$$

Thus we have

$$\int_{\mathbf{H}_n} \big|K_0(x\cdot y^{-1}) - K_0(x)\big| dV(x) \;\leq\; C|y|_h, \qquad \text{if} \quad |y|_h \geq 1.$$

Now suppose that $|y|_h < 1$. We may consider K_0 as a function on $\mathbb{R}^{2n+1}$. We use the notation $\widetilde{K}_0$ to denote such a function. Since y^{-1}, the inverse of y in $\mathbf{H}_n$, corresponds to $-y$ in $\mathbb{R}^{2n+1}$, we get

$$\big|K_0(x\cdot y^{-1}) - K_0(x)\big| \;=\; \big|\widetilde{K}_0(x-y) - \widetilde{K}_0(x)\big|.$$

Thus, using the mean value theorem, we have

$$\big|\widetilde{K}_0(x-y) - \widetilde{K}_0(x)\big| \;\leq\; C|y|_e.$$

Therefore, from Lemma 7.1, one has

$$\big|\widetilde{K}_0(x-y) - \widetilde{K}_0(x)\big| \;\leq\; C|y|_h.$$

Hence,

$$\begin{aligned}
&\int_{\mathbf{H}_n} \big|K_0(x\cdot y^{-1}) - K_0(x)\big| dV(x)\\
&\leq C \cdot |y|_h \int_{\mathbf{H}_n} \chi_{\text{supp}(|K_0(x\cdot y^{-1}) - K_0(x)|)} dV(x) \qquad (8.7)\\
&\leq C \cdot |y|_h \left(\int_{\mathbf{H}_n} \chi_{\text{supp}(K_0(x\cdot y^{-1}))} dV(x) + \int_{\mathbf{H}_n} \chi_{\text{supp}(K_0(x\cdot y^{-1}))} dV(x) \right).
\end{aligned}$$

We certainly have

$$\text{supp}(K_0(x\cdot y^{-1})) \;\subset\; \big\{2^{-1} \leq |x\cdot y^{-1}| \leq 2\big\} \quad \text{and} \quad \text{supp}(K_0(x)) \;\subset\; \big\{2^{-1} \leq |x\cdot y^{-1}| \leq 2\big\}.$$

Since $|y|_h < 1$, if $x \in \text{supp}(K_0(x\cdot y^{-1}))$, we have

$$|x|_h \;\leq\; \gamma\big(|x\cdot y^{-1}|_h + |y|_h\big) \;\leq\; 3\gamma.$$

Therefore,

$$\text{supp}(K_0(x\cdot y^{-1})) \;\subset\; \{|x|_h \leq 3\gamma\}.$$

Hence we can rewrite (8.7) as follows:

$$\int_{\mathbf{H}_n} \big|K_0(x\cdot y^{-1}) - K_0(x)\big| dV(x) \;\leq\; C \cdot |y|_h (3\gamma)^{2n+2} \;=\; C' \cdot |y|_h.$$

Thus the **Claim 2** is proved.

Now let us concentrate on $\|K_j * K_k^*\|_{L^1}$:

$$\begin{aligned}\|K_j * K_k^*\|_{L^1} &= \int_{\mathbf{H}_n} \Big| \int_{\mathbf{H}_n} \big|K_j(x\cdot y^{-1}) - K_j(x)\big| K_k^*(y) dV(y) \Big| dV(x) \\ &\le \int_{\mathbf{H}_n} \big|K_k^*(y)\big| \int_{\mathbf{H}_n} \big|K_j(x\cdot y^{-1}) - K_j(x)\big| dV(x) dV(y) \\ &\le \int_{\mathbf{H}_n} \big|K_k^*(y)\big| (C\cdot 2^{-j}|y|_h) dV(y) \\ &= C\cdot 2^{-j} \int_{\mathbf{H}_n} \big|\overline{K_k(y^{-1})}\big| \cdot |y|_h dV(y) \\ &= C\cdot 2^{-j} \int_{\mathbf{H}_n} (-1)^{2n+2} \big|K_k(y)\big| \cdot |y|_h dV(y) \\ &= C\cdot 2^{-j} \int_{\Sigma}\int_0^\infty \big|K_k(r\xi)\big| \cdot r\cdot r^{2n+1} dr\, d\sigma(\xi) \\ &= C\cdot 2^{-j} \int_{\Sigma}\int_0^\infty \big|K(r\xi)\big| \cdot |\psi_k(r\xi)| r^{2n+2} dr\, d\sigma(\xi) \\ &= C\cdot 2^{-j} \int_{\Sigma} \big|K_k(\xi)\big| \int_0^\infty r^{-2n-2}\cdot r^{2n+2} |\psi_k(r\xi)| dr\, d\sigma(\xi) \\ &\le C\cdot 2^{-j} \int_{2^{k-1}}^{2^{j+1}} dr \ \le\ C\cdot 2^{k-j}.\end{aligned}$$

Hence the second assertion is proved. The proof of the third assertion is similar.

Now we invoke the Cotlar-Knapp-Stein lemma and get

$$\Big\| \sum_{\ell=1}^{M} T_{j_\ell} \Big\| \ \le\ C, \qquad \forall\ M\in\mathbb{N}.$$

We actually wish to consider

$$T_\varepsilon^N(f)(x) \ =\ \int_{\varepsilon\le |y|_h\le N} f(x\cdot y^{-1}) K(y) dV(y)$$

and let $\varepsilon\to 0$ and $N\to\infty$.

Claim 3.

$$\big\|T_\varepsilon^N(f)\big\|_{L^2} \ \le\ C\|f\|_{L^2},$$

where C is independent of ε and N.

For the proof of the **Claim 3**, let

$$K_\varepsilon^N(y) \ =\ K(y)\chi_{[\varepsilon,N]}(|y|_h).$$

Then

$$T_\varepsilon^N(f)(x) \ =\ \int_{\mathbf{H}_n} f(x\cdot y^{-1}) K_\varepsilon^N(y) dV(y).$$

Therefore, to prove the **Claim 3**, we will show that $\|K_\varepsilon^N\|_{L^1}\le C$, where C is independent of ε and N.

We may find $j,k\in\mathbb{Z}$ such that

$$2^{j-1}\le\varepsilon<2^j \qquad \text{and} \qquad 2^k\le\varepsilon<2^{k+1}$$

and want to compare $\sum_{j\le\ell\le k} T_\ell$ and T_ε^N. So we look at

$$\left(\sum_{j\le\ell\le k} K_\ell\right) - K_\varepsilon^N.$$

Note that

$$\begin{aligned}\sum_{j\le\ell\le k} K_\ell &= K(x)\big[\psi_j(x)+\cdots+\psi_k(x)\big]\\ &= K(x)\big[\phi(2^{-j}x)-\phi(2^{-j+1}x)+\phi(2^{-j-1}x)-\phi(2^{-j}x)\pm\cdots\\ &\quad +\phi(2^{-k+1}x)-\phi(2^{-k+2}x)+\phi(2^{-k}x)-\phi(2^{-k+1}x)\big]\\ &= K(x)\big[\phi(2^{-j}x)-\phi(2^{-j+1}x)\big].\end{aligned}$$

It follows that

$$\operatorname{supp}\left(\sum_{j\le\ell\le k} K_\ell\right) \subset \left\{2^{j-1}\le |x|_h \le 2^{k+1}\right\}$$

and

$$\sum_{j\le\ell\le k} K_\ell(x) = K(x), \qquad \text{if} \quad 2^j \le |x|_h \le 2^k.$$

Hence,

$$\begin{aligned}\operatorname{supp}\left(\Big[\sum_{j\le\ell\le k} K_\ell\Big] - K_\varepsilon^N\right) &\subset \Big[\big\{2^{j-1}\le |x|_h\le 2^j\big\}\bigcup\big\{2^k\le |x|_h\le 2^{k+1}\big\}\Big]\\ &\subset \Big[\big\{2^{-1}\varepsilon\le |x|_h\le 2\varepsilon\big\}\bigcup\big\{2^{-1}N\le |x|_h\le 2N\big\}\Big].\end{aligned}$$

Therefore,

$$\begin{aligned}\Big\|\sum_{j\le\ell\le k} K_\ell - K_\varepsilon^N\Big\|_{L^1} &\precsim \int_{2^{-1}\varepsilon\le|x|_h\le 2\varepsilon} |K(x)|dV(x) + \int_{2^{-1}N\le|x|_h\le 2N} |K(x)|dV(x)\\ &= \int_\Sigma\int_{2^{-1}\varepsilon}^{2\varepsilon} |K(r\xi)|r^{2n+1}dr\,d\sigma(\xi) + \int_\Sigma\int_{2^{-1}N}^{2N} |K(r\xi)|r^{2n+1}dr\,d\sigma(\xi)\\ &= \int_{2^{-1}\varepsilon}^{2\varepsilon} r^{2n+1}r^{-2n-2}dr\int_\Sigma |K(\xi)|d\sigma(\xi)\\ &\quad + \int_{2^{-1}N}^{2N} r^{2n+1}r^{-2n-2}dr\int_\Sigma |K(\xi)|d\sigma(\xi)\\ &= C'\big(\log 4+\log 4\big) = C.\end{aligned}$$

It follows that

$$\Big\|K_\varepsilon^N\Big\|_{L^1} \le \Big\|\sum_{j\le\ell\le k} K_\ell\Big\|_{L^1} + \Big\|\sum_{j\le\ell\le k} K_\ell - K_\varepsilon^N\Big\|_{L^1} \le C.$$

Hence, applying Functional Analysis Principle I (Theorem 8.3), the proof of the theorem is therefore complete. □

As we can see in the proof of Theorem 8.1, one needs to apply the following Functional Analysis Principle. Readers can find the proof of it in many standard functional analysis book.

Theorem 8.3. *Let X be a Banach space and S a dense subset. Let $T_j : X \to X$ be linear operators. Suppose that*
(a). For each $s \in S$, $\lim_{j\to\infty} T_j(s)$ exists in the Banach space norm;
(b). There is a finite constant $C > 0$, independent of x, such that

$$\|T_j(x)\|_X \le C \cdot \|x\|, \qquad \forall\ x \in X \quad \text{and} \quad \forall\ j.$$

Then $\lim_{j\to\infty} T_j(x)$ exists for every $x \in X$.

9. The Szegö Kernel on the Siegel Upper Half Space

Let us begin with the classical upper half-plane in $\mathbb{C}^1$, $U = \{x + iy : y > 0\}$, and it associate Hardy space:

$$\mathcal{H}^2(U) = \Big\{ f \in H(\mathbb{R}^2_+) : \sup_{y>0} \int_{\infty}^{\infty} |f(x+iy)|^2 dx < \infty \Big\}.$$

Obvious, $\mathcal{H}^2(U)$ is a Hilbert space with norm given by

$$\|f\|_{\mathcal{H}^2} = \sup_{y>0} \Big(\int_{\infty}^{\infty} |f(x+iy)|^2 dx \Big)^{\frac{1}{2}}.$$

The classical structure for this space is the following Paley-Wiener theorem:

Theorem 9.1. *The equation*

$$f \mapsto F(z) \equiv \int_0^{\infty} e^{2\pi z\cdot\lambda} f(\lambda) d\lambda \tag{9.1}$$

yields an isomorphism between $L^2(\mathbb{R}_+)$ and $\mathcal{H}^2(U)$.

Observe that one direction of this theorem is easy: given a function $f \in L^2(\mathbb{R}_+)$, the integral in (9.1) converges absolutely as long as $y = \text{Im}(z) > 0$. Furthermore, for any $y > 0$, we set $F_y(x) = F(x+iy)$ and see that

$$\begin{aligned}\|F_y\|^2_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} \Big| \int_0^{\infty} e^{-2\pi y\lambda} e^{2\pi i x\lambda} f(\lambda) d\lambda \Big|^2 dx \\ &\le \int_{\mathbb{R}} \Big| \int_0^{\infty} e^{2\pi i x\lambda} f(\lambda) d\lambda \Big|^2 dx \\ &= \|f(x)\|^2_{L^2(\mathbb{R})} = \|f(x)\|^2_{L^2(\mathbb{R}_+)}.\end{aligned}$$

It is also clear that

$$\|F\|_{\mathcal{H}^2(\mathbb{R}^2_+)} = \sup_{y>0} \|F_y\|_{L^2(\mathbb{R})} = \|f(x)\|_{L^2(\mathbb{R}_+)}.$$

The more difficult direction is the assertion that the map $f(\lambda) \mapsto F(z)$ is actually **onto** $\mathcal{H}^2(\mathbb{R}^2_+)$. We shall not treat it in detail. but refer the reader instead to Stein and Weiss's book [21] or or Katznelson's book [15].

We would like to develop an analogue for the Paley-Wiener theorem on the Siegel upper half space Ω_{n+1}. First we must discuss integration on $\mathbf{H}_n$. Recall that a measure dm on a topological group is the Haar measure (unique up to multiplication by a constant) if it is a Borel measure that is invariant under left translation. Our measure $dw'ds$ (the usual Lebesgue measure) turns out to be both left and right invariant, *i.e.,* it is *unimodular*. The proof is simply a matter of carrying out the integration:

$$\begin{aligned}\int_{\mathbf{H}_n} f((z',t)\cdot(w',s)) dw'ds &= \int_{\mathbf{H}_n} f(z'+w', t+s+2\text{Im}(z'\cdot\overline{w'})) dw'ds \\ &= \int_{\mathbf{H}_n} f(z'+w', s) dw'ds = \int_{\mathbf{H}_n} f(w', s) dw'ds.\end{aligned}$$

Observe now that the map $(w', s) \mapsto (-w', -s)$ preserves the measure but also sends an element of $\mathbf{H}_n$ onto its inverse. Thus it sends left translation into right translation, and so that left invariance of the measure implies its right invariance. With that preliminary step out of the way, we can make the following.

Definition 9.2. Define

$$\|f\|_{\mathcal{H}^2(\Omega_{n+1})} = \sup_{\rho>0} \left(\int_{\mathbb{C}^n} \int_{\mathbb{R}} \left| f(z', t + i|z'|^2 + i\rho) \right|^2 dz' dt \right)^{\frac{1}{2}}.$$

Then we set

$$\mathcal{H}^2(\Omega_{n+1}) = \left\{ f \in H(\Omega_{n+1}) : \ \|f\|_{\mathcal{H}^2(\Omega_{n+1})} < \infty \right\}.$$

Here ρ is the "height function" $\rho = \mathrm{Im}(w_{n+1}) - |w'|^2$ on Ω_{n+1}.

Now, just as in the case of $U \subset \mathbb{C}^1$, where we integrated over level sets $\{x+iy \in \mathbb{R}^2_+ : \ y = \text{constant}\}$ which are parallel to $\mathbb{R}^1 = \partial\mathbb{R}^2_+$, so here we integrate over level sets ρ =constant which are parallel to $\mathbf{H}_n = \partial\Omega_{n+1}$.

Let us prove that $\mathcal{H}^2(\Omega_{n+1})$ is a Hilbert space first. The substitute for $L^2(\mathbb{R}_+)$ in the present case will be $\widetilde{\mathcal{H}}^2$ which consists of all function $\tilde{f} = \tilde{f}(z', \lambda)$ with $z' \in \mathbb{C}^n$ and $\lambda \in \mathbb{R}_+$ such that

(1). $\tilde{f}$ is jointly measurable in z' and λ;
(2). For almost every λ, $z' \mapsto \tilde{f}(z', \lambda)$ is entire on $\mathbb{C}^n$;
(3).

$$\|\tilde{f}\|^2_{\widetilde{\mathcal{H}}^2} = \int_{\mathbb{C}^n} \int_0^\infty |\tilde{f}(z', \lambda)|^2 e^{-4\pi\lambda|z'|^2} dz' d\lambda < \infty \big\}.$$

We have the following basic structure theorem:

Theorem 9.3. *Consider the equation*

$$F(z) = F(z', z_{n+1}) = \int_0^\infty e^{2\pi i\lambda z_{n+1}} f(z', \lambda) d\lambda. \tag{9.2}$$

(1). *Given an $f \in \widetilde{\mathcal{H}}^2$, the integral in (9.2) converges absolutely for $z \in \Omega_{n+1}$ and uniformly for $z \in K \subset\subset \Omega_{n+1}$. Thus we can interchange the order of differentiation and integration, and we see that the function F given by the integral is holomorphic.*
(2). *The function F defined in part (9.2) from an $f \in \widetilde{\mathcal{H}}^2$ is an element of $\mathcal{H}^2(\Omega_{n+1})$, and the resulting map $\tilde{f} \mapsto F$ is an isometry of $\widetilde{\mathcal{H}}^2$ onto $\mathcal{H}^2(\Omega_{n+1})$; i.e., it is an isomorphism of Hilbert spaces.*
(3). *Let $\hat{\imath} = (0, \dots, 0, i) \in \mathbb{C}^{n+1}$ and let $f \in \mathcal{H}^2(\Omega_{n+1})$. Set $f_\varepsilon = f(z + \varepsilon\hat{\imath})\Big|_{\partial\Omega_{n+1}}$. Then f_ε is a function on $\mathbf{H}_n$,*

$$f_\varepsilon \to f_0 \quad \text{in} \quad L^2(\mathbf{H}_n) \quad \text{as } \varepsilon \to 0,$$

and

$$\|f_0\|_{L^2(\mathbf{H}_n)} = \|f\|_{\mathcal{H}^2}.$$

The idea of the proof is to freeze the z' variable and look at the Paley-Wiener representation of the half-space $\mathrm{Im}(z_{n+1}) < |z'|^2$. However, thee are several nontrivial technical problems with this process, so we shall have to develop the proof stages. First, we want to show that $\widetilde{\mathcal{H}}^2$ is a Hilbert space.

Lemma 9.4. *$\widetilde{\mathcal{H}}^2$ is a Hilbert space.*

Proof. Since $\widetilde{\mathcal{H}}^2$ is defined as $L^2(\mathbb{C}^n \times \mathbb{R}_+, dm)$ for a certain measure m, its inner product is already determined. It remains to proof the completeness of the space.

Since we are dealing with holomorphic functions, the L^2 convergence will lead to a very strong (*i.e.,* uniform on compact subsets) type of convergence on the interior of $\mathbb{C}^n \times \mathbb{R}$. Now suppose we are given a Cauchy sequence in $\widetilde{\mathcal{H}}^2$; we must show that some subsequence converges to an element of $\widetilde{\mathcal{H}}^2$. Since $\mathcal{H}^2$ is an L^2 space and L^2 being complete, some subsequence converges in L^2, and we can extract from that a subsequence converging both in L^2 and pointwise almost everywhere. Next take a compact set $K \subset\subset \mathbb{C}^n$ which is the closure of an open set and $L \subset\subset \mathbb{R}_+$, and a subsequence $\{f_k\}$ such that

$$\int_J \int_K \left|f_k(z', \lambda) - f_{k+1}(z', \lambda)\right|^2 dz' d\lambda \le \frac{1}{2^{2k}}.$$

It follows that

$$\sum_k \left\|f_k - f_{k+1}\right\|_{K,L}^2 = \sum_k \int_J \int_K \left|f_k(z', \lambda) - f_{k+1}(z', \lambda)\right|^2 dz' d\lambda < \infty.$$

If we set

$$\Delta_k(\lambda) = \int_K \left|f_k(z', \lambda) - f_{k+1}(z', \lambda)\right|^2 dz'.$$

then one has

$$\int_L \sum_k \Delta_k(\lambda)\, d\lambda < \infty.$$

Thus $\sum_k \Delta_k(\lambda) < \infty$ for almost every $\lambda \in L$. Passing to a subset $K' \subset\subset K$ we find a number $\delta > 0$ such that $B(z'; \delta) \subset\subset K$ for all $z' \in K'$. Since, for a fixed λ, the functions f_k are holomorphic on K, they obey the mean value property. Hence,

$$\begin{aligned}
\left|f_k(z', \lambda) - f_{k+1}(z', \lambda)\right| &\le \frac{1}{c_n \delta^{2n}} \int_{B(z';\delta)} \left|f_k(w', \lambda) - f_{k+1}(w', \lambda)\right| dw' \\
&\le \frac{C_n}{\delta^{2n}} \left(\int_{B(z';\delta)} \left|f_k(w', \lambda) - f_{k+1}(w', \lambda)\right|^2 dw' \right)^{\frac{1}{2}} \\
&\le C_\delta \cdot \sqrt{\Delta_k(\lambda)}
\end{aligned}$$

for all $z' \in K'$ and λ fixed. Therefore, the sequence $\{f_k(\cdot, \lambda)\}$ converges uniformly on compact subsets of $\mathbb{C}^n$ for almost every λ. Since, for almost every λ the functions f_k are holomorphic, the limit is then holomorphic. Since the functions f_k already converge in $L^2(\mathbb{C}^n \times \mathbb{R}_+, dm)$ and pointwise almost everywhere, the limit is in $\widetilde{\mathcal{H}}^2$. The proof of the lemma is therefore complete. □

Next we need to prove the following result.

Lemma 9.5. *If $f \in \widetilde{\mathcal{H}}^2$, then for $(z', z_{n+1}) \in K \subset\subset \Omega_{n+1}$, we have that*

$$\int_0^\infty e^{2\pi i \lambda z_{n+1}} f(z', \lambda)\, d\lambda$$

converges absolutely. Its absolute value is $\le C_K \|f\|_{\widetilde{\mathcal{H}}^2}$.

Proof. For $(z', z_{n+1}) \in K \subset\subset \Omega_{n+1}$, there is an $\varepsilon > 0$ such that $\mathrm{Im}(z_{n+1}) - |z'|^2 \ge \varepsilon$. Since $f(z', \lambda)$ is entire in z' for almost every λ , we have by mean value property that

$$\left|f(z', \lambda)\right| \le \frac{1}{|\mathcal{B}(z', \delta)|} \int_{\mathcal{B}(z', \delta)} |f(w', \lambda)| dw'. \tag{9.3}$$

The number $\delta > 0$ will be determined later.

Since $\mathrm{Im}(z_{n+1}) \le -|z'|^2 - \varepsilon$, we calculate that

$$\begin{aligned}
&\left|\int_0^\infty e^{2\pi i\lambda z_{n+1}} f(z',\lambda)\,d\lambda\right| \\
&\le \int_0^\infty e^{-2\pi\varepsilon\lambda} e^{-2\pi\lambda|z'|^2} |f(z',\lambda)|\,d\lambda \\
&\le \left(\int_0^\infty e^{-2\pi\varepsilon\lambda}\,d\lambda\right)^{\frac{1}{2}} \cdot \left(\int_0^\infty e^{-2\pi\varepsilon\lambda} e^{-4\pi\lambda|z'|^2} |f(z',\lambda)|^2 d\lambda\right)^{\frac{1}{2}}
\end{aligned}$$

by Schwarz's inequality. Now set

$$\tilde{C} = \left(\int_0^\infty e^{-2\pi\varepsilon\lambda}\,d\lambda\right)^{\frac{1}{2}}$$

and apply (9.3) to obtain

$$\begin{aligned}
&\left|\int_0^\infty e^{2\pi i\lambda z_{n+1}} f(z',\lambda)\,d\lambda\right| \\
&\le \frac{C^\star}{|\mathcal{B}(z',\delta)|} \left(\int_0^\infty e^{-2\pi\varepsilon\lambda} e^{-4\pi\lambda|z'|^2} \Big(\int_{\mathcal{B}(z',\delta)} |f(w',\lambda)|dw'\Big)^2 d\lambda\right)^{\frac{1}{2}}.
\end{aligned}$$

But an application of Schwarz's inequality to the w' integration yields

$$\begin{aligned}
&\left|\int_0^\infty e^{2\pi i\lambda z_{n+1}} f(z',\lambda)\,d\lambda\right|^2 \\
&\le \tilde{C}^2 \cdot \int_0^\infty e^{-2\pi\varepsilon\lambda} e^{-4\pi\lambda|z'|^2} \cdot \int_{\mathcal{B}(z',\delta)} |f(w',\lambda)|dw'\,d\lambda.
\end{aligned}$$

Now we would like to replace the expression $e^{-4\pi\lambda|z'|^2}$ by $e^{-4\pi\lambda|w'|^2}$ and then apply condition (3) of the Definition 9.2. Since $w' \in \mathcal{B}(z';\delta)$, we see that

$$e^{-4\pi\lambda|w'|^2} \ge e^{-4\pi\lambda|z'|^2} \cdot e^{-4\pi\lambda\delta}.$$

We choose $0 < \delta < \frac{\varepsilon}{2}$. It follows that

$$e^{-2\pi\varepsilon\lambda} e^{-4\pi\lambda|z'|^2} = e^{-2\pi\varepsilon\lambda} e^{4\pi\lambda\delta} e^{-4\pi\lambda\delta} e^{-4\pi\lambda|z'|^2} \le e^{-4\pi\lambda|w'|^2}$$

and we find that

$$\left|\int_0^\infty e^{2\pi i\lambda z_{n+1}} f(z',\lambda)\,d\lambda\right|^2 \le \tilde{C}^2 \cdot \int_0^\infty \int_{\mathcal{B}(z',\delta)} e^{-4\pi\lambda|w'|^2} |f(w',\lambda)|^2 dw'\,d\lambda.$$

Hence we have

$$\int_0^\infty e^{2\pi i\lambda z_{n+1}} f(z',\lambda)\,d\lambda \le \tilde{C}^2 \cdot \le \|f\|_{\widetilde{\mathcal{H}}^2}.$$

This completes the proof of the lemma. □

Now that we have the absolute convergence of our integral and uniform convergence for $z \in K \subset\subset \Omega_{n+1}$, we are allowed to differentiate under the integral sign and it is clear that the function F which is created from $f \in \widetilde{\mathcal{H}}^2$ is holomorphic. We may continue to prove Theorem 9.3. Here are two observations.

(a). There would appear to be an ambiguity in the definition

$$f(z',\lambda) = \int_{-\infty}^\infty e^{-2\pi\lambda(x_{n+1}+iy_{n+1})} f(z', x_{n+1}+iy_{n+1})\,dx_{n+1} \quad \text{where} \quad y_{n+1} > |z'|^2.$$

After all, the right-hand side explicitly depends on y_{n+1}, and yet the left-hand side is independent of y_{n+1}. The fact is that the right-hand side is also independent of y_{n+1}. After all, f is holomorphic in the variable $x_{n+1} + iy_{n+1}$, as it ranges over the half-plane $y_{n+1} > |z'|^2$. Then our claim is simply that the integral of f over a line parallel to the x-axis is independent of the particular line we choose (as long as $y_{n+1} > |z'|^2$). This statement is a consequence of Cauchy's integral theorem: the difference of the integral of f over two parallel horizontal lines is the limit of the integral of f over long horizontal rectangles: from $-N$ to N say. Now the integral of f over a rectangle is zero, and we will see that f has sufficiently rapid decrease at ∞ so that then integrals over the ends of the rectangle tend to zero as $N \to \infty$. Therefore,

$$\int_{-\infty}^{\infty} e^{-2\pi\lambda(x_{n+1}+iy_{n+1})} f(z', x_{n+1} + iy_{n+1})\, dx_{n+1}$$
$$= \int_{-\infty}^{\infty} e^{-2\pi\lambda(x_{n+1}+iy'_{n+1})} f(z', x_{n+1} + iy'_{n+1})\, dx_{n+1}$$

for $0 < y_{n+1} < y'_{n+1}$.

(b). Fix a point $(z', z_{n+1}) \in \Omega_{n+1}$. Consider the functional which sends $f \in \widetilde{\mathcal{H}}^2$ to $F(z', z_{n+1})$, where F is the function created by the Fourier integral of f. Then this functional is continuous on $\widetilde{\mathcal{H}}^2$. However, the integral of f which yields F is taken over a 1-dimensional set, so how can the result be well-defined pointwise as a function?

The answer is that for almost every λ we are careful pick an almost everywhere equivalent of $f(z', \lambda)$ which is entire in z', so that the resulting F is holomorphic. Thus the precise definition of our linear functional is "evaluation at the point (z', z_{n+1}) of the holomorphic function which is an almost everywhere equivalent of the function F arising from f." We next prove the following lemma.

Lemma 9.6. *Let $F \in \mathcal{H}^2(\Omega_{n+1})$. Then, for a fixed z', $F_\varepsilon(z', \cdot) \in \mathcal{H}^2\big(\{y_{n+1} > |z'|^2\}\big)$ (as a function of one complex variable) where*

$$F_\varepsilon(z', x_{n+1} + iy_{n+1}) \;=\; F(z', x_{n+1} + iy_{n+1} + i\varepsilon), \qquad \text{for} \quad \varepsilon \;>\; 0.$$

Proof. We may assume $z' = 0$. Apply the mean value theorem to $F(0, x_{n+1} + iy_{n+1} + i\varepsilon)$ on $D(0; \delta') \times \mathcal{B}(0, \delta)$, where D is a disc in the plane and $\mathcal{B} \subset \mathbb{C}^n$. We see that

$$\big|F(0, x_{n+1} + iy_{n+1} + i\varepsilon)\big|^2 \;=\; C_{\delta',\delta} \cdot \int_{D(x_{n+1}+iy_{n+1}+i\varepsilon,\delta')} \int_{\mathcal{B}(0;\delta)} |F|^2 dz'\, dw.$$

Hence,

$$\int_{-\infty}^{\infty} \big|F(0, x_{n+1} + iy_{n+1} + i\varepsilon)\big|^2 dx_{n+1}$$
$$\leq C_{\delta',\delta} \cdot \int_{-\infty}^{\infty} \int_{D(x_{n+1}+iy_{n+1}+i\varepsilon,\delta')} \int_{\mathcal{B}(0;\delta)} |F|^2 dz'\, dw\, dx_{n+1}.$$

But

$$\int_{-\infty}^{\infty} \int_{D(x_{n+1}+iy_{n+1}+i\varepsilon,\delta')} |F(x + w)| dw\, dx \;\leq\; \int_{-\infty}^{\infty} \int_{-\delta'}^{\delta} |F(x + iv)| dv\, dx$$

for any F. We may choose $\delta' = \frac{\varepsilon}{3}$ and set $\varepsilon' = \frac{2\varepsilon}{3}$, to obtain

$$\int_{-\delta'}^{\delta} F_\varepsilon(x + iv) dv \;=\; \int_{0}^{2\delta'} F_{\varepsilon'}(x + iv) dv.$$

Thus

$$\int_{-\infty}^{\infty} |F_\varepsilon(0, x_{n+1} + iy_{n+1})|^2 dx_{n+1}$$
$$\leq C \cdot \int_{-\infty}^{\infty} \int_0^{2\delta'} \int_{\mathcal{B}(0;\delta)} |F(z', x_{n+1} + iy_{n+1} + iv + i\varepsilon')| dz' dv\, dx_{n+1}.$$

Now $|z'| < \delta$; we choose $\delta = \sqrt{\frac{\varepsilon'}{2}}$ so that $|z'|^2 < \frac{\varepsilon'}{2}$. Therefore

$$\big\|F_\varepsilon(0, \cdot)\big\|^2_{\mathcal{H}^2}$$
$$\leq C \cdot \int_{-\infty}^{\infty} \int_{\mathcal{B}(0;\delta)} \int_0^{2\delta'} \big|F(x_{n+1} + i(z', |z'|^2 + y_{n+1}) + i(v + \frac{\varepsilon'}{2} + \frac{\varepsilon'}{2} - |z'|^2)\big|^2 dv\, dz' dx_{n+1}.$$

Next set $\tilde{v} = v + \frac{\varepsilon'}{2} - |z'|^2$ and observe that

$$\int_0^{2\delta'} \big|F(0, v + \frac{\varepsilon'}{2} - |z'|^2)\big| dv \leq \int_0^{2\delta' + \frac{\varepsilon'}{2}} \big|F(0, \tilde{v})\big| d\tilde{v}.$$

But we know that $2\delta' + \frac{\varepsilon'}{2} = \varepsilon$ so we have

$$\begin{aligned}\big\|F_\varepsilon(0, \cdot)\big\|^2_{\mathcal{H}^2} &\leq C \cdot \int_0^{\varepsilon} \int_{-\infty}^{\infty} \int_{\mathcal{B}(0;\delta)} \Big|F\big(z', x_{n+1} + i(z', |z'|^2 + y_{n+1} + v + \frac{\varepsilon'}{2}\Big|^2 dz' dx_{n+1} dv \\ &\leq C \cdot \int_0^{\varepsilon} \|F\|^2_{\mathcal{H}^2(\Omega_{n+1})} dv = C_\delta \cdot \varepsilon \cdot \|F\|^2_{\mathcal{H}^2(\Omega_{n+1})} < \infty.\end{aligned}$$

This completes the proof of the lemma. □

Remark. Lemma 9.6 is not necessarily true for the boundary limit function $F(z', x_{n+1} + i|z'|^2)$. For the constant $C_\delta \sim \delta^{-n}$, hence the right-hand side blows up as $\varepsilon \to 0$ (and hence $\delta \to 0$). We are finally in a position to bring our calculations together and to prove Theorem 9.3. We have seen that from a given $f \in \widetilde{\mathcal{H}}^2$ we obtain a function $F(z', z_{n+1})$, holomorphic in Ω_{n+1}. We now show that it is in $\mathcal{H}^2$ and in fact that its $\mathcal{H}^2$ norm equals $\|f\|_{\widetilde{\mathcal{H}}^2}$. Now

$$\begin{aligned}&\int_{\mathbb{C}^n} \int_{-\infty}^{\infty} \big|F(z', x_{n+1} + i|z'|^2 + i\rho)\big|^2 dx_{n+1} dz' \\ &= \int_{\mathbb{C}^n} \int_0^{\infty} \big|\widetilde{F}(z', \lambda)\big|^2 e^{-4\pi\lambda(|z'|^2 + \rho)} d\lambda\, dz'\end{aligned}$$

and the integral on the right increases to

$$\int_{\mathbb{C}^n} \int_0^{\infty} \big|\widetilde{F}(z', \lambda)\big|^2 e^{-4\pi\lambda|z'|^2} d\lambda\, dz', \qquad \text{as} \qquad \rho \to 0.$$

Hence,

$$\begin{aligned}\|F\|^2_{\mathcal{H}^2(\Omega_{n+1})} &= \sup_{\rho > 0} \int_{\mathbb{C}^n} \int_{-\infty}^{\infty} \big|F(z', x_{n+1} + i|z'|^2 + i\rho)\big|^2 dx_{n+1} dz' \\ &= \int_{\mathbb{C}^n} \int_0^{\infty} \big|\widetilde{F}(z', \lambda)\big|^2 e^{-4\pi\lambda|z'|^2} d\lambda\, dz' = \|f\|^2_{\widetilde{\mathcal{H}}^2}.\end{aligned}$$

Furthermore, this equality of norms implies that our map from $\widetilde{\mathcal{H}}^2$ to $\mathbf{H}^2$ is injective. All that remains is to show that an arbitrary $F \in \mathcal{H}^2$ has such a representation.

Given $F = F(z', z_{n+1}) \in \mathcal{H}^2(\Omega_{n+1})$, Lemma 9.6 tells us that for any fixed $\varepsilon > 0$ and $z' \in \mathbb{C}^n$, the function $F_\varepsilon(z', z_{n+1}) \equiv F(z', z_{n+1} + i\varepsilon)$ has a Paley-Wiener representation. Moreover, the function $F_\varepsilon(z', \lambda)$ is holomorphic in the z'-variable. Since we have the relation

$$F_\varepsilon(z', z_{n+1}) \;=\; \int_0^\infty \widetilde{F}_\varepsilon(z', \lambda) e^{2\pi\lambda z_{n+1}} d\lambda$$

and since the functions $\{F_\varepsilon\}$ are uniformly bounded in $\widetilde{\mathcal{H}}^2$ as $\varepsilon \to 0$, it follows that the functions $\{\widetilde{F}_\varepsilon\}$ are uniformly bounded in $\widetilde{\mathcal{H}}^2$. We can therefore extract a subsequence $\{F_{\varepsilon_j}\}$ such that $F_{\varepsilon_j} \to f_0$ weakly as $j \to \infty$. Observe that since $f_0 \in \widetilde{\mathcal{H}}^2$ we can recover from it $F_0 \in \mathcal{H}^2(\Omega)$.

Lemma 9.5 tells us that for $(z', z_{n+1}) \in K \subset\subset \Omega_{n+1}$, the (continuous) linear functional on $\widetilde{\mathcal{H}}^2$ given by Fourier inversion and then evaluation at the point (z', z_{n+1}) is uniformly bounded:

$$\left|F(z', z_{n+1})\right| \;\leq\; C_K \cdot \|\tilde{F}\|_{\widetilde{\mathcal{H}}^2}.$$

Thus $F_{\varepsilon_j}(z', z_{n+1}) \to F_0(z', z_{n+1})$ uniformly on compact subsets of Ω_{n+1}. However,

$$F_{\varepsilon_j}(z', z_{n+1}) \;\equiv\; F(z', z_{n+1} + i\varepsilon_j) \;\to\; F(z', z_{n+1})$$

pointwise, so we know that $f_0 \equiv F$. Thus F has a representation in terms of a function in $\widetilde{\mathcal{H}}^2$ because f_0 does.

Finally we must show that, if F_ε is defined as above, then F_ε converges to a function f in $L^2(\partial\Omega_{n+1})$. But we see that

$$F_\varepsilon(z) \;=\; F(z', z_{n+1} + i\varepsilon) \;=\; \int_0^\infty e^{2\pi\lambda z_{n+1}} \cdot e^{-2\pi\lambda\varepsilon} \tilde{F}(z', \lambda) d\lambda$$

so that

$$\int_{\mathbf{H}_n} \left|F(z', z_{n+1} + i\varepsilon)\right|^2 dz' dx_{n+1} \;=\; \int_0^\infty \int_{\mathbb{C}^n} e^{-4\pi\lambda\varepsilon} \left|\tilde{F}(z', \lambda)\right|^2 e^{-4\pi\lambda|z'|^2} dz' d\lambda$$

and

$$\int_{\mathbf{H}_n} \left|F_{\varepsilon_1}(z) - F_{\varepsilon_2}(z)\right|^2 dz' dx_{n+1} \;=\; \int_0^\infty \int_{\mathbb{C}^n} \left|e^{-2\pi\lambda\varepsilon_1} - e^{-2\pi\lambda\varepsilon_2}\right|^2 \cdot \left|\tilde{F}(z', \lambda)\right|^2 e^{-4\pi\lambda|z'|^2} dz' d\lambda.$$

Thus the Lebesgue Dominated Convergence Theorem tells us that $\{F_\varepsilon\}$ is a Cauchy sequence in $L^2(\mathbf{H}_n)$. It follows that f has boundary values in $L^2(\mathbf{H}_n)$.

As a direct consequence of Lemma 9.5, we have the following corollary:

Lemma 9.7. *$\mathcal{H}^2(\Omega_{n+1})$ is a Hilbert space with reproducing kernel.*

The reproducing kernel for $\mathcal{H}^2$ is the Cauchy-Szegö kernel; we shall see, by symmetry considerations, that it is uniquely determined up to a constant. Define $S(z, w)$ to be the reproducing kernel for $\mathcal{H}^2(\Omega_{n+1})$.

Theorem 9.8. *On the Siegel upper half space Ω_{n+1}, the Szegö kernel $S(z, w)$ is*

$$S(z, w) \;=\; \frac{C_n}{\rho^{n+1}(z, w)} \;=\; \frac{C_n}{\left[\frac{i}{2}(\bar{w}_{n+1} - z_{n+1}) - \sum_{k=1}^n z_k \bar{w}_k\right]^{n+1}},$$

where

$$C_n \;=\; \frac{n!}{4\pi^{n+1}}.$$

Observe that ρ is a polarization of our "height function" $\rho(z) = \mathrm{Im}(z_{n+1}) - |z'|^2$, for $\rho(z,w)$ is holomorphic in z, antiholomorphic in w, and $\rho(z,z) = \rho(z)$. It is common to refer to new function ρ as an "*almost analytic continuation*" of the old function.

Before we prove that theorem we will formulate an important corollary. Since all our constructs are canonical, the Cauchy-Szegö representation ought to be modeled on a simple convolution operator on the Heisenberg group. Let us determine how to write the reproducing formula as a convolution.

A function F defined on Ω_{n+1} induces, for each value of the "height" ρ, a function on the Heisenberg group:

$$F_\rho(z',t) \;=\; F\big(z', t + i(|z'|^2 + \rho)\big).$$

Since $S(z,w)$ is the reproducing kernel, we know that

$$F(z) \;=\; \int_{\mathbf{H}_n} F(w)S(z,w)dm(w) \tag{9.4}$$

where $dm(w) = dw'ds$ is the Haar measure on $\mathbf{H}_n$ with $w = (w', s + i|w'|^2)$. Recall part (3) of Theorem 9.3 guarantees the existence of L^2 boundary values for F, and the boundary of $\partial\Omega_{n+1}$ is $\mathbf{H}_n$. Thus the integral (9.4) is well-defined. This is the corollary.

Corollary 9.9. *We have that*

$$F_\rho(z',t) \;=\; F_0 * K_\rho(z',t),$$

where F_0 is the L^2 boundary limit of F, and

$$K_\rho(z',t) \;=\; \frac{2^{n-1}n!}{\pi^{n+1}}\frac{1}{(|z'|^2 - it + \rho)^{n+1}}.$$

Proof. We write

$$F_\rho(z',t) \;=\; \int_{\mathbf{H}_n} S\big((z', t + i|z'|^2 + i\rho), (w', s + i|w'|^2)\big)F(w', s + i|w'|^2)dm(w).$$

Therefore,

$$\begin{aligned}
&F_\rho(z',t)\\
&= C_n \int_{\mathbf{H}_n} \frac{F(w', s + i|w'|^2)}{\big(\frac{i}{2}(s - i|w'|^2 - t - i\rho - i|z'|^2) - \sum_{k=1}^n z_k\bar{w}_k\big)^{n+1}} dm(w)\\
&= 2^{n+1}C_n \int_{\mathbf{H}_n} \frac{F(w', s + i|w'|^2)}{\big[|z'|^2 + |w'|^2 - 2\mathrm{Re}(z'\cdot\overline{w'}) + \rho - i(s - t + 2\mathrm{Im}(z'\cdot\overline{w'}))\big]^{n+1}} dm(w)\\
&= \int_{\mathbf{H}_n} F_0(w',s)K_\rho\big((z',t)^{-1}\cdot(w',s)\big)dm(w).
\end{aligned}$$

This completes the proof of the corollary. □

Proof of Theorem 9.8: First we need the following elementary uniqueness result from complex analysis. We know that if $\rho(z,w)$ is holomorphic in z and antiholomorphic in w then it is uniquely determined by $\rho(z,z) = \rho(z)$. Next we demonstrate

Claim (1). If g is an element of $\mathbf{H}_n$ then $S(g\circ z, g\circ w) \equiv S(z,w)$.

After all, if $F \in \mathcal{H}^2(\Omega_{n+1})$ then the map $F \mapsto F_g$ (where $F_g(z) = F(g\circ z)$) is a unitary map of $\mathcal{H}^2(\Omega_{n+1})$ to itself. Now

$$F(g\circ z) \;=\; \int_{\mathbf{H}_n} S(z,w)F(g\circ w)dm(w).$$

We make the change of variables $\tilde{w} = g \circ w$; since dm is the Haar measure on $\mathbf{H}_n$, it follows that $dm(\tilde{w}) = dm(w)$. Thus

$$F(g \circ z) \;=\; \int_{\mathbf{H}_n} S(z, g^{-1} \circ \tilde{w}) F(\tilde{w}) dm(\tilde{w})$$

therefore,

$$F(z) \;=\; \int_{\mathbf{H}_n} S(g^{-1} \circ z, g^{-1} \circ w) F(w) dm(w).$$

We conclude that $S(z, w)$ and $S(g \circ z, g \circ w)$ are both reproducing kernels for $\mathcal{H}^2(\Omega_{n+1})$. Hence they are equal.

Claim (2). If δ is the natural dilation on Ω_{n+1} by

$$\delta(z', z_{n+1}) \;=\; (\delta z', \delta^2 z_{n+1})$$

then

$$S(\delta z, \delta w) \;=\; \delta^{-2n-2} S(z, w).$$

The proof is just as above:

$$\begin{aligned} F(\delta z) \;&=\; \int_{\mathbf{H}_n} S(z, w) F(\delta w) dm(w) \\ &=\; \int_{\mathbf{H}_n} S(z, \delta^{-1}\tilde{w}) F(\tilde{w}) \cdot \delta^{-2n-2} dm(w) \end{aligned}$$

so that

$$F(z) \;=\; \delta^{-(2n+2)} \int_{\mathbf{H}_n} S(\delta^{-1} z, \delta^{-1} w) F(w) dm(w)$$

Then the uniqueness of the reproducing kernel yields

$$S(z, w) \;=\; \delta^{-(2n+4)} S(\delta^{-1} z, \delta^{-1} w) \qquad \text{for all} \quad \delta > 0.$$

Now the uniqueness result following Theorem 6.4 shows that $S(z, w)$ will be completely determined if we can prove that

$$S(z, z) \;=\; \frac{C_n}{\rho^{n+1}(z)}.$$

However, $\rho(z)$ is invariant under translation of Ω_{n+1} by element of the Heisenberg group, *i.e.*, $\rho(g \circ z) = \rho(z)$ for all $g \in \mathbf{H}_n$ and

$$\rho(\delta z) \;=\; \text{Im}(\delta^2 z_{n+1}) - \delta^2 |z'|^2 \;=\; \delta^2 \rho(z).$$

Therefore the function

$$S(z, z) \cdot \rho^{n+1}(z)$$

has homogeneity zero and is invariant under the action of the Heisenberg group. Since the Heisenberg group acts simply transitively on "parallels" to $\partial\Omega_{n+1}$, and since dilations enables us to move from any one parallel to another, any function with these two invariance properties must be constant. Hence we have

$$S(z, z) \;\equiv\; C_n \cdot \rho^{-(n+1)}(z).$$

It follows that

$$S(z, w) \;\equiv\; C_n \cdot \rho^{-(n+1)}(z, w).$$

At long last we have proved Theorem 9.8. We have not taken the trouble to calculate the exact value of the constant in front of the canonical kernel. That value has no practical significance for us at this moment.

10. Estimates for Singular Integral Operators in Hardy Spaces

In this section, we shall prove that if a singular integral operator T defined by a kernel K satisfying $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ satisfies the generalized Hörmander's condition:

$$\int_{|x|\leq a|y|} |K(x-y)-K(x)|dx \leq C_1 \cdot a^{-\theta}$$

for all $a \geq 2$ and some $\theta \geq 0$ then the operator T extends to a bounded operator on $H^p(\mathbb{R}^n)$ for $p_\theta < p < \infty$. In fact, we have the following theorem.

Theorem 10.1. *Suppose that the kernel K satisfies the following condition:*

$$\int_{|x|\leq a|y|} |K(x-y)-K(x)|dx \leq C_1 \cdot a^{-\theta} \tag{10.1}$$

for all $a \geq 2$ and some $\theta > 0$. We define the mapping T by

$$T(f)(x) = \int_{\mathbb{R}^n} K(x-y)f(y)dy.$$

Assume that $\|T(f)\|_{L^2(\mathbb{R}^n)} \leq C_2 \cdot \|f\|_{L^2(\mathbb{R}^n)}$ holds. Then T extends to a bounded operator on $H^p(\mathbb{R}^n)$ for $p_\theta < p < \infty$, where p_θ is an index depending on θ such that $0 < p_\theta < 1$. The operator norm of T only depends on C_1 and C_2.

In order to prove Theorem 10.1, let us prove the following Kolmogoroff's Theorem first.

Theorem 10.2. (1). *If $\mathbf{K}$ is an operator of weak type (p,q), $1 \leq p, q < \infty$, with constant A, then for all $0 < r < q$, $|\mathbf{K}f|^r$ is locally integrable for each $f \in L^p(\mathbb{R}^n)$ and, furthermore, the Kolmogoroff inequality*

$$\left(\int_E |\mathbf{K}f(y)|^r dy\right)^{\frac{1}{r}} \leq M\left(\frac{q}{q-r}\right)^{\frac{1}{r}} |E|^{\frac{1}{r}-\frac{1}{q}} \|f\|_{L^p(\mathbb{R}^n)} \tag{10.2}$$

holds, where $E \subset \mathbb{R}^n$ is any compact subsets in $\mathbb{R}^n$.
(2). *Conversely, if there exist an r for which $0 < r < q$ and a positive constant A_1 such that for all $E \subset \mathbb{R}^n$,*

$$\left(\int_E |\mathbf{K}f(y)|^r dy\right)^{\frac{1}{r}} \leq A_1 \cdot |E|^{\frac{1}{r}-\frac{1}{q}} \|f\|_{L^p(\mathbb{R}^n)} \tag{10.3}$$

holds for all $f \in L^p(\mathbb{R}^n)$, then $\mathbf{K}$ is of weak type (p,q) with constant $A \leq A_1$.

Proof. (1). Let

$$(\mathbf{K}f)_*^E(\alpha) = |\{y \in E : |\mathbf{K}f(y)| > \alpha\}| = |(\mathbf{K}f)_*(\alpha) \cap E|.$$

Thus we have $(\mathbf{K}f)_*^E(\alpha) \leq \nu(E)$ and

$$(\mathbf{K}f)_*^E(\alpha) \leq (\mathbf{T}f)_*(\alpha) \leq \left(\frac{A\|f\|_{L^p}}{\alpha}\right)^q.$$

Then

$$\begin{aligned}\int_E |\mathbf{K}f(y)|^r dy &= r\int_0^\infty \alpha^{r-1}(\mathbf{K}f)_*^E(\alpha)d\alpha = r\int_0^N + r\int_N^\infty \\ &\leq r\int_0^N \alpha^{r-1}|E|d\alpha + r\int_N^\infty \alpha^{r-1}\left(\frac{A\|f\|_{L^p}}{\alpha}\right)^q d\alpha \\ &= |E| \cdot N^r + (A\|f\|_{L^p})^q\left(\frac{r}{q-r}\right) \cdot N^{r-q}.\end{aligned}$$

The last sum is minimized by $N = A \cdot \|f\|_{L^p} |E|^{-\frac{1}{q}}$ and this value of N gives (10.2).
(2). Assume that (10.3) holds. Let $E \subset (\mathbf{K}f)_*(\alpha)$ be any subset with $|E| < \infty$. Then we have

$$\alpha \cdot |E|^{\frac{1}{r}} \leq \left(\int_E |\mathbf{K}f(y)| dy \right)^{\frac{1}{r}} \leq A_1 \cdot |E|^{1-\frac{r}{q}} \|f\|_{L^p(\mathbb{R}^n)},$$

by Chebysheff's inequality. Therefore,

$$|E| \leq \left(\frac{A_1 \|f\|_{L^p}}{\alpha} \right)^q \quad \text{for all} \quad \alpha > 0.$$

That is, $\mathbf{K}$ is of weak type (p, q). □

Proof of Theorem 10.1. The result for $p > 1$ follows from the standard theory of singular integrals (see *e.g.*, Stein's books [19] and [20]). For the fact that T maps $H^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$, see R. Fefferman [10]. It suffices to check the assertion on a (p, ∞)-atom a for some p, since the conditions for this theorem are invariant under translations and dilations. Hence we may assume that a is supported on the unit ball $B(0; 1)$. We know that $|a| \leq C$ and

$$\int_{\mathbb{R}^n} a(x) dx = 0.$$

Let $A(x) = K * a(x)$. It follows that we need to estimate

$$A^*(x) \equiv \mathcal{M}_\varphi(A)(x) = \sup_{\varepsilon > 0} |\varphi_\varepsilon * A(x)|.$$

There are two cases:

Case 1. $|x| \leq 4$. Then we have

$$\begin{aligned}
\int_{|x|\leq 4} |A^*(x)|^p dx &\leq \left(\int_{|x|\leq 4} |A^*(x)|^2 dx \right)^{\frac{p}{2}} \cdot \left(\int_{|x|\leq 4} 1 \cdot dx \right)^{\frac{2-p}{2}} \\
&\leq \left(\int_{|x|\leq 4} |A(x)|^2 dx \right)^{\frac{p}{2}} \cdot \left(\int_{|x|\leq 4} 1 \cdot dx \right)^{\frac{2-p}{2}} \\
&\leq C \cdot C_2^p \cdot \left(\int_{|x|\leq 4} |a(x)|^2 dx \right)^{\frac{p}{2}} \cdot 4 \frac{n(2-p)}{2} \\
&\leq C \cdot C_2^p \cdot 4^{\frac{np}{2}} \cdot 4^{\frac{n(2-p)}{2}} \equiv C.
\end{aligned}$$

The second inequality above follows from the Hardy-Littlewood maximal theorem.

Case 2. $|x| > 4$. Then we have

$$\begin{aligned}
|\varphi_\varepsilon * A(x)| &= \left| \int_{\mathbb{R}^n} \varphi_\varepsilon(x-y) \int_{\mathbb{R}^n} a(z) K(y-z) dz dy \right| \\
&= \left| \int_{\mathbb{R}^n} \varphi_\varepsilon(x-y) \int_{\mathbb{R}^n} a(z) [K(y-z) - K(y)] dz dy \right| \\
&= \left| \int_{\mathbb{R}^n} \varphi_\varepsilon(y) \int_{\mathbb{R}^n} a(z) [K(x-y-z) - K(x-y)] dz dy \right|.
\end{aligned}$$

The second equality holds by the moment condition for a. Now we define

$$J(x-y) = \int_{|z|\leq 1} a(z) [K(x-y-z) - K(x-y)] dz$$

then

$$J(x) = \int_{|z|\le 1} a(z)[K(x-z) - K(x)]dz.$$

Suppose $2^k \le |x| \le 2^{k+1}$, $k \ge 2$, then by the assumption, we have

$$\begin{aligned}\int_{2^k\le|x|\le 2^{k+1}} |J(x)|dx &\le \int_{2^k\le|x|\le 2^{k+1}}\int_{|z|\le 1} |a(z)|\cdot|K(x-z)-K(x)|dzdx\\ &\le C\cdot\int_{2^k\le|x|\le 2^{k+1}} |K(x-z)-K(x)|dx\\ &\le C\cdot C_1\cdot 2^{-k\theta}.\end{aligned}$$

This tells us that $J \in L^1(\mathbb{R}^n)$. Thus we have

$$\sup_{\varepsilon>0}|\varphi_\varepsilon * A(x)| \le \sup_{\varepsilon>0}\frac{1}{\varepsilon^n}\int_{|y|\le\varepsilon}|J(x-y)|dy \;=\; M_{HL}(J)(x).$$

Here $M_{HL}(J)(x)$ is the Hardy-Littlewood maximal function of $J(x)$. If $0<p<1$ and

$$I_k = \{x\in\mathbb{R}^n : 2^k \le |x| \le 2^{k+1}\} \quad \text{for} \quad k\in\mathbb{Z}_+,$$

we have

$$\left[\int_{I_k}(M_{HL}(J)(x))^p dx\right] \le C\cdot|I_k|^{-p+1}\cdot\|J(x)\|^p_{L^1(I_k^*)}.$$

Here $|I_k|$ is the measure of the cube I_k and $I_k^* = \{x\in\mathbb{R}^n : 2^{k-1}\le|x|\le 2^{k+2}\}$ is the "doubling" of I_k. Now by Kolmogoroff's theorem, we have

$$\begin{aligned}\int_{|x|>4}(M_{HL}(J)(x))^p dx &= \sum_{k=2}^\infty\int_{2^k\le|x|\le 2^{k+1}}(M_{HL}(J)(x))^p dx \;\le\; C\cdot\sum_{k=2}^\infty|I_k|^{-p+1}\cdot\|J\|^p_{L^1(I_k^*)}\\ &\le C\cdot\sum_{k=1}^\infty(2^k)^{n(-p+1)}\cdot 2^{-kp\theta} \;\le\; C\cdot\sum_{k=1}^\infty 2^{-k[(n+\theta)p-n]}.\end{aligned}\tag{10.4}$$

If $p > \frac{n}{n+\theta}$, then we have $(n+\theta)p - n > 0$. So the infinite series (10.4) converges and we have proved that the mapping

$$a \;\mapsto\; a * K$$

is bounded on $H^p(\mathbb{R}^n)$ to itself for $p_\theta < p < 1$. Since we know that $a\mapsto a*K$ is bounded on $H^p(\mathbb{R}^n)$ to itself for $1<p<\infty$, then by a interpolation theorem (see Folland and Stein [12]), we know that the mapping is bounded on $H^p(\mathbb{R}^n)$ to itself for $p_\theta < p < \infty$. From the computation above, it is easy to see the operator norm of T depends on C_1 and C_2 only. □

Remarks.

(1). In the proof of Theorem 10.1, we just use the moment condition

$$\int_{\mathbb{R}^n} a(x)dx = 0$$

to obtain our result for $p > \frac{n}{n+\theta}$. For those $p << 1$, we have to use the higher moment condition for $a(x)$ so the assumption (10.1) in Theorem 10.1 is not enough to prove $T : H^p(\mathbb{R}^n)\to H^p(\mathbb{R}^n)$. In fact, the correct condition for K such that T can be extended as a bounded operator on $H^p(\mathbb{R}^n)$ should be

$$\int_{|x|\ge a|y|}\left|K(x-y) - K(x) - \sum_{|\beta|=1}^{n_p}\frac{\partial^\beta K(x)}{\partial x^\beta}y^\beta\right|dx \le C\cdot a^{-\theta}$$

for all $a \geq 2$ and for some $\theta > 0$. Here $n_p \geq \left[n\left(\frac{1}{p}-1\right)\right]$ is the integral part of the number $n\left(\frac{1}{p}-1\right)$.

(2). From Theorem 10.1, we can prove that a singular integral defined by a Calderón-Zygmund kernel

$$T(f)(x) = \lim_{\varepsilon \to 0} c \cdot \int_{|y|>\varepsilon} f(x-y) \frac{\Omega(y')}{|y|^n} dy$$

can be extended as a bounded from $H^p(\mathbb{R}^n)$ to $H^p(\mathbb{R}^n)$ for $0 < p < \infty$.

(3). Theorem 10.1 can be generalized to nilpotent Lie groups. The best example will be the Heisenberg group $\mathbf{H}^n$. Let $|\cdot|_h$ be the "non-isotropic norm" function defined in (7.3) in Section 6.3. Let T be a singular integral operator defined by a kernel K. Suppose further that $T : L^2(\mathbf{H}^2) \to L^2(\mathbf{H}^n)$ and that the kernel K satisfying supp(K) has compact support in the variable x,

$$\int_{|x|_h \geq a|y|_h} |K(x \cdot y) - K(x)| dV(x) < C_1 \cdot a^{-\theta}$$

and

$$\int_{|x|_h \geq a|y|_h} |K(y \cdot x) - K(x)| dV(x) < C_2 \cdot a^{-\theta}$$

for all $a \geq 2$ and some $\theta > 0$. Then the operator T extends to a bounded operator on Hardy space $\mathcal{H}^p(\mathbf{H}^n)$ for $p_\theta < p < \infty$ and the operator norm of T depends on the constants C_1 and C_2 only.

11. Local Hardy Spaces $h^p(\mathbb{R}^n)$

As we have seen before, the real Hardy spaces are designed to behave well under the Calderón-Zygmund operators. In particular, they respect translations, rotations, and dilations. It is known that order zero pseudo-differential operators (which are not translation invariant) are not bounded on the space $H^1(\mathbb{R}^n)$. Closely related to this shortcoming is the fact the Hardy spaces are not closed under compositions with diffeomorphisms nor under multiplication by smooth functions with compact support.

It is with these technical difficulties in mind that in order to deal with pseudo-differential operators of order zero, we need first to introduce the definitions associated with the local Hardy theory developed in Goldberg [13]. We first describe local Hardy spaces $h^p(\mathbb{R}^n)$ in terms of "local" p-atoms. As before, denote $\ell(Q)$ the side length of a cube Q.

Definition 11.1. Let $0 < p \leq 1$. A bounded, measurable function a on $\mathbb{R}^n$ is called a local $(p, 2)$-atom if

- a is supported on a cube $Q \subseteq \mathbb{R}^n$;
- $\|a\|_{L^2(\mathbb{R}^n)} \leq |Q|^{\frac{1}{2}-\frac{1}{p}}$;
- Either $\ell(Q) \leq 1$ and $\int_Q a(x)x^\alpha dx = 0$ for all multi-indices α with $|\alpha| \leq \left[n\left(\frac{1}{p}-1\right)\right]$; or $\ell(Q) > 1$.

Definition 11.2. A distribution f on $\mathbb{R}^n$ is said to be in the local p-Hardy space, written $f \in h^p(\mathbb{R}^n)$ if and only if there are a sequence $\{\lambda_j\} \in \ell^p$ and local $(p, 2)$-atoms a_j such that

$$f = \sum_{j=1}^{n} \lambda_j a_j.$$

The infimum of the norms $\|\{\lambda_j\}\|_{\ell^p}$, taken over all possible atomic decompositions, is comparable to the h^p norm of f.

Similar to $H^p(\mathbb{R}^n)$, we may define $h^p(\mathbb{R}^n)$ by maximal function. Let ϕ be a fixed C_0^∞ function with integral 1 and $\phi_\varepsilon(x) \equiv \varepsilon^{-n}\phi(x/\varepsilon)$, $\varepsilon > 0$, is a standard approximation to the identity. The local maximal function is defined by

$$M_{loc}f(x) \equiv \sup_{0<t<1} |\phi_t * f(x)|.$$

The "locality" enters the picture because we calculate the supremum only over $0 < t < 1$. By a theorem of Goldberg [13], we know that

Theorem 11.3. *Let $0 < p \le 1$. A distribution f is in $h^p(\mathbb{R}^n)$ if and only if the maximal function $M_{loc}f(x)$ lies in $L^p(\mathbb{R}^n)$. Moreover,*

$$\|M_{loc}f\|_{L^p(\mathbb{R}^n)} \approx \|\{\lambda_j\}\|_{\ell^p}.$$

Obviously, one has $H^p(\mathbb{R}^n) \subset h^p(\mathbb{R}^n)$. The local Hardy spaces $h^p(\mathbb{R}^n)$ enjoy many attractive properties: they are preserved by composition with a diffeomorphism which is the identity map for sufficiently large x; also if $f \in h^p(\mathbb{R}^n)$ and $\psi \in C_0^\infty(\mathbb{R}^n)$ then $\psi \cdot f \in h^p(\mathbb{R}^n)$. For details, we refer the reader to Goldberg's paper. One result is that the local Hardy spaces may be defined on a manifold and are acted on in a natural way by pseudo-differential operators.

The last theorem in this lecture notes that we are going to discuss is the h^p regularity property for pseudo-differential operators of order zero. Most of the material in this section is taken from joint works of Chang, Krantz and Stein [4] and [5].

Theorem 11.4. *Let E be a bounded domain with smooth boundary. Suppose K is a distribution that is smooth away from the diagonal $\Sigma = \{(x,y) \in E \times \mathbb{R}^n : x = y\}$ and let $K(x,y) = k(x, x-y)$. For each $x \in E$, we assume that the distribution $k(x,z)$ is a smooth function when $z \neq 0$ which satisfies the following conditions:*

(1)

$$|\widehat{k}(x,\xi)| = \left| \int_{\mathbb{R}^n} e^{-2\pi i z\cdot\xi} k(x,\cdot)dz \right| \le c;$$

(2)

$$\left| \frac{\partial^\alpha k(x,z)}{\partial z^\alpha} \right| \le \frac{c_\alpha}{|z|^{n+|\alpha|}} \qquad (x,z) \in E \times \mathbb{R}^n, \quad z \neq 0$$

for all $|\alpha| \ge 0$. We also assume that $x \mapsto k(x,\cdot)$ is a smooth function and satisfies the following conditions:

(3)

$$\left| \frac{\partial^\beta \widehat{k}(x,\xi)}{\partial x^\beta} \right| \le c_\beta;$$

and

(4)

$$\left| \frac{\partial^{\alpha+\beta} k(x,z)}{\partial z^\alpha \partial x^\beta} \right| \le \frac{c_{\alpha\beta}}{|z|^{n+|\alpha|}}, \qquad z \neq 0$$

for all $|\alpha| \ge 0$, $|\beta| \ge 0$. We define the mapping T by

$$T(f)(x) = \int_{\mathbb{R}^n} k(x, x-z) f(z) dz.$$

We use the terminology "norm" even though these are not norms when $p < 1$.

Then we have

$$\|Tf\|_{h^p(\mathbb{R}^n)} \le C \cdot \|f\|_{h^p(\mathbb{R}^n)} \quad \textit{for} \quad 0 < p < \infty.$$

The constant C depends on $c_{\alpha\beta}$ and E only.

Proof. The result for $p > 1$ follows from the standard theory of singular integrals. When $p \le 1$ and the kernel k are translation invariant, see Theorem 10.1. Here we shall treat in detail only the case of $p \le 1$ and T is not a convolution operator.

To prove the proposition, we need only to check the assertion on a local p-atom $a \in h^p(\mathbb{R}^n)$. Let us fix once and for all a function $\phi \in C_0^\infty(\mathbb{R}^n)$, with $\phi \ge 0$, $\operatorname{supp}(\phi) \subseteq \{x \in \mathbb{R}^n : |x| \le 1\}$, and $\int \phi(x)dx = 1$. From Theorem 11.3, we see that in order to prove the theorem it is necessary to show that

$$\|M_{loc}A\|_{L^p(\mathbb{R}^n)} \equiv \|A^*\|_{L^p(\mathbb{R}^n)} \le C,$$

where $A = T(a)$.

Without loss of generality, we may assume that

$$\operatorname{supp}(a) \subseteq \{y \in \mathbb{R}^n : |y| \le \varepsilon\}.$$

If the atom a is supported in a cube with diameter greater than 1, *i.e.*, $\varepsilon > 1$, then a does not necessarily satisfy a moment condition. Since $\|a\|_{L^2(\mathbb{R}^n)} \le |Q|^{\frac{1}{2}-\frac{1}{p}} \le 1$, it is obvious that $a \in L^2(\mathbb{R}^n)$. Hence by part (3) of our hypotheses, we find that

$$\int_E |A(x)|^2 dx \le \int_{\mathbb{R}^n} |A(x)|^2 dx = \int_{\mathbb{R}^n} |T(a)(x)|^2 dx \le c \cdot \int_{\mathbb{R}^n} |a(x)|^2 dx \le c.$$

It follows that $\|A\|_{L^2(E)} \le \frac{C_E}{|E|^{\frac{1}{2}-\frac{1}{p}}}$ is a "large" local p-atom and hence $A \in h^p(\mathbb{R}^n)$.

If the atom a is supported in a cube with diameter less than or equal to 1, *i.e.*, a is a classical p-atom, then there are two cases:

Case (1). $|x| \le 4\varepsilon$. Then we have

$$\begin{aligned}
\int_{|x|\le 4\varepsilon} |A^*(x)|^p dx &\le \left(\int_{|x|\le 4\varepsilon} |A^*(x)|^2 dx\right)^{\frac{p}{2}} \left(\int_{|x|\le 4\varepsilon} dx\right)^{\frac{2-p}{2}} \\
&\le \left(\int_{|x|\le 4\varepsilon} |M_{HL}(A)(x)|^2 dx\right)^{\frac{p}{2}} \left(\int_{|x|\le 4\varepsilon} dx\right)^{\frac{2-p}{2}} \\
&\le c \left(\int_{\mathbb{R}^n} |A(x)|^2 dx\right)^{\frac{p}{2}} \left(\int_{|x|\le 4\varepsilon} dx\right)^{\frac{2-p}{2}} \\
&\le c^p \left(\int_{|x|\le 4\varepsilon} |a(x)|^2 dx\right)^{\frac{p}{2}} \cdot (4\varepsilon)^{\frac{n(2-p)}{2}} \\
&\le c^p \cdot \varepsilon^{\frac{n(p-2)}{2}} \cdot (4\varepsilon)^{\frac{n(2-p)}{2}} \equiv C.
\end{aligned}$$

Here $M_{HL}(A)$ is the Hardy-Littlewood maximal function of A.

Case (2). $|x| > 4\varepsilon$. Then we have

$$\begin{aligned}|\phi_t * A(x)| &= \left|\int \phi_t(x-y)\int a(u)k(y,y-u)dudy\right| \\ &= \left|\int a(u)\int \phi_t(y)k(x-y,x-y-u)dydu\right| \\ &\equiv \left|\int a(u)k_t(x,x-u)du\right|.\end{aligned}$$

□

In order to finish the proof of this theorem, we must have good control of the kernels $k_t(x,u)$, for $0 < t \leq 1$. We need the following two lemmas:

Lemma 11.5. *Let k be a distribution in $\mathbb{R}^n$ satisfying the following estimates*

(1)

$$\left|\widehat{k}(\xi)\right| = \left|\int_{\mathbb{R}^n} e^{-2\pi i z\cdot\xi}k(\cdot)dz\right| \leq c;$$

and

(2)

$$\left|\frac{\partial^\alpha k(z)}{\partial z^\alpha}\right| \leq \frac{c_\alpha}{|z|^{n+|\alpha|}}, \qquad \text{for all } |\alpha| \geq 0, \;\; z \neq 0.$$

Then the function

$$k_t(z) = \int_{\mathbb{R}^n} \phi_t(u)k(z-u)du$$

satisfies the size estimate

$$\left|\frac{\partial^\alpha k_t(z)}{\partial z^\alpha}\right| \leq \frac{c'_\alpha}{|z|^{n+|\alpha|}}$$

for all $|\alpha| \geq 0$ and uniformly in $0 < t < \infty$.

Proof. By rescaling, we may assume that $t = 1$. It is easy to see, by taking the Fourier transform that, the kernel $k_1(z)$ satisfies

$$|k_1(z)| \leq c \qquad \text{and} \qquad \left|\frac{\partial^\alpha k_1(z)}{\partial z^\alpha}\right| \leq c_\alpha.$$

It remains to show that

$$\left|\frac{\partial^\alpha k_1(z)}{\partial z^\alpha}\right| \leq c_\alpha |z|^{-n-|\alpha|}$$

for $|z|$ large. It suffices to show this for $|z| \geq 2$, *i.e.*, $|z-u| \approx |z|$. By the assumptions on the distribution k we have

$$\begin{aligned}\left|\frac{\partial^\alpha k_1(z)}{\partial z^\alpha}\right| &= \left|\int_{|u|\leq 1} \phi_1(u)\frac{\partial^\alpha k(z)}{\partial z^\alpha}(z-u)du\right| \\ &\leq A\int_{|u|\leq 1} \frac{\phi(u)}{|z-u|^{n+|\alpha|}}du \;\leq\; \frac{A}{|z|^{n+|\alpha|}}.\end{aligned}$$

This completes the proof of the lemma. □

Lemma 11.6. *Let* $k(x,z) \in C_0^\infty(\mathbb{R}^n) \times \mathcal{S}'(\mathbb{R}^n)$ *satisfy the following estimates*

(1)

$$\left|\widehat{k}(x,\xi)\right| = \left|\int_{\mathbb{R}^n} e^{-2\pi i z\cdot\xi} k(x,\cdot)dz\right| \leq c, \quad \textit{uniformly in} \quad x$$

and

(2)

$$\left|\frac{\partial^{\alpha+\beta}k(x,z)}{\partial x^\beta \partial z^\alpha}\right| \leq \frac{A}{|z|^{n+|\alpha|}}, \qquad z \neq 0$$

for all $|\alpha|, |\beta| \geq 0$. *Let* ϕ *be as in the beginning of the proof of Theorem 11.4. Then the function*

$$k_t(x,z) = \int_{\mathbb{R}^n} \phi_t(u)k(x-u, z-u)du$$

satisfies the estimate

$$\left|\frac{\partial^\alpha k_t(x,z)}{\partial z^\alpha}\right| \leq \frac{A}{|z|^{n+|\alpha|}}$$

for all $|\alpha|, |\beta| \geq 0$ *and* $0 < t \leq 1$.

Proof. We may rewrite the kernel $k_t(x,z)$ as follows:

$$\begin{aligned} k_t(z) &= \int_{\mathbb{R}^n} \phi_t(u)k(x-u, z-u)du \\ &= \int_{\mathbb{R}^n} \phi_t(u)k(x-z, z-u)du \\ &\quad + \int_{\mathbb{R}^n} \phi_t(u)[k(x-u, z-u) - k(x-z, z-u)]du. \end{aligned} \tag{11.1}$$

The first term on the right satisfies the correct estimates by Lemma 8.5. By the assumptions on $k(x,z)$ and the mean value theorem, the second term is dominated by

$$\left|\int_{|u|\leq t} \phi_t(u)[k(x-u,z-u) - k(x-z,z-u)]du\right| \leq A \cdot \int_{|u|\leq t} \phi_t(u)|z-u|^{-n+1}du. \tag{11.2}$$

If $|z| \leq 2$ and $|z| \leq 4t$, then this is majorized by

$$ct^{-n}\int_{|u|\leq 5t} |u|^{-n+1}du \leq ct^{-n+1} \leq c_1|z|^{-n+1}.$$

If $|z| \leq 2$ and $|z| > 4t$, then we have $|z-u| \geq |z| - |u| \geq |z| - \frac{1}{4}|z| = \frac{3}{4}|z|$. Hence (11.2) is instead majorized by

$$\int_{|u|\leq t} \phi_t(u)|z-u|^{-n+1}du \leq c_1|z|^{-n+1} \cdot \int_{|u|\leq t} \phi_t(u)du \leq c_1|z|^{-n+1}.$$

It follows that the second term in (8.1) is dominated by $\frac{c_1'}{|z|^n}$ since $|z| \leq 2$. When $|z| \geq 2$, we just need to look at the size estimate

$$\left|\int_{|u|<t} \phi_t(u)k(x-u,z-u)du\right| \leq \frac{c}{|z|^n}.$$

In this case we know that $|u| \leq t \leq 1 \leq \frac{|z|}{2}$. It follows that $|z-u| \approx |z|$.

To prove that the function $k_t(x,z)$ satisfying the estimate $\left|\frac{\partial^\alpha k_t(x,z)}{\partial z^\alpha}\right| \leq \frac{A}{|z|^{n+|\alpha|}}$ is similar. We omit the details. ☐

Conclusion of the proof of Theorem 11.4. We know that

$$|\phi_t * A(x)| = \left| \int a(u)[k_t(x, x-u) - k_t(x, x)]du \right| \tag{11.3}$$

for $0 < t \leq 1$. However, by Lemma 11.6,

$$k_t(x, x-u) = \sum_{|\alpha| \leq M} \frac{(-u)^\alpha}{\alpha!} \frac{\partial^\alpha k_t(x, x)}{\partial z^\alpha} + \mathcal{O}(|u|^{M+1} \cdot |x|^{-n-M-1}). \tag{11.4}$$

Inserting the expression in $\sum_{|\alpha| \leq M}$ into (11.3) gives zero because of the moment condition satisfied by a.

Now the integral that results from substituting the error term of (11.4) into (11.3) is majorized by

$$\frac{c_M}{|x|^{n+M+1}} \cdot \left(\int_{|u| \leq \varepsilon} |a(u)|^2 du \right)^{\frac{1}{2}} \left(\int_{|u| \leq \varepsilon} |u|^{2M+2} du \right)^{\frac{1}{2}} \leq \frac{c_M}{|x|^{n+M+1}} \cdot \varepsilon^{\frac{n}{2} - \frac{n}{p}} \varepsilon^{\frac{n}{2}+M+1}.$$

Then we have

$$\begin{aligned} \int_{|x|>4\varepsilon} |\phi_t * A(x)|^p \, dx &= \int_{|x|>4\varepsilon} \left(\frac{c_M}{|x|^{n+M+1}} \cdot \varepsilon^{\frac{n}{2} - \frac{n}{p}} \varepsilon^{\frac{n}{2}+M+1} \right)^p dx \\ &= C(c_M) \varepsilon^{p(-n-M-1)+n} \cdot \varepsilon^{np-n+Mp+p} = C(c_M) < \infty \end{aligned}$$

This concludes Case (2) and Theorem 11.4 is proved. □

Final Remark. From Theorem 9.8, Theorem 10.1, and Theorem 11.3, we may conclude that the Hilbert transform on $\mathbb{R}$, Riesz transforms on $\mathbb{R}^{n+1}$ and the Szegö projection on the Heisenberg group $\mathbf{H}_n$ originally defined on Schwartz space $\mathcal{S}$ can be extended as bounded operators from H^p to H^p (and h^p to h^p) for $0 < p < \infty$.

Statements and Declarations

The authors declare that they have no conflict of interest, and the manuscript has no associated data.

Acknowledgments

The first author is partially supported by a McDevitt Endowment Fund at Georgetown University. The second author is partially supported by a NNSF grant of China (No. 12071473). The third author is partially supported by a MOST grant of Taiwan (No. 113-2115-M-039-001).

References

[1] L. Ahlfors. Zur Theorie der Überlagerungsfächen. *Acta Mathematica*, 65: 157-194, 1935.

[2] A.P. Calderón. Cauchy integrals on Lipschitz curves and related operators. *Proceedings of the National Academy of Sciences of the United States of America*, 74(4): 1324-1327, 1977.

[3] A.P. Calderón and A. Zygmund. On the existence of certain singular integrals. *Acta Mathematica*, 88: 85-139, 1952.

[4] D.-C. Chang, S. Krantz, and E.M. Stein. Hardy spaces and elliptic boundary value problems. *Contemporary Mathematics*, 137: 119-131, 1992.

[5] D.-C. Chang, S. Krantz, and E.M. Stien. H^p theory on a smooth domain in $\mathbb{R}^n$ and elliptic boundary value problems. *Journal of Functional Analysis*, 114: 286-347, 1993.

[6] R. Coifman and G. Weiss. *Analyse Harmonique Non-Commutative sur Certains Espaces Homogenes*. Springer Lecture Notes, 242: Springer-Verlag, Berlin, 1971.

[7] M. Cotlar. A combinatorial inequality and its application to L^2 spaces. *Revista Matemática Cuyana*, 1: 41-55, 1955.

[8] G. David and J. Journé. A Boundedness Criterion for generalized Calderón-Zygmund operators. *Annals of Mathematics*, 120(2): 371-397, 1984.

[9] G. David and S. Semmes. Regular mappings between dimensions. *Publicacions Matemátiques*, 44: 369-417, 2000.

[10] R. Fefferman. Calderón-Zygmund theory for product domains: H^p spaces. *Proceeding of the National Academy of Science*, USA, 83: 840-843, 1986.

[11] G.B. Folland and E.M. Stein. Estimate for the $\bar{\partial}_b$-complex and analysis on the Heisenberg group. *Communication on Pure and Applied Mathematics*, 27: 429-522, 1974.

[12] G.B. Folland and E.M.Stein. *Hardy Spaces on Homogeneous Groups.* Princeton University Press, Princeton, New Jersey, 1982.

[13] W. Goldberg. A local version of real Hardy spaces. *Duke Mathematical Journal*, 46: 27-42, 1979.

[14] L. Hormander. Hypoelliptic second-order differential equations. *Acta Mathematica*, 119: 147-171, 1967.

[15] Katznelson. *An Introduction to Harmonic Analysis.* John Wiley, New York, 1968.

[16] N. Kerzman and E.M. Stein. The Cauchy kernel, the Szegö kernel, and the Riemann mapping function. *Mathematische Annalen*, 236: 85-94, 1978.

[17] A. Knapp and E.M. Stein. Intertwining operators for semi-simple groups. *Annal of Mathematics*, 93: 489-578, 1971.

[18] C. Sadosky. *Interpolation of operators and singular integrals.* Marcel Dekker, New York-Basel, 1979.

[19] E. M. Stein. *Singular Integrals and Differentiation Properties of Functions.* Princeton University Press, Princeton, New Jersey, 1971.

[20] E. M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals.* (PMS- 43). Princeton University Press, Princeton, New Jersey, 1993.

[21] E.M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, Princeton, New Jersey, 1971.