Chengchun Hao
Institute of Mathematics, AMSS, CAS

LECTURE NOTES ON
Introduction to Harmonic Analysis

调和分析引论·讲义

郝成春

• 2012 •

Contents

Chapter 1
The Fourier Transform and Tempered Distributions

In this chapter, we introduce the Fourier transform and study its more elementary properties, and extend the definition to the space of tempered distributions. We also give some characterizations of operators commuting with translations.

1.1 The L^1 theory of the Fourier transform

We begin by introducing some notation that will be used throughout this work. $\mathbb{R}^n$ denotes n-dimensional real Euclidean space. We consistently write $x = (x_1, x_2, \cdots, x_n)$, $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$, $\cdots$ for the elements of $\mathbb{R}^n$. The inner product of x, $\xi \in \mathbb{R}^n$ is the number $x \cdot \xi = \sum_{j=1}^n x_j \xi_j$, the norm of $x \in \mathbb{R}^n$ is the nonnegative number $|x| = \sqrt{x \cdot x}$. Furthermore, $dx = dx_1 dx_2 \cdots dx_n$ denotes the element of ordinary Lebesgue measure.

We will deal with various spaces of functions defined on $\mathbb{R}^n$. The simplest of these are the $L^p = L^p(\mathbb{R}^n)$ spaces, $1 \leqslant p < \infty$, of all measurable functions f such that $\|f\|_p = \left(\int_{\mathbb{R}^n} |f(x)|^p dx\right)^{1/p} < \infty$. The number $\|f\|_p$ is called the L^p norm of f. The space $L^\infty(\mathbb{R}^n)$ consists of all essentially bounded functions on $\mathbb{R}^n$ and, for $f \in L^\infty(\mathbb{R}^n)$, we let $\|f\|_\infty$ be the essential supremum of $|f(x)|$, $x \in \mathbb{R}^n$. Often, the space $\mathcal{C}_0(\mathbb{R}^n)$ of all continuous functions vanishing at infinity, with the L^∞ norm just described, arises more naturally than $L^\infty = L^\infty(\mathbb{R}^n)$. Unless otherwise specified, all functions are assumed to be complex valued; it will be assumed, throughout the note, that all functions are (Borel) measurable.

In addition to the vector-space operations, $L^1(\mathbb{R}^n)$ is endowed with a "multiplication" making this space a Banach algebra. This operation, called *convolution*, is defined in the following way: If both f and g belong to $L^1(\mathbb{R}^n)$, then their convolution $h = f * g$ is the function whose value at $x \in \mathbb{R}^n$ is

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

One can show by an elementary argument that $f(x-y)g(y)$ is a measurable function of the two variables x and y. It then follows immediately from Fibini's theorem on the interchange of the order of integration that $h \in L^1(\mathbb{R}^n)$ and $\|h\|_1 \leqslant \|f\|_1 \|g\|_1$. Furthermore, this operation is commutative and associative. More generally, we have, with the help of Minkowski's integral inequality $\| \int F(x,y)dy\|_{L^p_x} \leqslant \int \|F(x,y)\|_{L^p_x} dy$, the following result:

Theorem 1.1. *If $f \in L^p(\mathbb{R}^n)$, $p \in [1, \infty]$, and $g \in L^1(\mathbb{R}^n)$ then $h = f * g$ is well defined and belongs to $L^p(\mathbb{R}^n)$. Moreover,*

$$\|h\|_p \leqslant \|f\|_p \|g\|_1.$$

Now, we first consider the Fourier[1] transform of L^1 functions.

Definition 1.2. Let $\omega \in \mathbb{R}\setminus\{0\}$ be a constant. If $f \in L^1(\mathbb{R}^n)$, then its *Fourier transform* is $\mathscr{F} f$ or $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ defined by

$$\mathscr{F} f(\xi) = \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} f(x)dx \tag{1.1}$$

for all $\xi \in \mathbb{R}^n$.

We now continue with some properties of the Fourier transform. Before doing this, we shall introduce some notations. For a measurable function f on $\mathbb{R}^n$, $x \in \mathbb{R}^n$ and $a \neq 0$ we define the *translation* and *dilation* of f by

$$\tau_y f(x) = f(x-y), \tag{1.2}$$

$$\delta_a f(x) = f(ax). \tag{1.3}$$

Proposition 1.3. *Given $f, g \in L^1(\mathbb{R}^n)$, $x, y, \xi \in \mathbb{R}^n$, α multiindex, $a, b \in \mathbb{C}$, $\varepsilon \in \mathbb{R}$ and $\varepsilon \neq 0$, we have*

(i) *Linearity: $\mathscr{F}(af + bg) = a\mathscr{F} f + b\mathscr{F} g$.*
(ii) *Translation: $\mathscr{F}\tau_y f(\xi) = e^{-\omega i y\cdot\xi}\hat{f}(\xi)$.*
(iii) *Modulation: $\mathscr{F}(e^{\omega i x\cdot y} f(x))(\xi) = \tau_y \hat{f}(\xi)$.*
(iv) *Scaling: $\mathscr{F}\delta_\varepsilon f(\xi) = |\varepsilon|^{-n}\delta_{\varepsilon^{-1}}\hat{f}(\xi)$.*
(v) *Differentiation: $\mathscr{F}\partial^\alpha f(\xi) = (\omega i\xi)^\alpha \hat{f}(\xi)$, $\partial^\alpha \hat{f}(\xi) = \mathscr{F}((-\omega i x)^\alpha f(x))(\xi)$.*
(vi) *Convolution: $\mathscr{F}(f * g)(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.*
(vii) *Transformation: $\mathscr{F}(f\circ A)(\xi) = \hat{f}(A\xi)$, where A is an orthogonal matrix and ξ is a column vector.*
(viii) *Conjugation: $\widehat{\overline{f(x)}} = \overline{\hat{f}(-\xi)}$.*

Proof. These results are easy to be verified. We only prove (vii). In fact,

$$\begin{aligned}\mathscr{F}(f\circ A)(\xi) &= \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} f(Ax)dx = \int_{\mathbb{R}^n} e^{-\omega i A^{-1}y\cdot\xi} f(y)dy \\ &= \int_{\mathbb{R}^n} e^{-\omega i A^\top y\cdot\xi} f(y)dy = \int_{\mathbb{R}^n} e^{-\omega i y\cdot A\xi} f(y)dy = \hat{f}(A\xi),\end{aligned}$$

where we used the change of variables $y = Ax$ and the fact that $A^{-1} = A^\top$ and $|\det A| = 1$. □

Corollary 1.4. *The Fourier transform of a radial function is radial.*

Proof. Let $\xi, \eta \in \mathbb{R}^n$ with $|\xi| = |\eta|$. Then there exists some orthogonal matrix A such that $A\xi = \eta$. Since f is radial, we have $f = f\circ A$. Then, it holds

$$\mathscr{F} f(\eta) = \mathscr{F} f(A\xi) = \mathscr{F}(f\circ A)(\xi) = \mathscr{F} f(\xi),$$

by (vi) in Proposition 1.3. □

It is easy to establish the following results:

Theorem 1.5 (Uniform continuity). (i) *The mapping $\mathscr{F}$ is a bounded linear transformation from $L^1(\mathbb{R}^n)$ into $L^\infty(\mathbb{R}^n)$. In fact, $\|\mathscr{F} f\|_\infty \leqslant \|f\|_1$.*
(ii) *If $f \in L^1(\mathbb{R}^n)$, then $\mathscr{F} f$ is uniformly continuous.*

[1] Jean Baptiste Joseph Fourier (21 March 1768 – 16 May 1830) was a French mathematician and physicist best known for initiating the investigation of Fourier series and their applications to problems of heat transfer and vibrations. The Fourier transform and Fourier's Law are also named in his honor. Fourier is also generally credited with the discovery of the greenhouse effect.

Proof. (i) is obvious. We now prove (ii). By

$$\hat{f}(\xi+h)-\hat{f}(\xi)=\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi}[e^{-\omega i x\cdot h}-1]f(x)dx,$$

we have

$$\begin{aligned}|\hat{f}(\xi+h)-\hat{f}(\xi)| \leqslant& \int_{\mathbb{R}^n} |e^{-\omega i x\cdot h}-1||f(x)|dx\\ \leqslant& \int_{|x|\leqslant r} |e^{-\omega i x\cdot h}-1||f(x)|dx+2\int_{|x|>r}|f(x)|dx\\ \leqslant& \int_{|x|\leqslant r} |\omega| r|h||f(x)|dx+2\int_{|x|>r}|f(x)|dx\\ =:& I_1+I_2,\end{aligned}$$

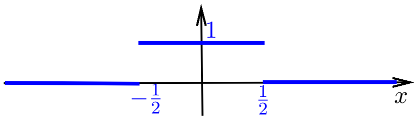
since for any $\theta\geqslant 0$

$$|e^{i\theta}-1|=\sqrt{(\cos\theta-1)^2+\sin^2\theta}=\sqrt{2-2\cos\theta}=2|\sin(\theta/2)|\leqslant|\theta|.$$

Given any $\varepsilon>0$, we can take r so large that $I_2<\varepsilon/2$. Then, we fix this r and take $|h|$ small enough such that $I_1<\varepsilon/2$. In other words, for given $\varepsilon>0$, there exists a sufficiently small $\delta>0$ such that $|\hat{f}(\xi+h)-\hat{f}(\xi)|<\varepsilon$ when $|h|\leqslant\delta$, where ε is independent of ξ. □

Example 1.6. Suppose that a signal consists of a single rectangular pulse of width 1 and height 1. Let's say that it gets turned on at $x=-\frac{1}{2}$ and turned off at $x=\frac{1}{2}$. The standard name for this "normalized" rectangular pulse is

$$\Pi(x)\equiv\operatorname{rect}(x):=\begin{cases}1, & \text{if } -\frac{1}{2}<x<\frac{1}{2},\\ 0, & \text{otherwise.}\end{cases}$$



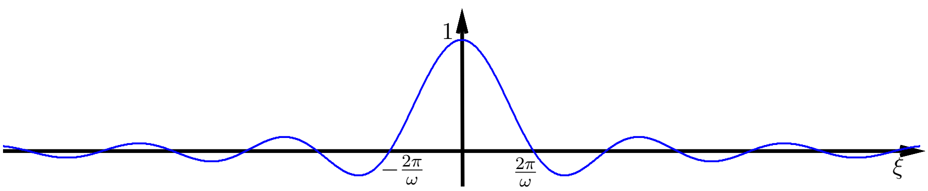
It is also called, variously, the normalized boxcar function, the top hat function, the indicator function, or the characteristic function for the interval $(-1/2,1/2)$. The Fourier transform of this signal is

$$\widehat{\Pi}(\xi)=\int_{\mathbb{R}} e^{-\omega i x\xi}\Pi(x)dx=\int_{-1/2}^{1/2} e^{-\omega i x\xi}dx=\left.\frac{e^{-\omega i x\xi}}{-\omega i\xi}\right|_{-1/2}^{1/2}=\frac{2}{\omega\xi}\sin\frac{\omega\xi}{2}$$

when $\xi\neq 0$. When $\xi=0$, $\widehat{\Pi}(0)=\int_{-1/2}^{1/2}dx=1$. By l'Hôpital's rule,

$$\lim_{\xi\to 0}\widehat{\Pi}(\xi)=\lim_{\xi\to 0}2\frac{\sin\frac{\omega\xi}{2}}{\omega\xi}=\lim_{\xi\to 0}2\frac{\frac{\omega}{2}\cos\frac{\omega\xi}{2}}{\omega}=1=\widehat{\Pi}(0),$$

so $\widehat{\Pi}(\xi)$ is continuous at $\xi=0$. There is a standard function called "sinc"[2] that is defined by $\operatorname{sinc}(\xi)=\frac{\sin\xi}{\xi}$. In this notation $\widehat{\Pi}(\xi)=\operatorname{sinc}\frac{\omega\xi}{2}$. Here is the graph of $\widehat{\Pi}(\xi)$.



Remark 1.7. The above definition of the Fourier transform in (1.1) extends immediately to finite Borel measures: if μ is such a measure on $\mathbb{R}^n$, we define $\mathscr{F}\mu$ by letting

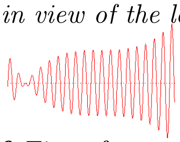
$$\mathscr{F}\mu(\xi)=\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi}d\mu(x).$$

[2] The term "sinc" (English pronunciation:['sɪŋk]) is a contraction, first introduced by Phillip M. Woodward in 1953, of the function's full Latin name, the sinus cardinalis (cardinal sine).

Theorem 1.5 is valid for these Fourier transform if we replace the L^1 norm by the total variation of μ.

The following theorem plays a central role in Fourier Analysis. It takes its name from the fact that it holds even for functions that are integrable according to the definition of Lebesgue. We prove it for functions that are absolutely integrable in the Riemann sense.[3]

Theorem 1.8 (Riemann-Lebesgue lemma). *If $f \in L^1(\mathbb{R}^n)$ then $\mathscr{F}f \to 0$ as $|\xi| \to \infty$; thus, in view of the last result, we can conclude that $\mathscr{F}f \in \mathcal{C}_0(\mathbb{R}^n)$.*



The Riemann-Lebesgue lemma states that the integral of a function like the left is small. The integral will approach zero as the number of oscillations increases.

Proof. First, for $n = 1$, suppose that $f(x) = \chi_{(a,b)}(x)$, the characteristic function of an interval. Then

$$\hat{f}(\xi) = \int_a^b e^{-\omega i x \xi} dx = \frac{e^{-\omega i b \xi} - e^{-\omega i b \xi}}{\omega i \xi} \to 0, \quad \text{as } |\xi| \to 0.$$

Similarly, the result holds when f is the characteristic function of the n-dimensional rectangle $I = \{x \in \mathbb{R}^n : a_1 \leqslant x_1 \leqslant b_1, \cdots, a_n \leqslant x_n \leqslant b_n\}$ since we can calculate $\mathscr{F}f$ explicitly as an iterated integral. The same is therefore true for a finite linear combination of such characteristic functions (i.e., simple functions). Since all such simple functions are dense in L^1, the result for a general $f \in L^1(\mathbb{R}^n)$ follows easily by approximating f in the L^1 norm by such a simple function g, then $f = g + (f - g)$, where $\mathscr{F}f - \mathscr{F}g$ is uniformly small by Theorem 1.5, while $\mathscr{F}g(\xi) \to 0$ as $|\xi| \to \infty$. $\square$

Theorem 1.8 gives a necessary condition for a function to be a Fourier transform. However, that belonging to $\mathcal{C}_0$ is not a sufficient condition for being the Fourier transform of an integrable function. See the following example.

Example 1.9. Suppose, for simplicity, that $n = 1$. Let

$$g(\xi) = \begin{cases} \dfrac{1}{\ln \xi}, & \xi > e, \\ \dfrac{\xi}{e}, & 0 \leqslant \xi \leqslant e, \end{cases}$$
$$g(\xi) = -\,g(-\xi), \quad \xi < 0.$$

It is clear that $g(\xi)$ is uniformly continuous on $\mathbb{R}$ and $g(\xi) \to 0$ as $|\xi| \to \infty$.

Assume that there exists an $f \in L^1(\mathbb{R})$ such that $\hat{f}(\xi) = g(\xi)$, i.e.,

$$g(\xi) = \int_{-\infty}^{\infty} e^{-\omega i x \xi} f(x) dx.$$

Since $g(\xi)$ is an odd function, we have

$$g(\xi) = \int_{-\infty}^{\infty} e^{\omega i x \xi} f(x) dx = i \int_{-\infty}^{\infty} \sin(\omega x \xi) f(x) dx = \int_0^{\infty} \sin(\omega x \xi) F(x) dx,$$

where $F(x) = i[f(x) - f(-x)] \in L^1(\mathbb{R})$. Integrating $\frac{g(\xi)}{\xi}$ over (e, N) yields

$$\int_e^N \frac{g(\xi)}{\xi} d\xi = \int_0^{\infty} F(x) \left(\int_e^N \frac{\sin(\omega x \xi)}{\xi} d\xi \right) dx.$$

[3] Let us very briefly recall what this means. A bounded function f on a finite interval $[a, b]$ is integrable if it can be approximated by Riemann sums from above and below in such a way that the difference of the integrals of these sums can be made as small as we wish. This definition is then extended to unbounded functions and infinite intervals by taking limits; these cases are often called improper integrals. If I is any interval and f is a function on I such that the (possibly improper) integral $\int_I |f(x)| dx$ has a finite value, then f is said to be absolutely integrable on I.

Noticing that

$$\left|\int_a^b \frac{\sin t}{t}dt\right| \leqslant C, \qquad \lim_{N\to\infty}\int_0^N \frac{\sin t}{t}dt = \frac{\pi}{2},$$

and by Lebesgue dominated convergence theorem,we get that the integral of r.h.s. is convergent as $N \to \infty$. However,

$$\lim_{N\to\infty}\int_e^N \frac{g(\xi)}{\xi}d\xi = \lim_{N\to\infty}\int_e^N \frac{d\xi}{\xi\ln\xi} = \infty.$$

This contradiction indicates that the assumption was invalid.

We now turn to the problem of inverting the Fourier transform. That is, we shall consider the question: *Given the Fourier transform $\hat{f}$ of an integrable function f, how do we obtain f back again from $\hat{f}$?* The reader, who is familiar with the elementary theory of Fourier series and integrals, would expect $f(x)$ to be equal to the integral

$$C\int_{\mathbb{R}^n} e^{\omega i x\cdot\xi}\hat{f}(\xi)d\xi. \tag{1.4}$$

Unfortunately, $\hat{f}$ need not be integrable (for example, let $n = 1$ and f be the characteristic function of a finite interval). In order to get around this difficulty, we shall use certain summability methods for integrals. We first introduce the *Abel method of summability*, whose analog for series is very well-known. For each $\varepsilon > 0$, we define the Abel mean $A_\varepsilon = A_\varepsilon(f)$ to be the integral

$$A_\varepsilon(f) = A_\varepsilon = \int_{\mathbb{R}^n} e^{-\varepsilon|x|}f(x)dx. \tag{1.5}$$

It is clear that if $f \in L^1(\mathbb{R}^n)$ then $\lim\limits_{\varepsilon\to 0} A_\varepsilon(f) = \int_{\mathbb{R}^n} f(x)dx$. On the other hand, these Abel means are well-defined even when f is not integrable (e.g., if we only assume that f is bounded, then $A_\varepsilon(f)$ is defined for all $\varepsilon > 0$). Moreover, their limit

$$\lim_{\varepsilon\to 0} A_\varepsilon(f) = \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} e^{-\varepsilon|x|}f(x)dx \tag{1.6}$$

may exist even when f is not integrable. A classical example of such a case is obtained by letting $f(x) = \mathrm{sinc}(x)$ when $n = 1$. Whenever the limit in (1.6) exists and is finite we say that $\int_{\mathbb{R}^n} f dx$ is Abel summable to this limit.

A somewhat similar method of summability is *Gauss summability*. This method is defined by the Gauss (sometimes called Gauss-Weierstrass) means

$$G_\varepsilon(f) = \int_{\mathbb{R}^n} e^{-\varepsilon|x|^2}f(x)dx. \tag{1.7}$$

We say that $\int_{\mathbb{R}^n} f dx$ is Gauss summable (to l) if

$$\lim_{\varepsilon\to 0} G_\varepsilon(f) = \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} e^{-\varepsilon|x|^2}f(x)dx \tag{1.6'}$$

exists and equals the number l.

We see that both (1.6) and (1.6') can be put in the form

$$M_{\varepsilon,\Phi}(f) = M_\varepsilon(f) = \int_{\mathbb{R}^n} \Phi(\varepsilon x)f(x)dx, \tag{1.8}$$

where $\Phi \in \mathcal{C}_0$ and $\Phi(0) = 1$. Then $\int_{\mathbb{R}^n} f(x)dx$ is summable to ℓ if $\lim_{\varepsilon\to 0} M_\varepsilon(f) = \ell$. We shall call $M_\varepsilon(f)$ the Φ means of this integral.

We shall need the Fourier transforms of the functions $e^{-\varepsilon|x|^2}$ and $e^{-\varepsilon|x|}$. The first one is easy to calculate.

Theorem 1.10. *For all $a > 0$, we have*

$$\mathscr{F} e^{-a|\omega x|^2}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{-n} (4\pi a)^{-n/2} e^{-\frac{|\xi|^2}{4a}}. \tag{1.9}$$

Proof. The integral in question is

$$\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} e^{-a|\omega x|^2} dx.$$

Notice that this factors as a product of one variable integrals. Thus it sufficient to prove the case $n = 1$. For this we use the formula for the integral of a Gaussian: $\int_{\mathbb{R}} e^{-\pi x^2} dx = 1$. It follows that

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-\omega i x \xi} e^{-a\omega^2 x^2} dx = &\int_{-\infty}^{\infty} e^{-a(\omega x + i\xi/(2a))^2} e^{-\frac{\xi^2}{4a}} dx \\ =&|\omega|^{-1} e^{-\frac{\xi^2}{4a}} \int_{-\infty+i\xi/(2a)}^{\infty+i\xi/(2a)} e^{-ax^2} dx \\ =&|\omega|^{-1} e^{-\frac{\xi^2}{4a}} \sqrt{\pi/a} \int_{-\infty}^{\infty} e^{-\pi y^2} dy \\ =&\left(\frac{|\omega|}{2\pi}\right)^{-1} (4\pi a)^{-1/2} e^{-\frac{\xi^2}{4a}},\end{aligned}$$

where we used contour integration at the next to last one. □

The second one is somewhat harder to obtain:

Theorem 1.11. *For all $a > 0$, we have*

$$\mathscr{F}(e^{-a|\omega x|}) = \left(\frac{|\omega|}{2\pi}\right)^{-n} \frac{c_n a}{(a^2 + |\xi|^2)^{(n+1)/2}}, \quad c_n = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}. \tag{1.10}$$

Proof. By a change of variables, i.e.,

$$\mathscr{F}(e^{-a|\omega x|}) = \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} e^{-a|\omega x|} dx = (a|\omega|)^{-n} \int_{\mathbb{R}^n} e^{-ix\cdot\xi/a} e^{-|x|} dx,$$

we see that it suffices to show this result when $a = 1$. In order to show this, we need to express the decaying exponential as a superposition of Gaussians, i.e.,

$$e^{-\gamma} = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-\eta}}{\sqrt{\eta}} e^{-\gamma^2/4\eta} d\eta, \quad \gamma > 0. \tag{1.11}$$

Then, using (1.9) to establish the third equality,

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-ix\cdot t} e^{-|x|} dx = &\int_{\mathbb{R}^n} e^{-ix\cdot t} \left(\frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-\eta}}{\sqrt{\eta}} e^{-|x|^2/4\eta} d\eta\right) dx \\ =&\frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-\eta}}{\sqrt{\eta}} \left(\int_{\mathbb{R}^n} e^{-ix\cdot t} e^{-|x|^2/4\eta} dx\right) d\eta \\ =&\frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-\eta}}{\sqrt{\eta}} \left((4\pi\eta)^{n/2} e^{-\eta|t|^2}\right) d\eta \\ =&2^n \pi^{(n-1)/2} \int_0^{\infty} e^{-\eta(1+|t|^2)} \eta^{\frac{n-1}{2}} d\eta \\ =&2^n \pi^{(n-1)/2} \left(1 + |t|^2\right)^{-\frac{n+1}{2}} \int_0^{\infty} e^{-\zeta} \zeta^{\frac{n+1}{2}-1} d\zeta \\ =&2^n \pi^{(n-1)/2} \Gamma\left(\frac{n+1}{2}\right) \frac{1}{(1 + |t|^2)^{(n+1)/2}}.\end{aligned}$$

Thus,

$$\mathscr{F}(e^{-a|\omega x|}) = \frac{(a|\omega|)^{-n}(2\pi)^n c_n}{(1+|\xi/a|^2)^{(n+1)/2}} = \left(\frac{|\omega|}{2\pi}\right)^{-n} \frac{c_n a}{(a^2+|\xi|^2)^{(n+1)/2}}.$$

Consequently, the theorem will be established once we show (1.11). In fact, by changes of variables, we have

$$\begin{aligned}
&\frac{1}{\sqrt{\pi}} e^{\gamma} \int_0^\infty \frac{e^{-\eta}}{\sqrt{\eta}} e^{-\gamma^2/4\eta} d\eta \\
=&\frac{2\sqrt{\gamma}}{\sqrt{\pi}} \int_0^\infty e^{-\gamma(\sigma-\frac{1}{2\sigma})^2} d\sigma \qquad (\text{by } \eta=\gamma\sigma^2) \\
=&\frac{2\sqrt{\gamma}}{\sqrt{\pi}} \int_0^\infty e^{-\gamma(\sigma-\frac{1}{2\sigma})^2} \frac{1}{2\sigma^2} d\sigma \qquad (\text{by } \sigma \mapsto \frac{1}{2\sigma}) \\
=&\frac{\sqrt{\gamma}}{\sqrt{\pi}} \int_0^\infty e^{-\gamma(\sigma-\frac{1}{2\sigma})^2} \left(1+\frac{1}{2\sigma^2}\right) d\sigma \qquad (\text{by averaging the last two formula}) \\
=&\frac{\sqrt{\gamma}}{\sqrt{\pi}} \int_{-\infty}^\infty e^{-\gamma u^2} du \qquad (\text{by } u=\sigma-\frac{1}{2\sigma}) \\
=&1, \qquad (\text{by } \int_{\mathbb{R}} e^{-\pi x^2} dx = 1)
\end{aligned}$$

which yields the desired identity (1.11). □

We shall denote the Fourier transform of $\left(\frac{|\omega|}{2\pi}\right)^n e^{-a|\omega x|^2}$ and $\left(\frac{|\omega|}{2\pi}\right)^n e^{-a|\omega x|}$, $a>0$, by W and P, respectively. That is,

$$W(\xi,a) = (4\pi a)^{-n/2} e^{-\frac{|\xi|^2}{4a}}, \quad P(\xi,a) = \frac{c_n a}{(a^2+|\xi|^2)^{(n+1)/2}}. \tag{1.12}$$

The first of these two functions is called the *Weierstrass* (or *Gauss-Weierstrass*) *kernel* while the second is called the *Poisson kernel.*

Theorem 1.12 (The multiplication formula)**.** *If $f,g\in L^1(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n} \hat{f}(\xi) g(\xi) d\xi = \int_{\mathbb{R}^n} f(x)\hat{g}(x) dx.$$

Proof. Using Fubini's theorem to interchange the order of the integration on $\mathbb{R}^{2n}$, we obtain the identity. □

Theorem 1.13. *If f and Φ belong to $L^1(\mathbb{R}^n)$, $\varphi=\hat{\Phi}$ and $\varphi_\varepsilon(x)=\varepsilon^{-n}\varphi(x/\varepsilon)$, then*

$$\int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} \Phi(\varepsilon\xi) \hat{f}(\xi) d\xi = \int_{\mathbb{R}^n} \varphi_\varepsilon(y-x) f(y) dy$$

for all $\varepsilon>0$. In particular,

$$\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} e^{-\varepsilon|\omega\xi|} \hat{f}(\xi) d\xi = \int_{\mathbb{R}^n} P(y-x,\varepsilon) f(y) dy,$$

and

$$\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} e^{-\varepsilon|\omega\xi|^2} \hat{f}(\xi) d\xi = \int_{\mathbb{R}^n} W(y-x,\varepsilon) f(y) dy.$$

Proof. From (iii) and (iv) in Proposition 1.3, it implies $(\mathscr{F} e^{\omega i x\cdot\xi}\Phi(\varepsilon\xi))(y) = \varphi_\varepsilon(y-x)$. The first result holds immediately with the help of Theorem 1.12. The last two follow from (1.9), (1.10) and (1.12). □

Lemma 1.14. (i) $\int_{\mathbb{R}^n} W(x,\varepsilon)dx = 1$ *for all* $\varepsilon>0$.

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

Proof. By a change of variable, we first note that

$$\int_{\mathbb{R}^n} W(x,\varepsilon)dx = \int_{\mathbb{R}^n} (4\pi\varepsilon)^{-n/2} e^{-\frac{|x|^2}{4\varepsilon}} dx = \int_{\mathbb{R}^n} W(x,1)dx,$$

and

$$\int_{\mathbb{R}^n} P(x,\varepsilon)dx = \int_{\mathbb{R}^n} \frac{c_n\varepsilon}{(\varepsilon^2+|x|^2)^{(n+1)/2}}dx = \int_{\mathbb{R}^n} P(x,1)dx.$$

Thus, it suffices to prove the lemma when $\varepsilon = 1$. For the first one, we use a change of variables and the formula for the integral of a Gaussian: $\int_{\mathbb{R}} e^{-\pi x^2}dx = 1$ to get

$$\int_{\mathbb{R}^n} W(x,1)dx = \int_{\mathbb{R}^n} (4\pi)^{-n/2} e^{-\frac{|x|^2}{4}} dx = \int_{\mathbb{R}^n} (4\pi)^{-n/2} e^{-\pi|y|^2} 2^n \pi^{n/2} dy = 1.$$

For the second one, we have

$$\int_{\mathbb{R}^n} P(x,1)dx = c_n \int_{\mathbb{R}^n} \frac{1}{(1+|x|^2)^{(n+1)/2}}dx.$$

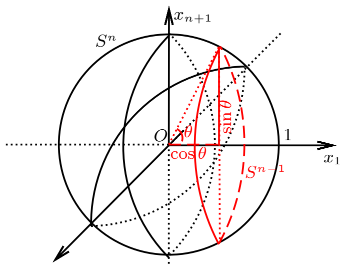
Letting $r = |x|$, $x' = x/r$ (when $x \neq 0$), $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$, dx' the element of surface area on S^{n-1} whose surface area[4] is denoted by ω_{n-1} and, finally, putting $r = \tan\theta$, we have

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{1}{(1+|x|^2)^{(n+1)/2}}dx =& \int_0^\infty \int_{S^{n-1}} \frac{1}{(1+r^2)^{(n+1)/2}} dx' r^{n-1} dr \\ =&\omega_{n-1}\int_0^\infty \frac{r^{n-1}}{(1+r^2)^{(n+1)/2}}dr = \omega_{n-1}\int_0^{\pi/2} \sin^{n-1}\theta d\theta.\end{aligned}$$

But $\omega_{n-1}\sin^{n-1}\theta$ is clearly the surface area of the sphere of radius $\sin\theta$ obtained by intersecting S^n with the hyperplane $x_1 = \cos\theta$. Thus, the area of the upper half of S^n is obtained by summing these $(n-1)$ dimensional areas as θ ranges from 0 to $\pi/2$, that is,

$$\omega_{n-1}\int_0^{\pi/2} \sin^{n-1}\theta d\theta = \frac{\omega_n}{2},$$

which is the desired result by noting that $1/c_n = \omega_n/2$. □



Theorem 1.15. *Suppose $\varphi \in L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n}\varphi(x)dx = 1$ and let $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$ for $\varepsilon > 0$. If $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, or $f \in \mathcal{C}_0(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$, then for $1 \leqslant p \leqslant \infty$*

$$\|f * \varphi_\varepsilon - f\|_p \to 0, \ \text{as } \varepsilon \to 0.$$

In particular, the Poisson integral of f*:*

$$u(x,\varepsilon) = \int_{\mathbb{R}^n} P(x-y,\varepsilon)f(y)dy$$

and the Gauss-Weierstrass integral of f*:*

$$s(x,\varepsilon) = \int_{\mathbb{R}^n} W(x-y,\varepsilon)f(y)dy$$

converge to f in the L^p norm as $\varepsilon \to 0$.

Proof. By a change of variables, we have

$$\int_{\mathbb{R}^n} \varphi_\varepsilon(y)dy = \int_{\mathbb{R}^n} \varepsilon^{-n}\varphi(y/\varepsilon)dy = \int_{\mathbb{R}^n} \varphi(y)dy = 1.$$

Hence,

$$(f * \varphi_\varepsilon)(x) - f(x) = \int_{\mathbb{R}^n} [f(x-y) - f(x)]\varphi_\varepsilon(y)dy.$$

[4] $\omega_{n-1} = 2\pi^{n/2}/\Gamma(n/2)$.

Therefore, by Minkowski's inequality for integrals and a change of variables, we get

$$\begin{aligned}\|f*\varphi_\varepsilon - f\|_p \leqslant & \int_{\mathbb{R}^n} \|f(x-y)-f(x)\|_p \varepsilon^{-n}|\varphi(y/\varepsilon)|dy \\ = & \int_{\mathbb{R}^n} \|f(x-\varepsilon y)-f(x)\|_p|\varphi(y)|dy.\end{aligned}$$

We point out that if $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, and denote $\|f(x-t)-f(x)\|_p = \Delta_f(t)$, then $\Delta_f(t) \to 0$, as $t \to 0$.[5] In fact, if $f_1 \in \mathscr{D}(\mathbb{R}^n) := C_0^\infty(\mathbb{R}^n)$ of all C^∞ functions with compact support, the assertion in that case is an immediate consequence of the uniform convergence $f_1(x-t) \to f_1(x)$, as $t \to 0$. In general, for any $\sigma > 0$, we can write $f = f_1 + f_2$, such that f_1 is as described and $\|f_2\|_p \leqslant \sigma$, since $\mathscr{D}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$. Then, $\Delta_f(t) \leqslant \Delta_{f_1}(t) + \Delta_{f_2}(t)$, with $\Delta_{f_1}(t) \to 0$ as $t \to 0$, and $\Delta_{f_2}(t) \leqslant 2\sigma$. This shows that $\Delta_f(t) \to 0$ as $t \to 0$ for general $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$.

For the case $p = \infty$ and $f \in \mathcal{C}_0(\mathbb{R}^n)$, the same argument gives us the result since $\mathscr{D}(\mathbb{R}^n)$ is dense in $C_0(\mathbb{R}^n)$ (cf. [Rud87, p.70, Proof of Theorem 3.17]).

Thus, by the Lebesgue dominated convergence theorem (due to $\varphi \in L^1$ and the fact $\Delta_f(\varepsilon y)|\varphi(y)| \leqslant 2\|f\|_p|\varphi(y)|$) and the fact $\Delta_f(\varepsilon y) \to 0$ as $\varepsilon \to 0$, we have

$$\lim_{\varepsilon\to 0}\|f*\varphi_\varepsilon - f\|_p \leqslant \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\Delta_f(\varepsilon y)|\varphi(y)|dy = \int_{\mathbb{R}^n}\lim_{\varepsilon\to 0}\Delta_f(\varepsilon y)|\varphi(y)|dy = 0.$$

This completes the proof. □

With the same argument, we have

Corollary 1.16. *Let $1 \leqslant p \leqslant \infty$. Suppose $\varphi \in L^1(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n}\varphi(x)dx = 0$, then $\|f * \varphi_\varepsilon\|_p \to 0$ as $\varepsilon \to 0$ whenever $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, or $f \in \mathcal{C}_0(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$.*

Proof. Once we observe that

$$\begin{aligned}(f*\varphi_\varepsilon)(x) =&(f*\varphi_\varepsilon)(x) - f(x)\cdot 0 = (f*\varphi_\varepsilon)(x) - f(x)\int_{\mathbb{R}^n}\varphi_\varepsilon(y)dy \\ =&\int_{\mathbb{R}^n}[f(x-y)-f(x)]\varphi_\varepsilon(y)dy,\end{aligned}$$

the rest of the argument is precisely that used in the last proof. □

In particular, we also have

Corollary 1.17. *Suppose $\varphi \in L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n}\varphi(x)dx = 1$ and let $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$ for $\varepsilon > 0$. Let $f(x) \in L^\infty(\mathbb{R}^n)$ be continuous at $\{0\}$. Then,*

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} f(x)\varphi_\varepsilon(x)dx = f(0).$$

Proof. Since $\int_{\mathbb{R}^n} f(x)\varphi_\varepsilon(x)dx - f(0) = \int_{\mathbb{R}^n}(f(x)-f(0))\varphi_\varepsilon(x)dx$, then we may assume without loss of generality that $f(0) = 0$. Since f is continuous at $\{0\}$, then for any $\eta > 0$, there exists a $\delta > 0$ such that

$$|f(x)| < \frac{\eta}{\|\varphi\|_1},$$

whenever $|x| < \delta$. Noticing that $|\int_{\mathbb{R}^n}\varphi(x)dx| \leqslant \|\varphi\|_1$, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n} f(x)\varphi_\varepsilon(x)dx\right| \leqslant & \frac{\eta}{\|\varphi\|_1}\int_{|x|<\delta}|\varphi_\varepsilon(x)|dx + \|f\|_\infty\int_{|x|\geqslant\delta}|\varphi_\varepsilon(x)|dx \\ \leqslant & \frac{\eta}{\|\varphi\|_1}\|\varphi\|_1 + \|f\|_\infty\int_{|y|\geqslant\delta/\varepsilon}|\varphi(y)|dy \\ =&\eta + \|f\|_\infty I_\varepsilon.\end{aligned}$$

[5] This statement is the continuity of the mapping $t \to f(x-t)$ of $\mathbb{R}^n$ to $L^p(\mathbb{R}^n)$.

But $I_\varepsilon \to 0$ as $\varepsilon \to 0$. This proves the result. □

From Theorems 1.13 and 1.15, we obtain the following solution to the Fourier inversion problem:

Theorem 1.18. *If both Φ and its Fourier transform $\varphi = \hat{\Phi}$ are integrable and $\int_{\mathbb{R}^n} \varphi(x)dx = 1$, then the Φ means of the integral $(|\omega|/2\pi)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} \hat{f}(\xi)d\xi$ converges to $f(x)$ in the L^1 norm. In particular, the Abel and Gauss means of this integral converge to $f(x)$ in the L^1 norm.*

We have singled out the Gauss-Weierstrass and the Abel methods of summability. The former is probably the simplest and is connected with the solution of the heat equation; the latter is intimately connected with harmonic functions and provides us with very powerful tools in Fourier analysis.

Since $s(x,\varepsilon) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} e^{-\varepsilon|\omega\xi|^2} \hat{f}(\xi)d\xi$ converges in L^1 to $f(x)$ as $\varepsilon > 0$ tends to 0, we can find a sequence $\varepsilon_k \to 0$ such that $s(x,\varepsilon_k) \to f(x)$ for a.e. x. If we further assume that $\hat{f} \in L^1(\mathbb{R}^n)$, the Lebesgue dominated convergence theorem gives us the following pointwise equality:

Theorem 1.19 (Fourier inversion theorem)**.** *If both f and $\hat{f}$ are integrable, then*

$$f(x) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} \hat{f}(\xi)d\xi,$$

for almost every x.

Remark 1.20. We know from Theorem 1.5 that $\hat{f}$ is continuous. If $\hat{f}$ is integrable, the integral $\int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} \hat{f}(\xi)d\xi$ also defines a continuous function (in fact, it equals $\hat{\hat{f}}(-x)$). Thus, by changing f on a set of measure 0, we can obtain equality in Theorem 1.19 for all x.

It is clear from Theorem 1.18 that if $\hat{f}(\xi) = 0$ for all ξ then $f(x) = 0$ for almost every x. Applying this to $f = f_1 - f_2$, we obtain the following uniqueness result for the Fourier transform:

Corollary 1.21 (Uniqueness)**.** *If f_1 and f_2 belong to $L^1(\mathbb{R}^n)$ and $\hat{f}_1(\xi) = \hat{f}_2(\xi)$ for $\xi \in \mathbb{R}^n$, then $f_1(x) = f_2(x)$ for almost every $x \in \mathbb{R}^n$.*

We will denote the inverse operation to the Fourier transform by $\mathscr{F}^{-1}$ or $\check{\ }$. If $f \in L^1$, then we have

$$\check{f}(x) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} f(\xi)d\xi. \tag{1.13}$$

We give a very useful result.

Theorem 1.22. *Suppose $f \in L^1(\mathbb{R}^n)$ and $\hat{f} \geqslant 0$. If f is continuous at 0, then*

$$f(0) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \hat{f}(\xi)d\xi.$$

Moreover, we have $\hat{f} \in L^1(\mathbb{R}^n)$ and

$$f(x) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} \hat{f}(\xi)d\xi,$$

for almost every x.

Proof. By Theorem 1.13, we have

$$\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{-\varepsilon|\omega\xi|} \hat{f}(\xi)d\xi = \int_{\mathbb{R}^n} P(y,\varepsilon)f(y)dy.$$

From Lemma 1.14, we get, for any $\delta > 0$,

$$\begin{aligned}&\left|\int_{\mathbb{R}^n} P(y,\varepsilon)f(y)dy - f(0)\right| = \left|\int_{\mathbb{R}^n} P(y,\varepsilon)[f(y)-f(0)]dy\right|\\ \leqslant&\left|\int_{|y|<\delta} P(y,\varepsilon)[f(y)-f(0)]dy\right| + \left|\int_{|y|\geqslant\delta} P(y,\varepsilon)[f(y)-f(0)]dy\right|\\ =&I_1 + I_2.\end{aligned}$$

Since f is continuous at 0, for any given $\sigma > 0$, we can choose δ small enough such that $|f(y) - f(0)| \leqslant \sigma$ when $|y| < \delta$. Thus, $I_1 \leqslant \sigma$ by Lemma 1.14. For the second term, we have, by a change of variables, that

$$\begin{aligned}I_2 \leqslant&\|f\|_1 \sup_{|y|\geqslant\delta} P(y,\varepsilon) + |f(0)| \int_{|y|\geqslant\delta} P(y,\varepsilon)dy\\ =&\|f\|_1 \frac{c_n\varepsilon}{(\varepsilon^2+\delta^2)^{(n+1)/2}} + |f(0)|\int_{|y|\geqslant\delta/\varepsilon} P(y,1)dy \to 0,\end{aligned}$$

as $\varepsilon \to 0$. Thus, $\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{-\varepsilon|\omega\xi|}\hat{f}(\xi)d\xi \to f(0)$ as $\varepsilon \to 0$. On the other hand, by Lebesgue dominated convergence theorem, we obtain

$$\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \hat{f}(\xi)d\xi = \left(\frac{|\omega|}{2\pi}\right)^n \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} e^{-\varepsilon|\omega\xi|}\hat{f}(\xi)d\xi = f(0),$$

which implies $\hat{f} \in L^1(\mathbb{R}^n)$ due to $\hat{f} \geqslant 0$. Therefore, from Theorem 1.19, it follows the desired result. □

An immediate consequence is

Corollary 1.23. i) $\int_{\mathbb{R}^n} e^{\omega ix\cdot\xi} W(\xi,\varepsilon)d\xi = e^{-\varepsilon|\omega x|^2}$.
ii) $\int_{\mathbb{R}^n} e^{\omega ix\cdot\xi} P(\xi,\varepsilon)d\xi = e^{-\varepsilon|\omega x|}$.

Proof. Noticing that

$$W(\xi,\varepsilon) = \mathscr{F}\left(\left(\frac{|\omega|}{2\pi}\right)^n e^{-\varepsilon|\omega x|^2}\right), \text{ and } P(\xi,\varepsilon) = \mathscr{F}\left(\left(\frac{|\omega|}{2\pi}\right)^n e^{-\varepsilon|\omega x|}\right),$$

we have the desired results by Theorem 1.22. □

We also have the semigroup properties of the Weierstrass and Poisson kernels.

Corollary 1.24. *If α_1 and α_2 are positive real numbers, then*
i) $W(\xi,\alpha_1+\alpha_2) = \int_{\mathbb{R}^n} W(\xi-\eta,\alpha_1)W(\eta,\alpha_2)d\eta$.
ii) $P(\xi,\alpha_1+\alpha_2) = \int_{\mathbb{R}^n} P(\xi-\eta,\alpha_1)P(\eta,\alpha_2)d\eta$.

Proof. It follows, from Corollary 1.23, that

$$\begin{aligned}W(\xi,\alpha_1+\alpha_2) =& \left(\frac{|\omega|}{2\pi}\right)^n (\mathscr{F}e^{-(\alpha_1+\alpha_2)|\omega x|^2})(\xi)\\ =&\left(\frac{|\omega|}{2\pi}\right)^n \mathscr{F}(e^{-\alpha_1|\omega x|^2}e^{-\alpha_2|\omega x|^2})(\xi)\\ =&\left(\frac{|\omega|}{2\pi}\right)^n \mathscr{F}\left(e^{-\alpha_1|\omega x|^2}\int_{\mathbb{R}^n} e^{\omega ix\cdot\eta}W(\eta,\alpha_2)d\eta\right)(\xi)\\ =&\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{-\omega ix\cdot\xi}e^{-\alpha_1|\omega x|^2}\int_{\mathbb{R}^n} e^{\omega ix\cdot\eta}W(\eta,\alpha_2)d\eta dx\\ =&\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{-\omega ix\cdot(\xi-\eta)}\left(\frac{|\omega|}{2\pi}\right)^n e^{-\alpha_1|\omega x|^2}dx\right)W(\eta,\alpha_2)d\eta\\ =&\int_{\mathbb{R}^n} W(\xi-\eta,\alpha_1)W(\eta,\alpha_2)d\eta.\end{aligned}$$

A similar argument can give the other equality. □

Finally, we give an example of the semigroup about the heat equation.

Example 1.25. Consider the Cauchy problem to the *heat equation*

$$u_t - \Delta u = 0, \quad u(0) = u_0(x), \quad t > 0,\ x \in \mathbb{R}^n.$$

Taking the Fourier transform, we have

$$\hat{u}_t + |\omega\xi|^2 \hat{u} = 0, \quad \hat{u}(0) = \hat{u}_0(\xi).$$

Thus, it follows, from Theorem 1.10, that

$$\begin{aligned} u =& \mathscr{F}^{-1} e^{-|\omega\xi|^2 t} \mathscr{F} u_0 = (\mathscr{F}^{-1} e^{-|\omega\xi|^2 t}) * u_0 = (4\pi t)^{-n/2} e^{-|x|^2/4t} * u_0 \\ =& W(x,t) * u_0 =: H(t) u_0. \end{aligned}$$

Then, we obtain

$$\begin{aligned} H(t_1+t_2)u_0 =& W(x, t_1+t_2) * u_0 = W(x,t_1) * W(x,t_2) * u_0 \\ =& W(x,t_1) * (W(x,t_2) * u_0) = W(x,t_1) * H(t_2) u_0 \\ =& H(t_1) H(t_2) u_0, \end{aligned}$$

i.e., $H(t_1+t_2) = H(t_1)H(t_2)$.

1.2 The L^2 theory and the Plancherel theorem

The integral defining the Fourier transform is not defined in the Lebesgue sense for the general function in $L^2(\mathbb{R}^n)$; nevertheless, the Fourier transform has a natural definition on this space and a particularly elegant theory.

If, in addition to being integrable, we assume f to be square-integrable then $\hat{f}$ will also be square-integrable. In fact, we have the following basic result:

Theorem 1.26 (Plancherel theorem)**.** *If* $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, *then* $\|\hat{f}\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|f\|_2$.

Proof. Let $g(x) = \overline{f(-x)}$. Then, by Theorem 1.1, $h = f * g \in L^1(\mathbb{R}^n)$ and, by Proposition 1.3, $\hat{h} = \hat{f}\hat{g}$. But $\hat{g} = \overline{\hat{f}}$, thus $\hat{h} = |\hat{f}|^2 \geqslant 0$. Applying Theorem 1.22, we have $\hat{h} \in L^1(\mathbb{R}^n)$ and $h(0) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \hat{h}(\xi) d\xi$. Thus, we get

$$\begin{aligned} \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 d\xi =& \int_{\mathbb{R}^n} \hat{h}(\xi) d\xi = \left(\frac{|\omega|}{2\pi}\right)^{-n} h(0) = \left(\frac{|\omega|}{2\pi}\right)^{-n} \int_{\mathbb{R}^n} f(x) g(0-x) dx \\ =& \left(\frac{|\omega|}{2\pi}\right)^{-n} \int_{\mathbb{R}^n} f(x)\overline{f(x)} dx = \left(\frac{|\omega|}{2\pi}\right)^{-n} \int_{\mathbb{R}^n} |f(x)|^2 dx, \end{aligned}$$

which completes the proof. □

Since $L^1 \cap L^2$ is dense in L^2, there exists a unique bounded extension, $\mathscr{F}$, of this operator to all of L^2. $\mathscr{F}$ will be called the Fourier transform on L^2; we shall also use the notation $\hat{f} = \mathscr{F} f$ whenever $f \in L^2(\mathbb{R}^n)$.

A linear operator on $L^2(\mathbb{R}^n)$ that is an isometry and maps onto $L^2(\mathbb{R}^n)$ is called a *unitary operator*. It is an immediate consequence of Theorem 1.26 that $\left(\frac{|\omega|}{2\pi}\right)^{n/2} \mathscr{F}$ is an isometry. Moreover, we have the additional property that $\left(\frac{|\omega|}{2\pi}\right)^{n/2} \mathscr{F}$ is onto:

Theorem 1.27. $\left(\frac{|\omega|}{2\pi}\right)^{n/2}$ *$\mathscr{F}$ is a unitary operator on* $L^2(\mathbb{R}^n)$.

Proof. Since $\left(\frac{|\omega|}{2\pi}\right)^{n/2}\mathscr{F}$ is an isometry, its range is a closed subspace of $L^2(\mathbb{R}^n)$. If this subspace were not all of $L^2(\mathbb{R}^n)$, we could find a function g such that $\int_{\mathbb{R}^n}\hat{f}g dx = 0$ for all $f \in L^2$ and $\|g\|_2 \neq 0$. Theorem 1.12 obviously extends to L^2; consequently, $\int_{\mathbb{R}^n} f\hat{g}dx = \int_{\mathbb{R}^n}\hat{f}gdx = 0$ for all $f \in L^2$. But this implies that $\hat{g}(x) = 0$ for almost every x, contradicting the fact that $\|\hat{g}\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\|g\|_2 \neq 0$. □

Theorem 1.27 is a major part of the basic theorem in the L^2 theory of the Fourier transform:

Theorem 1.28. *The inverse of the Fourier transform, $\mathscr{F}^{-1}$, can be obtained by letting*

$$(\mathscr{F}^{-1}f)(x) = \left(\frac{|\omega|}{2\pi}\right)^n (\mathscr{F}f)(-x)$$

for all $f \in L^2(\mathbb{R}^n)$.

We can also extend the definition of the Fourier transform to other spaces, such as Schwartz space, tempered distributions and so on.

1.3 Schwartz spaces

Distributions (generalized functions) aroused mostly due to Paul Dirac and his delta function δ. The Dirac delta gives a description of a point of unit mass (placed at the origin). The mass density function is such that if its integrated on a set not containing the origin it vanishes, but if the set does contain the origin it is 1. No function (in the traditional sense) can have this property because we know that the value of a function at a particular point does not change the value of the integral.

In mathematical analysis, distributions are objects which generalize functions and probability distributions. They extend the concept of derivative to all integrable functions and beyond, and are used to formulate generalized solutions of partial differential equations. They are important in physics and engineering where many non-continuous problems naturally lead to differential equations whose solutions are distributions, such as the Dirac delta distribution.

"Generalized functions" were introduced by Sergei Sobolev in 1935. They were independently introduced in late 1940s by Laurent Schwartz, who developed a comprehensive theory of distributions.

The basic idea in the theory of distributions is to consider them as linear functionals on some space of "regular" functions — the so-called "testing functions". The space of testing functions is assumed to be well-behaved with respect to the operations (differentiation, Fourier transform, convolution, translation, etc.) we have been studying, and this is then reflected in the properties of distributions.

We are naturally led to the definition of such a space of testing functions by the following considerations. Suppose we want these operations to be defined on a function space, $\mathscr{S}$, and to preserve it. Then, it would certainly have to consist of functions that are indefinitely differentiable; this, in view of part (v) in Proposition 1.3, indicates that each function in $\mathscr{S}$, after being multiplied by a polynomial, must still be in $\mathscr{S}$. We therefore make the following definition:

Definition 1.29. The *Schwartz space* $\mathscr{S}(\mathbb{R}^n)$ of rapidly decaying functions is defined as

$$\mathscr{S}(\mathbb{R}^n) = \left\{ \varphi \in C^\infty(\mathbb{R}^n) : |\varphi|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\alpha (\partial^\beta \varphi)(x)| < \infty, \ \forall \alpha, \beta \in \mathbb{N}_0^n \right\}, \tag{1.14}$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

If $\varphi \in \mathscr{S}$, then $|\varphi(x)| \leqslant C_m(1+|x|)^{-m}$ for any $m \in \mathbb{N}_0$. The second part of next example shows that the converse is not true.

Example 1.30. $\varphi(x) = e^{-\varepsilon|x|^2}$, $\varepsilon > 0$, belongs to $\mathscr{S}$; on the other hand, $\varphi(x) = e^{-\varepsilon|x|}$ fails to be differential at the origin and, therefore, does not belong to $\mathscr{S}$.

Example 1.31. $\varphi(x) = e^{-\varepsilon(1+|x|^2)^\gamma}$ belongs to $\mathscr{S}$ for any $\varepsilon, \gamma > 0$.

Example 1.32. $\mathscr{S}$ contains the space $\mathscr{D}(\mathbb{R}^n)$.

But it is not immediately clear that $\mathscr{D}$ is nonempty. To find a function in $\mathscr{D}$, consider the function

$$f(t) = \begin{cases} e^{-1/t}, & t > 0, \\ 0, & t \leqslant 0. \end{cases}$$

Then, $f \in C^\infty$, is bounded and so are all its derivatives. Let $\varphi(t) = f(1+t)f(1-t)$, then $\varphi(t) = e^{-2/(1-t^2)}$ if $|t| < 1$, is zero otherwise. It clearly belongs to $\mathscr{D} = \mathscr{D}(\mathbb{R}^1)$. We can easily obtain n-dimensional variants from φ. For examples,

(i) For $x \in \mathbb{R}^n$, define $\psi(x) = \varphi(x_1)\varphi(x_2)\cdots\varphi(x_n)$, then $\psi \in \mathscr{D}(\mathbb{R}^n)$;

(ii) For $x \in \mathbb{R}^n$, define $\psi(x) = e^{-2/(1-|x|^2)}$ for $|x| < 1$ and 0 otherwise, then $\psi \in \mathscr{D}(\mathbb{R}^n)$;

(iii) If $\eta \in C^\infty$ and ψ is the function in (ii), then $\psi(\varepsilon x)\eta(x)$ defines a function in $\mathscr{D}(\mathbb{R}^n)$; moreover, $e^2\psi(\varepsilon x)\eta(x) \to \eta(x)$ as $\varepsilon \to 1$.

Example 1.33. We observe that the order of multiplication by powers of $x_1, \cdots, x_n$ and differentiation, in (1.14), could have been reversed. That is, $\varphi \in \mathscr{S}$ if and only if $\varphi \in C^\infty$ and $\sup_{x\in\mathbb{R}^n} |\partial^\beta(x^\alpha\varphi(x))| < \infty$ for all multi-indices α and β of nonnegative integers. This shows that if P is a polynomial in n variables and $\varphi \in \mathscr{S}$ then $P(x)\varphi(x)$ and $P(\partial)\varphi(x)$ are again in $\mathscr{S}$, where $P(\partial)$ is the associated differential operator (i.e., we replace x^α by ∂^α in $P(x)$).

Example 1.34. Sometimes $\mathscr{S}(\mathbb{R}^n)$ is called the space of rapidly decaying functions. But observe that the function $\varphi(x) = e^{-x^2}e^{ie^{x^2}}$ is not in $\mathscr{S}(\mathbb{R})$. Hence, rapid decay of the value of the function alone does not assure the membership in $\mathscr{S}(\mathbb{R})$.

Theorem 1.35. *The spaces $\mathcal{C}_0(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, contain $\mathscr{S}(\mathbb{R}^n)$. Moreover, both $\mathscr{S}$ and $\mathscr{D}$ are dense in $\mathcal{C}_0(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$.*

Proof. $\mathscr{S} \subset \mathcal{C}_0 \subset L^\infty$ is obvious by (1.14). The L^p norm of $\varphi \in \mathscr{S}$ is bounded by a finite linear combination of L^∞ norms of terms of the form $x^\alpha\varphi(x)$. In fact, by (1.14), we have

$$\begin{aligned}
&\left(\int_{\mathbb{R}^n} |\varphi(x)|^p dx\right)^{1/p} \\
\leqslant &\left(\int_{|x|\leqslant 1} |\varphi(x)|^p dx\right)^{1/p} + \left(\int_{|x|>1} |\varphi(x)|^p dx\right)^{1/p} \\
\leqslant &\|\varphi\|_\infty \left(\int_{|x|\leqslant 1} dx\right)^{1/p} + \||x|^{2n}|\varphi(x)|\|_\infty \left(\int_{|x|>1} |x|^{-2np} dx\right)^{1/p} \\
= &\left(\frac{\omega_{n-1}}{n}\right)^{1/p} \|\varphi\|_\infty + \left(\frac{\omega_{n-1}}{(2p-1)n}\right)^{1/p} \left\||x|^{2n}|\varphi|\right\|_\infty \\
< &\infty.
\end{aligned}$$

For the proof of the density, we only need to prove the case of $\mathscr{D}$ since $\mathscr{D} \subset \mathscr{S}$. We will use the fact that the set of finite linear combinations of characteristic functions of bounded measurable sets in $\mathbb{R}^n$ is dense in $L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$. This is a well-known fact from functional analysis.

Now, let $E \subset \mathbb{R}^n$ be a bounded measurable set and let $\varepsilon > 0$. Then, there exists a closed set F and an open set Q such that $F \subset E \subset Q$ and $\mathbb{m}(Q \setminus F) < \varepsilon^p$ (or only $\mathbb{m}(Q) < \varepsilon^p$ if there is no closed set $F \subset E$). Here $\mathbb{m}$ is the Lebesgue measure in $\mathbb{R}^n$. Next, let φ be a function from $\mathscr{D}$ such that $\operatorname{supp} \varphi \subset Q$, $\varphi|_F \equiv 1$ and $0 \leqslant \varphi \leqslant 1$. Then,

$$\|\varphi - \chi_E\|_p^p = \int_{\mathbb{R}^n} |\varphi(x) - \chi_E(x)|^p dx \leqslant \int_{Q \setminus F} dx = \mathbb{m}(Q \setminus F) < \varepsilon^p$$

or

$$\|\varphi - \chi_E\|_p < \varepsilon,$$

where χ_E denotes the characteristic function of E. Thus, we may conclude that $\overline{\mathscr{D}(\mathbb{R}^n)} = L^p(\mathbb{R}^n)$ with respect to L^p measure for $1 \leqslant p < \infty$.

For the case of $\mathcal{C}_0$, we leave it to the interested reader. □

Remark 1.36. The density is not valid for $p = \infty$. Indeed, for a nonzero constant function $f \equiv c_0 \neq 0$ and for any function $\varphi \in \mathscr{D}(\mathbb{R}^n)$, we have

$$\|f - \varphi\|_\infty \geqslant |c_0| > 0.$$

Hence we cannot approximate any function from $L^\infty(\mathbb{R}^n)$ by functions from $\mathscr{D}(\mathbb{R}^n)$. This example also indicates that $\mathscr{S}$ is not dense in L^∞ since $\lim_{|x| \to \infty} |\varphi(x)| = 0$ for all $\varphi \in \mathscr{S}$.

From part (v) in Proposition 1.3, we immediately have

Theorem 1.37. *If $\varphi \in \mathscr{S}$, then $\hat{\varphi} \in \mathscr{S}$.*

If $\varphi, \psi \in \mathscr{S}$, then Theorem 1.37 implies that $\hat{\varphi}, \hat{\psi} \in \mathscr{S}$. Therefore, $\hat{\varphi}\hat{\psi} \in \mathscr{S}$. By part (vi) in Proposition 1.3, i.e., $\mathscr{F}(\varphi * \psi) = \hat{\varphi}\hat{\psi}$, an application of the inverse Fourier transform shows that

Theorem 1.38. *If $\varphi, \psi \in \mathscr{S}$, then $\varphi * \psi \in \mathscr{S}$.*

The space $\mathscr{S}(\mathbb{R}^n)$ is not a normed space because $|\varphi|_{\alpha,\beta}$ is only a semi-norm for multi-indices α and β, i.e., the condition

$$|\varphi|_{\alpha,\beta} = 0 \text{ if and only if } \varphi = 0$$

fails to hold, for example, for constant function φ. But the space $(\mathscr{S}, \rho)$ is a metric space if the metric ρ is defined by

$$\rho(\varphi, \psi) = \sum_{\alpha,\beta \in \mathbb{N}_0^n} 2^{-|\alpha|-|\beta|} \frac{|\varphi - \psi|_{\alpha,\beta}}{1 + |\varphi - \psi|_{\alpha,\beta}}.$$

Theorem 1.39 (Completeness)**.** *The space $(\mathscr{S}, \rho)$ is a complete metric space, i.e., every Cauchy sequence converges.*

Proof. Let $\{\varphi_k\}_{k=1}^\infty \subset \mathscr{S}$ be a Cauchy sequence. For any $\sigma > 0$ and any $\gamma \in \mathbb{N}_0^n$, let $\varepsilon = \frac{2^{-|\gamma|}\sigma}{1+2\sigma}$, then there exists $N_0(\varepsilon) \in \mathbb{N}$ such that $\rho(\varphi_k, \varphi_m) < \varepsilon$ when $k, m \geqslant N_0(\varepsilon)$ since $\{\varphi_k\}_{k=1}^\infty$ is a Cauchy sequence. Thus, we have

$$\frac{|\varphi_k - \varphi_m|_{0,\gamma}}{1 + |\varphi_k - \varphi_m|_{0,\gamma}} < \frac{\sigma}{1 + \sigma},$$

and then

$$\sup_{x \in K} |\partial^\gamma(\varphi_k - \varphi_m)| < \sigma$$

for any compact set $K \subset \mathbb{R}^n$. It means that $\{\varphi_k\}_{k=1}^{\infty}$ is a Cauchy sequence in the Banach space $C^{|\gamma|}(K)$. Hence, there exists a function $\varphi \in C^{|\gamma|}(K)$ such that

$$\lim_{k\to\infty} \varphi_k = \varphi, \text{ in } C^{|\gamma|}(K).$$

Thus, we can conclude that $\varphi \in C^{\infty}(\mathbb{R}^n)$. It only remains to prove that $\varphi \in \mathscr{S}$. It is clear that for any $\alpha, \beta \in \mathbb{N}_0^n$

$$\begin{aligned}\sup_{x\in K} |x^\alpha \partial^\beta \varphi| \leqslant& \sup_{x\in K} |x^\alpha \partial^\beta (\varphi_k - \varphi)| + \sup_{x\in K} |x^\alpha \partial^\beta \varphi_k| \\ \leqslant& C_\alpha(K) \sup_{x\in K} |\partial^\beta (\varphi_k - \varphi)| + \sup_{x\in K} |x^\alpha \partial^\beta \varphi_k|.\end{aligned}$$

Taking $k \to \infty$, we obtain

$$\sup_{x\in K} |x^\alpha \partial^\beta \varphi| \leqslant \limsup_{k\to\infty} |\varphi_k|_{\alpha,\beta} < \infty.$$

The last inequality is valid since $\{\varphi_k\}_{k=1}^{\infty}$ is a Cauchy sequence, so that $|\varphi_k|_{\alpha,\beta}$ is bounded. The last inequality doesn't depend on K either. Thus, $|\varphi|_{\alpha,\beta} < \infty$ and then $\varphi \in \mathscr{S}$. □

Moreover, some easily established properties of $\mathscr{S}$ and its topology, are the following:

Proposition 1.40. i) *The mapping $\varphi(x) \mapsto x^\alpha \partial^\beta \varphi(x)$ is continuous.*

ii) *If $\varphi \in \mathscr{S}$, then $\lim_{h\to 0} \tau_h \varphi = \varphi$.*

iii) *Suppose $\varphi \in \mathscr{S}$ and $h = (0, \cdots, h_i, \cdots, 0)$ lies on the i-th coordinate axis of $\mathbb{R}^n$, then the difference quotient $[\varphi - \tau_h \varphi]/h_i$ tends to $\partial\varphi/\partial x_i$ as $|h| \to 0$.*

iv) *The Fourier transform is a homeomorphism of $\mathscr{S}$ onto itself.*

v) *$\mathscr{S}$ is separable.*

Finally, we describe and prove a fundamental result of Fourier analysis that is known as the uncertainty principle. In fact this theorem was "discovered" by W. Heisenberg in the context of quantum mechanics. Expressed colloquially, the uncertainty principle says that it is not possible to know both the position and the momentum of a particle at the same time. Expressed more precisely, the uncertainty principle says that the position and the momentum cannot be simultaneously localized.

In the context of harmonic analysis, the uncertainty principle implies that one cannot at the same time localize the value of a function and its Fourier transform. The exact statement is as follows.

Theorem 1.41 (The Heisenberg uncertainty principle)**.** *Suppose ψ is a function in $\mathscr{S}(\mathbb{R})$. Then*

$$\|x\psi\|_2 \|\xi\hat{\psi}\|_2 \geqslant \left(\frac{|\omega|}{2\pi}\right)^{-1/2} \frac{\|\psi\|_2^2}{2|\omega|},$$

and equality holds if and only if $\psi(x) = Ae^{-Bx^2}$ where $B > 0$ and $A \in \mathbb{R}$.

Moreover, we have

$$\|(x - x_0)\psi\|_2 \|(\xi - \xi_0)\hat{\psi}\|_2 \geqslant \left(\frac{|\omega|}{2\pi}\right)^{-1/2} \frac{\|\psi\|_2^2}{2|\omega|}$$

for every x_0, $\xi_0 \in \mathbb{R}$.

Proof. The last inequality actually follows from the first by replacing $\psi(x)$ by $e^{-\omega i x \xi_0}\psi(x+x_0)$ (whose Fourier transform is $e^{\omega i x_0(\xi+\xi_0)}\hat{\psi}(\xi + \xi_0)$ by parts (ii) and (iii) in Proposition 1.3) and changing variables. To prove the first inequality, we argue as follows.

Since $\psi \in \mathscr{S}$, we know that ψ and ψ' are rapidly decreasing. Thus, an integration by parts gives

$$\|\psi\|_2^2 = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = -\int_{-\infty}^{\infty} x \frac{d}{dx} |\psi(x)|^2 dx$$

$$= -\int_{-\infty}^{\infty} \left(x\psi'(x)\overline{\psi(x)} + x\overline{\psi'(x)}\psi(x)\right) dx.$$

The last identity follows because $|\psi|^2 = \psi\overline{\psi}$. Therefore,

$$\|\psi\|_2^2 \leqslant 2\int_{-\infty}^{\infty} |x||\psi(x)||\psi'(x)|dx \leqslant 2\|x\psi\|_2\|\psi'\|_2,$$

where we have used the Cauchy-Schwarz inequality. By part (v) in Proposition 1.3, we have $\mathscr{F}(\psi')(\xi) = \omega i\xi\hat{\psi}(\xi)$. It follows, from the Plancherel theorem, that

$$\|\psi'\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{1/2} \|\mathscr{F}(\psi')\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{1/2} |\omega|\|\xi\hat{\psi}\|_2.$$

Thus, we conclude the proof of the inequality in the theorem.

If equality holds, then we must also have equality where we applied the Cauchy-Schwarz inequality, and as a result, we find that $\psi'(x) = \beta x\psi(x)$ for some constant β. The solutions to this equation are $\psi(x) = Ae^{\beta x^2/2}$, where A is a constant. Since we want ψ to be a Schwartz function, we must take $\beta = -2B < 0$. $\square$

1.4 The class of tempered distributions

The collection $\mathscr{S}'$ of all continuous linear functionals on $\mathscr{S}$ is called the *space of tempered distributions*. That is

Definition 1.42. The functional $T : \mathscr{S} \to \mathbb{C}$ is a *tempered distribution* if

i) T is linear, i.e., $\langle T, \alpha\varphi + \beta\psi\rangle = \alpha\langle T, \varphi\rangle + \beta\langle T, \psi\rangle$ for all $\alpha, \beta \in \mathbb{C}$ and $\varphi, \psi \in \mathscr{S}$.

ii) T is continuous on $\mathscr{S}$, i.e., there exist $n_0 \in \mathbb{N}_0$ and a constant $c_0 > 0$ such that

$$|\langle T, \varphi\rangle| \leqslant c_0 \sum_{|\alpha|,|\beta|\leqslant n_0} |\varphi|_{\alpha,\beta}$$

for any $\varphi \in \mathscr{S}$.

In addition, for $T_k, T \in \mathscr{S}'$, the convergence $T_k \to T$ in $\mathscr{S}'$ means that $\langle T_k, \varphi\rangle \to \langle T, \varphi\rangle$ in $\mathbb{C}$ for all $\varphi \in \mathscr{S}$.

Remark 1.43. Since $\mathscr{D} \subset \mathscr{S}$, the space of tempered distributions $\mathscr{S}'$ is more narrow than the space of distributions $\mathscr{D}'$, i.e., $\mathscr{S}' \subset \mathscr{D}'$. Another more narrow distribution space $\mathscr{E}'$ which consists of continuous linear functionals on the (widest test function) space $\mathscr{E} := C^\infty(\mathbb{R}^n)$. In short, $\mathscr{D} \subset \mathscr{S} \subset \mathscr{E}$ implies that

$$\mathscr{E}' \subset \mathscr{S}' \subset \mathscr{D}'.$$

Example 1.44. Let $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, and define $T = T_f$ by letting

$$\langle T, \varphi\rangle = \langle T_f, \varphi\rangle = \int_{\mathbb{R}^n} f(x)\varphi(x)dx$$

for $\varphi \in \mathscr{S}$. It is clear that T_f is a linear functional on $\mathscr{S}$. To show that it is continuous, therefore, it suffices to show that it is continuous at the origin. Then, suppose $\varphi_k \to 0$ in $\mathscr{S}$ as $k \to \infty$. From the proof of Theorem 1.35, we have seen that for any $q \geqslant 1$, $\|\varphi_k\|_q$ is dominated by a finite linear combination of L^∞ norms of terms of the form $x^\alpha\varphi_k(x)$. That is, $\|\varphi_k\|_q$ is dominated by a finite linear combination of semi-norms $|\varphi_k|_{\alpha,0}$. Thus, $\|\varphi_k\|_q \to 0$ as $k \to \infty$. Choosing $q = p'$, i.e., $1/p + 1/q = 1$, Hölder's inequality shows that $|\langle T, \varphi_k\rangle| \leqslant \|f\|_p\|\varphi_k\|_{p'} \to 0$ as $k \to \infty$. Thus, $T \in \mathscr{S}'$.

Example 1.45. We consider the case $n = 1$. Let $f(x) = \sum_{k=0}^{m} a_k x^k$ be a polynomial, then $f \in \mathscr{S}'$ since

$$\begin{aligned}|\langle T_f, \varphi\rangle| =& \left|\int_{\mathbb{R}} \sum_{k=0}^{m} a_k x^k \varphi(x) dx\right| \\ \leqslant& \sum_{k=0}^{m} |a_k| \int_{\mathbb{R}} (1+|x|)^{-1-\varepsilon}(1+|x|)^{1+\varepsilon}|x|^k|\varphi(x)|dx \\ \leqslant& C\sum_{k=0}^{m} |a_k||\varphi|_{k+1+\varepsilon,0} \int_{\mathbb{R}} (1+|x|)^{-1-\varepsilon} dx,\end{aligned}$$

so that the condition ii) of the definition is satisfied for $\varepsilon = 1$ and $n_0 = m + 2$.

Example 1.46. Fix $x_0 \in \mathbb{R}^n$ and a multi-index $\beta \in \mathbb{N}_0^n$. By the continuity of the semi-norm $|\cdot|_{\alpha,\beta}$ in $\mathscr{S}$, we have that $\langle T, \varphi\rangle = \partial^\beta \varphi(x_0)$, for $\varphi \in \mathscr{S}$, defines a tempered distribution. A special case is the Dirac δ-function: $\langle T_\delta, \varphi\rangle = \varphi(0)$.

The tempered distributions of Examples 1.44-1.46 are called functions or measures. We shall write, in these cases, f and δ instead of T_f and T_δ. These functions and measures may be considered as embedded in $\mathscr{S}'$. If we put on $\mathscr{S}'$ the weakest topology such that the linear functionals $T \to \langle T, \varphi\rangle$ ($\varphi \in \mathscr{S}$) are continuous, it is easy to see that the spaces $L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, are continuously embedded in $\mathscr{S}'$. The same is true for the space of all finite Borel measures on $\mathbb{R}^n$, i.e., $\mathscr{B}(\mathbb{R}^n)$.

There exists a simple and important characterization of tempered distributions:

Theorem 1.47. *A linear functional T on $\mathscr{S}$ is a tempered distribution if and only if there exists a constant $C > 0$ and integers ℓ and m such that*

$$|\langle T, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant \ell, |\beta|\leqslant m} |\varphi|_{\alpha,\beta}$$

for all $\varphi \in \mathscr{S}$.

Proof. It is clear that the existence of C, ℓ, m implies the continuity of T.

Suppose T is continuous. It follows from the definition of the metric that a basis for the neighborhoods of the origin in $\mathscr{S}$ is the collection of sets $N_{\varepsilon,\ell,m} = \{\varphi : \sum_{|\alpha|\leqslant \ell,|\beta|\leqslant m} |\varphi|_{\alpha,\beta} < \varepsilon\}$, where $\varepsilon > 0$ and ℓ and m are integers, because $\varphi_k \to \varphi$ as $k \to \infty$ if and only if $|\varphi_k - \varphi|_{\alpha,\beta} \to 0$ for all (α, β) in the topology induced by this system of neighborhoods and their translates. Thus, there exists such a set $N_{\varepsilon,\ell,m}$ satisfying $|\langle T, \varphi\rangle| \leqslant 1$ whenever $\varphi \in N_{\varepsilon,\ell,m}$.

Let $\|\varphi\| = \sum_{|\alpha|\leqslant \ell,|\beta|\leqslant m} |\varphi|_{\alpha,\beta}$ for all $\varphi \in \mathscr{S}$. If $\sigma \in (0, \varepsilon)$, then $\psi = \sigma\varphi/\|\varphi\| \in N_{\varepsilon,\ell,m}$ if $\varphi \neq 0$. From the linearity of T, we obtain

$$\frac{\sigma}{\|\varphi\|}|\langle T, \varphi\rangle| = |\langle T, \psi\rangle| \leqslant 1.$$

But this is the desired inequality with $C = 1/\sigma$. □

Example 1.48. Let $T \in \mathscr{S}'$ and $\varphi \in \mathscr{D}(\mathbb{R}^n)$ with $\varphi(0) = 1$. Then the product $\varphi(x/k)T$ is well-defined in $\mathscr{S}'$ by

$$\langle \varphi(x/k)T, \psi\rangle := \langle T, \varphi(x/k)\psi\rangle,$$

for all $\psi \in \mathscr{S}$. If we consider the sequence $T_k := \varphi(x/k)T$, then

$$\langle T_k, \psi\rangle \equiv \langle T, \varphi(x/k)\psi\rangle \to \langle T, \psi\rangle$$

as $k \to \infty$ since $\varphi(x/k)\psi \to \psi$ in $\mathscr{S}$. Thus, $T_k \to T$ in $\mathscr{S}'$ as $k \to \infty$. Moreover, T_k has compact support as a tempered distribution in view of the compactness of $\varphi_k = \varphi(x/k)$.

Now we are ready to prove more serious and more useful fact.

Theorem 1.49. *Let $T \in \mathscr{S}'$, then there exists a sequence $\{T_k\}_{k=0}^{\infty} \subset \mathscr{S}$ such that*

$$\langle T_k, \varphi\rangle = \int_{\mathbb{R}^n} T_k(x)\varphi(x)dx \to \langle T, \varphi\rangle, \quad \text{as } k \to \infty,$$

where $\varphi \in \mathscr{S}$. In short, $\mathscr{S}$ is dense in $\mathscr{S}'$ with respect to the topology on $\mathscr{S}'$.

Proof. If h and g are integrable functions and $\varphi \in \mathscr{S}$, then it follows, from Fubini's theorem, that

$$\begin{aligned}\langle h * g, \varphi\rangle &= \int_{\mathbb{R}^n} \varphi(x) \int_{\mathbb{R}^n} h(x-y)g(y)dydx = \int_{\mathbb{R}^n} g(y) \int_{\mathbb{R}^n} h(x-y)\varphi(x)dxdy \\ &= \int_{\mathbb{R}^n} g(y) \int_{\mathbb{R}^n} Rh(y-x)\varphi(x)dxdy = \langle g, Rh * \varphi\rangle,\end{aligned}$$

where $Rh(x) := h(-x)$ is the reflection of h.

Let now $\psi \in \mathscr{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \psi(x)dx = 1$ and $\psi(-x) = \psi(x)$. Let $\zeta \in \mathscr{D}(\mathbb{R}^n)$ with $\zeta(0) = 1$. Denote $\psi_k(x) := k^n\psi(kx)$. For any $T \in \mathscr{S}'$, denote $T_k := \psi_k * \tilde{T}_k$, where $\tilde{T}_k = \zeta(x/k)T$. From above considerations, we know that $\langle \psi_k * \tilde{T}_k, \varphi\rangle = \langle \tilde{T}_k, R\psi_k * \varphi\rangle$.

Let us prove that these T_k meet the requirements of the theorem. In fact, we have

$$\begin{aligned}\langle T_k, \varphi\rangle \equiv& \langle \psi_k * \tilde{T}_k, \varphi\rangle = \langle \tilde{T}_k, R\psi_k * \varphi\rangle = \langle \zeta(x/k)T, \psi_k * \varphi\rangle \\ =& \langle T, \zeta(x/k)(\psi_k * \varphi)\rangle \to \langle T, \varphi\rangle, \quad \text{as } k \to \infty,\end{aligned}$$

by the fact $\psi_k * \varphi \to \varphi$ in $\mathscr{S}$ as $k \to \infty$ in view of Theorem 1.15, and the fact $\zeta(x/k) \to 1$ pointwise as $k \to \infty$ since $\zeta(0) = 1$ and $\zeta(x/k)\varphi \to \varphi$ in $\mathscr{S}$ as $k \to \infty$. Finally, since ψ_k, $\zeta \in \mathscr{D}(\mathbb{R}^n)$, it follows that $T_k \in \mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$. □

Definition 1.50. Let $L : \mathscr{S} \to \mathscr{S}$ be a linear continuous mapping. Then, the dual/conjugate mapping $L' : \mathscr{S}' \to \mathscr{S}'$ is defined by

$$\langle L'T, \varphi\rangle := \langle T, L\varphi\rangle, \quad T \in \mathscr{S}', \ \varphi \in \mathscr{S}.$$

Clearly, L' is also a linear continuous mapping.

Corollary 1.51. *Any linear continuous mapping (or operator) $L : \mathscr{S} \to \mathscr{S}$ admits a linear continuous extension $\tilde{L} : \mathscr{S}' \to \mathscr{S}'$.*

Proof. If $T \in \mathscr{S}'$, then by Theorem 1.49, there exists a sequence $\{T_k\}_{k=0}^{\infty} \subset \mathscr{S}$ such that $T_k \to T$ in $\mathscr{S}'$ as $k \to \infty$. Hence,

$$\langle LT_k, \varphi\rangle = \langle T_k, L'\varphi\rangle \to \langle T, L'\varphi\rangle := \langle \tilde{L}T, \varphi\rangle, \quad \text{as } k \to \infty,$$

for any $\varphi \in \mathscr{S}$. □

Now, we can list the properties of tempered distributions about the multiplication, differentiation, translation, dilation and Fourier transform.

Theorem 1.52. *The following linear continuous operators from $\mathscr{S}$ into $\mathscr{S}$ admit unique linear continuous extensions as maps from $\mathscr{S}'$ into $\mathscr{S}'$: For $T \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$,*

i) $\langle \psi T, \varphi\rangle := \langle T, \psi\varphi\rangle$, $\psi \in \mathscr{S}$.
ii) $\langle \partial^\alpha T, \varphi\rangle := \langle T, (-1)^{|\alpha|}\partial^\alpha\varphi\rangle$, $\alpha \in \mathbb{N}_0^n$.
iii) $\langle \tau_h T, \varphi\rangle := \langle T, \tau_{-h}\varphi\rangle$, $h \in \mathbb{R}^n$.
iv) $\langle \delta_\lambda T, \varphi\rangle := \langle T, |\lambda|^{-n}\delta_{1/\lambda}\varphi\rangle$, $0 \neq \lambda \in \mathbb{R}$.
v) $\langle \mathscr{F}T, \varphi\rangle := \langle T, \mathscr{F}\varphi\rangle$.

Proof. See the previous definition, Theorem 1.49 and its corollary. □

Remark 1.53. Since $\langle \mathscr{F}^{-1}\mathscr{F}T, \varphi\rangle = \langle \mathscr{F}T, \mathscr{F}^{-1}\varphi\rangle = \langle T, \mathscr{F}\mathscr{F}^{-1}\varphi\rangle = \langle T, \varphi\rangle$, we get $\mathscr{F}^{-1}\mathscr{F} = \mathscr{F}\mathscr{F}^{-1} = I$ in $\mathscr{S}'$.

Example 1.54. Since for any $\varphi \in \mathscr{S}$,

$$\begin{aligned}\langle \mathscr{F}1, \varphi\rangle = \langle 1, \mathscr{F}\varphi\rangle = \int_{\mathbb{R}^n} (\mathscr{F}\varphi)(\xi) d\xi = \left(\frac{|\omega|}{2\pi}\right)^{-n} \left(\frac{|\omega|}{2\pi}\right)^{n} \int_{\mathbb{R}^n} e^{\omega i 0\cdot\xi} (\mathscr{F}\varphi)(\xi) d\xi \\ = \left(\frac{|\omega|}{2\pi}\right)^{-n} \mathscr{F}^{-1}\mathscr{F}\varphi(0) = \left(\frac{|\omega|}{2\pi}\right)^{-n} \varphi(0) = \left(\frac{|\omega|}{2\pi}\right)^{-n} \langle \delta, \varphi\rangle,\end{aligned}$$

we have

$$\hat{1} = \left(\frac{|\omega|}{2\pi}\right)^{-n} \delta, \quad \text{in } \mathscr{S}'.$$

Moreover, $\check{\delta} = \left(\frac{|\omega|}{2\pi}\right)^{n} \cdot 1$.

Example 1.55. For $\varphi \in \mathscr{S}$, we have

$$\langle \hat{\delta}, \varphi\rangle = \langle \delta, \mathscr{F}\varphi\rangle = \hat{\varphi}(0) = \int_{\mathbb{R}^n} e^{-\omega i x\cdot 0} \varphi(x) dx = \langle 1, \varphi\rangle.$$

Thus, $\hat{\delta} = 1$ in $\mathscr{S}'$.

Example 1.56. Since

$$\begin{aligned}\langle \widehat{\partial^\alpha \delta}, \varphi\rangle =& \langle \partial^\alpha \delta, \hat{\varphi}\rangle = (-1)^{|\alpha|} \langle \delta, \partial^\alpha \hat{\varphi}\rangle = \langle \delta, \mathscr{F}[(\omega i \xi)^\alpha \varphi]\rangle \\ =& \langle \hat{\delta}, (\omega i \xi)^\alpha \varphi\rangle = \langle (\omega i \xi)^\alpha, \varphi\rangle,\end{aligned}$$

we have $\widehat{\partial^\alpha \delta} = (\omega i \xi)^\alpha$.

Now, we shall show that the convolution can be defined on the class $\mathscr{S}'$. We first recall a notation we have used: If g is any function on $\mathbb{R}^n$, we define its reflection, Rg, by letting $Rg(x) = g(-x)$. A direct application of Fubini's theorem shows that if u, φ and ψ are all in $\mathscr{S}$, then

$$\int_{\mathbb{R}^n} (u * \varphi)(x)\psi(x) dx = \int_{\mathbb{R}^n} u(x)(R\varphi * \psi)(x) dx.$$

The mappings $\psi \mapsto \int_{\mathbb{R}^n} (u * \varphi)(x)\psi(x)dx$ and $\theta \mapsto \int_{\mathbb{R}^n} u(x)\theta(x)dx$ are linear functionals on $\mathscr{S}$. If we denote these functionals by $u * \varphi$ and u, the last equality can be written in the form:

$$\langle u * \varphi, \psi\rangle = \langle u, R\varphi * \psi\rangle. \tag{1.15}$$

If $u \in \mathscr{S}'$ and φ, $\psi \in \mathscr{S}$, the right side of (1.15) is well-defined since $R\varphi * \psi \in \mathscr{S}$. Furthermore, the mapping $\psi \mapsto \langle u, R\varphi * \psi\rangle$, being the composition of two continuous functions, is continuous. Thus, we can define the convolution of the distribution u with the testing function φ, $u * \varphi$, by means of equality (1.15).

It is easy to show that this convolution is associative in the sense that $(u * \varphi) * \psi = u * (\varphi * \psi)$ whenever $u \in \mathscr{S}'$ and φ, $\psi \in \mathscr{S}$. The following result is a characterization of the convolution we have just described.

Theorem 1.57. *If $u \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$, then the convolution $u * \varphi$ is the function f, whose value at $x \in \mathbb{R}^n$ is $f(x) = \langle u, \tau_x R\varphi\rangle$, where τ_x denotes the translation by x operator. Moreover, f belongs to the class C^∞ and it, as well as all its derivatives, are slowly increasing.*

Proof. We first show that f is C^∞ slowly increasing. Let $h = (0, \cdots, h_j, \cdots, 0)$, then by part iii) in Proposition 1.40,

$$\frac{\tau_{x+h} R\varphi(y) - \tau_x R\varphi(y)}{h_j} \to -\tau_x \frac{\partial R\varphi}{\partial y_j}(y),$$

as $|h| \to 0$, in the topology of $\mathscr{S}$. Thus, since u is continuous, we have

$$\frac{f(x+h)-f(x)}{h_j} = \langle u, \frac{\tau_{x+h}R\varphi(y)-\tau_x R\varphi(y)}{h_j}\rangle \to \langle u, -\tau_x \frac{\partial R\varphi}{\partial y_j}(y)\rangle$$

as $h_j \to 0$. This, together with ii) in Proposition 1.40, shows that f has continuous first partial derivatives. Since $\partial R\varphi/\partial y_j \in \mathscr{S}$, we can iterate this argument and show that $\partial^\beta f$ exists and is continuous for all multi-index $\beta \in \mathbb{N}_0^n$. We observe that $\partial^\beta f(x) = \langle u, (-1)^{|\beta|}\tau_x \partial^\beta R\varphi\rangle$. Consequently, since $\partial^\beta R\varphi \in \mathscr{S}$, if f were slowly increasing, then the same would hold for all the derivatives of f. In fact, that f is slowly increasing is an easy consequence of Theorem 1.47: There exist $C > 0$ and integers ℓ and m such that

$$|f(x)| = |\langle u, \tau_x R\varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant \ell, |\beta|\leqslant m} |\tau_x R\varphi|_{\alpha,\beta}.$$

But $|\tau_x R\varphi|_{\alpha,\beta} = \sup_{y\in\mathbb{R}^n} |y^\alpha \partial^\beta R\varphi(y-x)| = \sup_{y\in\mathbb{R}^n} |(y+x)^\alpha \partial^\beta R\varphi(y)|$ and the latter is clearly bounded by a polynomial in x.

In order to show that $u * \varphi$ is the function f, we must show that $\langle u * \varphi, \psi\rangle = \int_{\mathbb{R}^n} f(x)\psi(x)dx$. But,

$$\begin{aligned}\langle u * \varphi, \psi\rangle =& \langle u, R\varphi * \psi\rangle = \langle u, \int_{\mathbb{R}^n} R\varphi(y-x)\psi(x)dx\rangle = \langle u, \int_{\mathbb{R}^n} \tau_x R\varphi(y)\psi(x)dx\rangle \\ =& \int_{\mathbb{R}^n} \langle u, \tau_x R\varphi(y)\rangle \psi(x)dx = \int_{\mathbb{R}^n} f(x)\psi(x)dx,\end{aligned}$$

since u is continuous and linear and the fact that the integral $\int_{\mathbb{R}^n} \tau_x R\varphi(y)\psi(x)dx$ converges in $\mathscr{S}$, which is the desired equality. □

1.5 Characterization of operators commuting with translations

Having set down these facts of distribution theory, we shall now apply them to the study of the basic class of linear operators that occur in Fourier analysis: the class of operators that commute with translations.

Definition 1.58. A vector space X of measurable functions on $\mathbb{R}^n$ is called *closed under translations* if for $f \in X$ we have $\tau_y f \in X$ for all $y \in \mathbb{R}^n$. Let X and Y be vector spaces of measurable functions on $\mathbb{R}^n$ that are closed under translations. Let also T be an operator from X to Y. We say that T *commutes with translations* or is *translation invariant* if

$$T(\tau_y f) = \tau_y(Tf)$$

for all $f \in X$ and all $y \in \mathbb{R}^n$.

It is automatic to see that convolution operators commute with translations. One of the main goals of this section is to prove the converse, i.e., every bounded linear operator that commutes with translations is of convolution type. We have the following:

Theorem 1.59. *Let $1 \leqslant p, q \leqslant \infty$. Suppose T is a bounded linear operator from $L^p(\mathbb{R}^n)$ into $L^q(\mathbb{R}^n)$ that commutes with translations. Then there exists a unique tempered distribution u such that*

$$Tf = u * f, \quad \forall f \in \mathscr{S}.$$

The theorem will be a consequence of the following lemma.

Lemma 1.60. *Let $1 \leqslant p \leqslant \infty$. If $f \in L^p(\mathbb{R}^n)$ has derivatives in the L^p norm of all orders $\leqslant n+1$, then f equals almost everywhere a continuous function g satisfying*

$$|g(0)| \leqslant C \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha f\|_p,$$

where C depends only on the dimension n and the exponent p.

Proof. Let $\xi \in \mathbb{R}^n$. Then there exists a C'_n such that

$$(1+|\xi|^2)^{(n+1)/2} \leqslant (1+|\xi_1|+\cdots+|\xi_n|)^{n+1} \leqslant C'_n \sum_{|\alpha| \leqslant n+1} |\xi^\alpha|.$$

Let us first suppose $p=1$, we shall show $\hat{f} \in L^1$. By part (v) in Proposition 1.3 and part (i) in Theorem 1.5, we have

$$\begin{aligned}|\hat{f}(\xi)| \leqslant & C'_n (1+|\xi|^2)^{-(n+1)/2} \sum_{|\alpha| \leqslant n+1} |\xi^\alpha||\hat{f}(\xi)| \\ = & C'_n (1+|\xi|^2)^{-(n+1)/2} \sum_{|\alpha| \leqslant n+1} |\omega|^{-|\alpha|} |\mathscr{F}(\partial^\alpha f)(\xi)| \\ \leqslant & C''(1+|\xi|^2)^{-(n+1)/2} \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha f\|_1.\end{aligned}$$

Since $(1+|\xi|^2)^{-(n+1)/2}$ defines an integrable function on $\mathbb{R}^n$, it follows that $\hat{f} \in L^1(\mathbb{R}^n)$ and, letting $C''' = C'' \int_{\mathbb{R}^N} (1+|\xi|^2)^{-(n+1)/2} d\xi$, we get

$$\|\hat{f}\|_1 \leqslant C''' \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha f\|_1.$$

Thus, by Theorem 1.19, f equals almost everywhere a continuous function g and by Theorem 1.5,

$$|g(0)| \leqslant \|f\|_\infty \leqslant \left(\frac{|\omega|}{2\pi}\right)^n \|\hat{f}\|_1 \leqslant C \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha f\|_1.$$

Suppose now that $p>1$. Choose $\varphi \in \mathscr{D}(\mathbb{R}^n)$ such that $\varphi(x)=1$ if $|x| \leqslant 1$ and $\varphi(x)=0$ if $|x|>2$. Then, it is clear that $f\varphi \in L^1(\mathbb{R}^n)$. Thus, $f\varphi$ equals almost everywhere a continuous function h such that

$$|h(0)| \leqslant C \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha (f\varphi)\|_1.$$

By Leibniz' rule for differentiation, we have $\partial^\alpha(f\varphi) = \sum_{\mu+\nu=\alpha} \frac{\alpha!}{\mu!\nu!} \partial^\mu f \partial^\nu \varphi$, and then

$$\begin{aligned}\|\partial^\alpha(f\varphi)\|_1 &\leqslant \int_{|x|\leqslant 2} \sum_{\mu+\nu=\alpha} \frac{\alpha!}{\mu!\nu!} |\partial^\mu f||\partial^\nu \varphi| dx \\ &\leqslant \sum_{\mu+\nu=\alpha} C \sup_{|x|\leqslant 2} |\partial^\nu \varphi(x)| \int_{|x|\leqslant 2} |\partial^\mu f(x)| dx \\ &\leqslant A \sum_{|\mu| \leqslant |\alpha|} \int_{|x|\leqslant 2} |\partial^\mu f(x)| dx \leqslant AB \sum_{|\mu|\leqslant|\alpha|} \|\partial^\mu f\|_p,\end{aligned}$$

where $A \geqslant \|\partial^\nu \varphi\|_\infty$, $|\nu| \leqslant |\alpha|$, and B depends only on p and n. Thus, we can find a constant K such that

$$|h(0)| \leqslant K \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha f\|_p.$$

Since $\varphi(x)=1$ if $|x| \leqslant 1$, we see that f is equal almost everywhere to a continuous function g in the sphere of radius 1 centered at 0, moreover,

$$|g(0)| = |h(0)| \leqslant K \sum_{|\alpha| \leqslant n+1} \|\partial^\alpha f\|_p.$$

But, by choosing φ appropriately, the argument clearly shows that f equals almost everywhere a continuous function on any sphere centered at 0. This proves the lemma. $\square$

Now, we turn to the proof of the previous theorem.

Proof of Theorem 1.59. We first prove that

$$\partial^\beta Tf = T\partial^\beta f, \quad \forall f \in \mathscr{S}(\mathbb{R}^n). \tag{1.16}$$

In fact, if $h = (0, \cdots, h_j, \cdots, 0)$ lies on the j-th coordinate axis, we have

$$\frac{\tau_h(Tf) - Tf}{h_j} = \frac{T(\tau_h f) - Tf}{h_j} = T\left(\frac{\tau_h f - f}{h_j}\right),$$

since T is linear and commuting with translations. By part iii) in Proposition 1.40, $\frac{\tau_h f - f}{h_j} \to -\frac{\partial f}{\partial x_j}$ in $\mathscr{S}$ as $|h| \to 0$ and also in L^p norm due to the density of $\mathscr{S}$ in L^p. Since T is bounded operator from L^p to L^q, it follows that $\frac{\tau_h(Tf)-Tf}{h_j} \to -\frac{\partial Tf}{\partial x_j}$ in L^q as $|h| \to 0$. By induction, we get (1.16). By Lemma 1.60, Tf equals almost everywhere a continuous function g_f satisfying

$$|g_f(0)| \leqslant C \sum_{|\beta|\leqslant n+1} \|\partial^\beta (Tf)\|_q = C \sum_{|\beta|\leqslant n+1} \|T(\partial^\beta f)\|_q \leqslant \|T\| C \sum_{|\beta|\leqslant n+1} \|\partial^\beta f\|_p.$$

From the proof of Theorem 1.35, we know that the L^p norm of $f \in \mathscr{S}$ is bounded by a finite linear combination of L^∞ norms of terms of the form $x^\alpha f(x)$. Thus, there exists a $m \in \mathbb{N}$ such that $|g_f(0)| \leqslant C \sum_{|\alpha|\leqslant m, |\beta|\leqslant n+1} \|x^\alpha \partial^\beta f\|_\infty = C \sum_{|\alpha|\leqslant m, |\beta|\leqslant n+1} |f|_{\alpha,\beta}$. Then, by Theorem 1.47, the mapping $f \mapsto g_f(0)$ is a continuous linear functional on $\mathscr{S}$, denoted by u_1. We claim that $u = Ru_1$ is the linear functional we are seeking. Indeed, if $f \in \mathscr{S}$, using Theorem 1.57, we obtain

$$\begin{aligned}(u * f)(x) =& \langle u, \tau_x Rf\rangle = \langle u, R(\tau_{-x} f)\rangle = \langle Ru, \tau_{-x} f\rangle = \langle u_1, \tau_{-x} f\rangle \\ =& (T(\tau_{-x} f))(0) = (\tau_{-x} Tf)(0) = Tf(x).\end{aligned}$$

We note that it follows from this construction that u is unique. The theorem is therefore proved. $\square$

Combining this result with Theorem 1.57, we obtain the fact that Tf, for $f \in \mathscr{S}$, is almost everywhere equal to a C^∞ function which, together with all its derivatives, is slowly increasing.

Now, we give a characterization of operators commuting with translations in $L^1(\mathbb{R}^n)$.

Theorem 1.61. *Let T be a bounded linear operator mapping $L^1(\mathbb{R}^n)$ to itself. Then a necessary and sufficient condition that T commutes with translations is that there exists a measure μ in $\mathscr{B}(\mathbb{R}^n)$ such that $Tf = \mu * f$, for all $f \in L^1(\mathbb{R}^n)$. One has then $\|T\| = \|\mu\|$.*

Proof. We first prove the sufficiency. Suppose that $Tf = \mu * f$ for a measure $\mu \in \mathscr{B}(\mathbb{R}^n)$ and all $f \in L^1(\mathbb{R}^n)$. Since $\mathscr{B} \subset \mathscr{S}'$, by Theorem 1.57, we have

$$\begin{aligned}\tau_h(Tf)(x) =& (Tf)(x-h) = \langle \mu, \tau_{x-h} Rf\rangle = \langle \mu(y), f(-y-x+h)\rangle \\ =& \langle \mu, \tau_x R\tau_h f\rangle = \mu * \tau_h f = T\tau_h f,\end{aligned}$$

i.e., $\tau_h T = T\tau_h$. On the other hand, we have $\|Tf\|_1 = \|\mu * f\|_1 \leqslant \|\mu\| \|f\|_1$ which implies $\|T\| = \|\mu\|$.

Now, we prove the necessariness. Suppose that T commutes with translations and $\|Tf\|_1 \leqslant \|T\| \|f\|_1$ for all $f \in L^1(\mathbb{R}^n)$. Then, by Theorem 1.59, there exists a unique tempered distribution μ such that $Tf = \mu * f$ for all $f \in \mathscr{S}$. The remainder is to prove $\mu \in \mathscr{B}(\mathbb{R}^n)$.

We consider the family of L^1 functions $\mu_\varepsilon = \mu * W(\cdot, \varepsilon) = TW(\cdot, \varepsilon)$, $\varepsilon > 0$. Then by assumption and Lemma 1.14, we get

$$\|\mu_\varepsilon\|_1 \leqslant \|T\|\|W(\cdot,\varepsilon)\|_1 = \|T\|.$$

That is, the family $\{\mu_\varepsilon\}$ is uniformly bounded in the L^1 norm. Let us consider $L^1(\mathbb{R}^n)$ as embedded in the Banach space $\mathscr{B}(\mathbb{R}^n)$. $\mathscr{B}(\mathbb{R}^n)$ can be identified with the dual of $\mathcal{C}_0(\mathbb{R}^n)$ by making each $\nu \in \mathscr{B}$ corresponding to the linear functional assigning to $\varphi \in \mathcal{C}_0$ the value $\int_{\mathbb{R}^n} \varphi(x)d\nu(x)$. Thus, the unit sphere of $\mathscr{B}$ is compact in the weak* topology. In particular, we can find a $\nu \in \mathscr{B}$ and a null sequence $\{\varepsilon_k\}$ such that $\mu_{\varepsilon_k} \to \nu$ as $k \to \infty$ in this topology. That is, for each $\varphi \in \mathcal{C}_0$,

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi(x)\mu_{\varepsilon_k}(x)dx = \int_{\mathbb{R}^n} \varphi(x)d\nu(x). \tag{1.17}$$

We now claim that ν, consider as a distribution, equals μ.

Therefore, we must show that $\langle \mu, \psi\rangle = \int_{\mathbb{R}^n} \psi(x)d\nu(x)$ for all $\psi \in \mathscr{S}$. Let $\psi_\varepsilon = W(\cdot,\varepsilon) * \psi$. Then, for all $\alpha \in \mathbb{N}_0^n$, we have $\partial^\alpha \psi_\varepsilon = W(\cdot,\varepsilon) * \partial^\alpha\psi$. It follows from Theorem 1.15 that $\partial^\alpha\psi_\varepsilon(x)$ converges to $\partial^\alpha\psi(x)$ uniformly in x. Thus, $\psi_\varepsilon \to \psi$ in $\mathscr{S}$ as $\varepsilon \to 0$ and this implies that $\langle \mu, \psi_\varepsilon\rangle \to \langle \mu, \psi\rangle$. But, since $W(\cdot,\varepsilon) = RW(\cdot,\varepsilon)$,

$$\langle \mu, \psi_\varepsilon\rangle = \langle \mu, W(\cdot,\varepsilon) * \psi\rangle = \langle \mu * W(\cdot,\varepsilon), \psi\rangle = \int_{\mathbb{R}^n} \mu_\varepsilon(x)\psi(x)dx.$$

Thus, putting $\varepsilon = \varepsilon_k$, letting $k \to \infty$ and applying (1.17) with $\varphi = \psi$, we obtain the desired equality $\langle \mu, \psi\rangle = \int_{\mathbb{R}^n} \psi(x)d\nu(x)$. Hence, $\mu \in \mathscr{B}$. This completes the proof. □

For L^2, we can also give a very simple characterization of these operators.

Theorem 1.62. *Let T be a bounded linear transformation mapping $L^2(\mathbb{R}^n)$ to itself. Then a necessary and sufficient condition that T commutes with translation is that there exists an $m \in L^\infty(\mathbb{R}^n)$ such that $Tf = u * f$ with $\hat{u} = m$, for all $f \in L^2(\mathbb{R}^n)$. One has then $\|T\| = \|m\|_\infty$.*

Proof. If $v \in \mathscr{S}'$ and $\psi \in \mathscr{S}$, we define their product, $v\psi$, to be the element of $\mathscr{S}'$ such that $\langle v\psi, \varphi\rangle = \langle v, \psi\varphi\rangle$ for all $\varphi \in \mathscr{S}$. With the product of a distribution with a testing function so defined we first observe that whenever $u \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$, then

$$\mathscr{F}(u * \varphi) = \hat{u}\hat{\varphi}. \tag{1.18}$$

To see this, we must show that $\langle \mathscr{F}(u * \varphi), \psi\rangle = \langle \hat{u}\hat{\varphi}, \psi\rangle$ for all $\psi \in \mathscr{S}$. It follows immediately, from (1.15), part (vi) in Proposition 1.3 and the Fourier inversion formula, that

$$\begin{aligned}\langle \mathscr{F}(u * \varphi), \psi\rangle =&\langle u * \varphi, \hat{\psi}\rangle = \langle u, R\varphi * \hat{\psi}\rangle = \langle \hat{u}, \mathscr{F}^{-1}(R\varphi * \hat{\psi})\rangle \\ =&\left\langle \hat{u}, \left(\frac{|\omega|}{2\pi}\right)^n (\mathscr{F}(R\varphi * \hat{\psi}))(-\xi)\right\rangle \\ =&\left\langle \hat{u}, \left(\frac{|\omega|}{2\pi}\right)^n (\mathscr{F}(R\varphi))(-\xi)(\mathscr{F}\hat{\psi})(-\xi)\right\rangle = \langle \hat{u}, \hat{\varphi}(\xi)\psi(\xi)\rangle \\ =&\langle \hat{u}\hat{\varphi}, \psi\rangle.\end{aligned}$$

Thus, (1.18) is established.

Now, we prove the necessariness. Suppose that T commutes with translations and $\|Tf\|_2 \leqslant \|T\|\|f\|_2$ for all $f \in L^2(\mathbb{R}^n)$. Then, by Theorem 1.59, there exists a unique tempered distribution u such that $Tf = u * f$ for all $f \in \mathscr{S}$. The remainder is to prove $\hat{u} \in L^\infty(\mathbb{R}^n)$.

Let $\varphi_0 = e^{-\frac{|\omega|}{2}|x|^2}$, then, we have $\varphi_0 \in \mathscr{S}$ and $\hat{\varphi}_0 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \varphi_0$ by Theorem 1.10 with $a = 1/2|\omega|$. Thus, $T\varphi_0 = u * \varphi_0 \in L^2$ and therefore $\Phi_0 := \mathscr{F}(u * \varphi_0) = \hat{u}\hat{\varphi}_0 \in L^2$ by (1.18) and the Plancherel theorem. Let $m(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} e^{\frac{|\omega|}{2}|\xi|^2}\Phi_0(\xi) = \Phi_0(\xi)/\hat{\varphi}_0(\xi)$.

We claim that

$$\mathscr{F}(u * \varphi) = m\hat{\varphi} \tag{1.19}$$

for all $\varphi \in \mathscr{S}$. By (1.18), it suffices to show that $\langle \hat{u}\hat{\varphi}, \psi\rangle = \langle m\hat{\varphi}, \psi\rangle$ for all $\psi \in \mathscr{D}$ since $\mathscr{D}$ is dense in $\mathscr{S}$. But, if $\psi \in \mathscr{D}$, then $(\psi/\hat{\varphi}_0)(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \psi(\xi)e^{\frac{|\omega|}{2}|\xi|^2} \in \mathscr{D}$; thus,

$$\begin{aligned}\langle \hat{u}\hat{\varphi}, \psi\rangle =&\langle \hat{u}, \hat{\varphi}\psi\rangle = \langle \hat{u}, \hat{\varphi}\hat{\varphi}_0\psi/\hat{\varphi}_0\rangle = \langle \hat{u}\hat{\varphi}_0, \hat{\varphi}\psi/\hat{\varphi}_0\rangle \\ =&\int_{\mathbb{R}^n} \Phi_0(\xi)\hat{\varphi}(\xi)\left(\frac{|\omega|}{2\pi}\right)^{n/2} \psi(\xi)e^{\frac{|\omega|}{2}|\xi|^2}d\xi \\ =&\int_{\mathbb{R}^n} m(\xi)\hat{\varphi}(\xi)\psi(\xi)d\xi = \langle m\hat{\varphi}, \psi\rangle.\end{aligned}$$

It follows immediately that $\hat{u} = m$: We have just shown that $\langle \hat{u}, \hat{\varphi}\psi\rangle = \langle m\hat{\varphi}, \psi\rangle = \langle m, \hat{\varphi}\psi\rangle$ for all $\varphi \in \mathscr{S}$ and $\psi \in \mathscr{D}$. Selecting φ such that $\hat{\varphi}(\xi) = 1$ for $\xi \in \operatorname{supp}\psi$, this shows that $\langle \hat{u}, \psi\rangle = \langle m, \psi\rangle$ for all $\psi \in \mathscr{D}$. Thus, $\hat{u} = m$.

Due to

$$\|m\hat{\varphi}\|_2 = \|\mathscr{F}(u * \varphi)\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\|u * \varphi\|_2 \leqslant \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\|T\|\|\varphi\|_2 = \|T\|\|\hat{\varphi}\|_2$$

for all $\varphi \in \mathscr{S}$, it follows that

$$\int_{\mathbb{R}^n} \left(\|T\|^2 - |m|^2\right)|\hat{\varphi}|^2 d\xi \geqslant 0,$$

for all $\varphi \in \mathscr{S}$. This implies that $\|T\|^2 - |m|^2 \geqslant 0$ for almost all $x \in \mathbb{R}^n$. Hence, $m \in L^\infty(\mathbb{R}^n)$ and $\|m\|_\infty \leqslant \|T\|$.

Finally, we can show the sufficiency easily. If $\hat{u} = m \in L^\infty(\mathbb{R}^n)$, the Plancherel theorem and (1.18) immediately imply that

$$\|Tf\|_2 = \|u * f\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\|m\hat{f}\|_2 \leqslant \|m\|_\infty\|f\|_2$$

which yields $\|T\| \leqslant \|m\|_\infty$.

Thus, if $m = \hat{u} \in L^\infty$, then $\|T\| = \|m\|_\infty$. □

For further results, one can see [SW71, p.30] and [Gra04, p.137-140].

Chapter 2
Interpolation of Operators

2.1 Riesz-Thorin's and Stein's interpolation theorems

We first present a notion that is central to complex analysis, that is, the holomorphic or analytic function.

Let Ω be an open set in $\mathbb{C}$ and f a complex-valued function on Ω. The function f is *holomorphic at the point* $z_0 \in \Omega$ if the quotient

$$\frac{f(z_0+h)-f(z_0)}{h} \tag{2.1}$$

converges to a limit when $h \to 0$. Here $h \in \mathbb{C}$ and $h \neq 0$ with $z_0 + h \in \Omega$, so that the quotient is well defined. The limit of the quotient, when it exists, is denoted by $f'(z_0)$, and is called the *derivative of* f *at* z_0:

$$f'(z_0) = \lim_{h\to 0} \frac{f(z_0+h)-f(z_0)}{h}.$$

It should be emphasized that in the above limit, h is a complex number that may approach 0 from any directions.

The function f is said to be *holomorphic on* Ω if f is holomorphic at every point of Ω. If C is a closed subset of $\mathbb{C}$, we say that f is *holomorphic on* C if f is holomorphic in some open set containing C. Finally, if f is holomorphic in all of $\mathbb{C}$ we say that f is *entire*.

Every holomorphic function is analytic, in the sense that it has a power series expansion near every point, and for this reason we also use the term *analytic* as a synonym for holomorphic. For more details, one can see [SS03, pp.8-10].

Example 2.1. The function $f(z) = z$ is holomorphic on any open set in $\mathbb{C}$, and $f'(z) = 1$.

Example 2.2. The function $1/z$ is holomorphic on any open set in $\mathbb{C}$ that does not contain the origin, and $f'(z) = -1/z^2$.

Example 2.3. The function $f(z) = \bar{z}$ is not holomorphic. Indeed, we have

$$\frac{f(z_0+h)-f(z_0)}{h} = \frac{\bar{h}}{h}$$

which has no limit as $h \to 0$, as one can see by first taking h real and then h purely imaginary.

The next result pertains to the size of a holomorphic function.

Theorem 2.4 (Maximum modulus principle)**.** *Suppose that Ω is a region with compact closure $\bar{\Omega}$. If f is holomorphic on Ω and continuous on $\bar{\Omega}$, then*

$$\sup_{z\in\Omega} |f(z)| \leqslant \sup_{z\in\bar{\Omega}\setminus\Omega} |f(z)|.$$

Proof. See [SS03, p.92]. □

For convenience, let $S = \{z \in \mathbb{C} : 0 \leqslant \Re z \leqslant 1\}$ be the closed strip, $S^\circ = \{z \in \mathbb{C} : 0 < \Re z < 1\}$ be the open strip, and $\partial S = \{z \in \mathbb{C} : \Re z \in \{0, 1\}\}$.

Theorem 2.5 (Phragmen-Lindelöf theorem/Maximum principle)**.** *Assume that $f(z)$ is analytic on S° and bounded and continuous on S. Then*

$$\sup_{z\in S} |f(z)| \leqslant \max\left(\sup_{t\in\mathbb{R}} |f(it)|,\ \sup_{t\in\mathbb{R}} |f(1+it)|\right).$$

Proof. Assume that $f(z) \to 0$ as $|\Im z| \to \infty$. Consider the mapping $h : S \to \mathbb{C}$ defined by

$$h(z) = \frac{e^{i\pi z} - i}{e^{i\pi z} + i}, \quad z \in S. \tag{2.2}$$

Then h is a bijective mapping from S onto $U = \{z \in \mathbb{C} : |z| \leqslant 1\} \setminus \{\pm 1\}$, that is analytic in S° and maps ∂S onto $\{|z| = 1\} \setminus \{\pm 1\}$. Therefore, $g(z) := f(h^{-1}(z))$ is bounded and continuous on U and analytic in the interior U°. Moreover, because of $\lim_{|\Im z|\to\infty} f(z) = 0$, $\lim_{z\to\pm 1} g(z) = 0$ and we can extend g to a continuous function on $\{z \in \mathbb{C} : |z| \leqslant 1\}$. Hence, by the maximum modulus principle (Theorem 2.4), we have

$$|g(z)| \leqslant \max_{|\omega|=1} |g(\omega)| = \max\left(\sup_{t\in\mathbb{R}} |f(it)|,\ \sup_{t\in\mathbb{R}} |f(1+it)|\right),$$

which implies the statement in this case.

Next, if f is a general function as in the assumption, then we consider

$$f_{\delta,z_0}(z) = e^{\delta(z-z_0)^2} f(z), \quad \delta > 0,\ z_0 \in S^\circ.$$

Since $|e^{\delta(z-z_0)^2}| \leqslant e^{\delta(x^2-y^2)}$ with $z - z_0 = x + iy$, $-1 \leqslant x \leqslant 1$ and $y \in \mathbb{R}$, we have $f_{\delta,z_0}(z) \to 0$ as $|\Im z| \to \infty$. Therefore

$$\begin{aligned}|f(z_0)| =&|f_{\delta,z_0}(z_0)| \leqslant \max\left(\sup_{t\in\mathbb{R}} |f_{\delta,z_0}(it)|,\ \sup_{t\in\mathbb{R}} |f_{\delta,z_0}(1+it)|\right)\\ \leqslant& e^{\delta} \max\left(\sup_{t\in\mathbb{R}} |f(it)|,\ \sup_{t\in\mathbb{R}} |f(1+it)|\right).\end{aligned}$$

Passing to the limit $\delta \to 0$, we obtain the desired result since $z_0 \in S$ is arbitrary. □

As a corollary we obtain the following three lines theorem, which is the basis for the proof of the Riesz-Thorin interpolation theorem and the complex interpolation method.

Theorem 2.6 (Hadamard three lines theorem)**.** *Assume that $f(z)$ is analytic on S° and bounded and continuous on S. Then*

$$\sup_{t\in\mathbb{R}} |f(\theta+it)| \leqslant \left(\sup_{t\in\mathbb{R}} |f(it)|\right)^{1-\theta} \left(\sup_{t\in\mathbb{R}} |f(1+it)|\right)^{\theta},$$

for every $\theta \in [0, 1]$.

Proof. Denote

$$A_0 := \sup_{t\in\mathbb{R}} |f(it)|, \quad A_1 := \sup_{t\in\mathbb{R}} |f(1+it)|.$$

Let $\lambda \in \mathbb{R}$ and define

$$F_\lambda(z) = e^{\lambda z} f(z).$$

Then by Theorem 2.5, it follows that

$$|F_\lambda(z)| \leqslant \max(A_0, e^\lambda A_1).$$

Hence,

$$|f(\theta+it)| \leqslant e^{-\lambda\theta} \max(A_0, e^\lambda A_1)$$

for all $t \in \mathbb{R}$. Choosing $\lambda = \ln \frac{A_0}{A_1}$ such that $e^\lambda A_1 = A_0$, we complete the proof. □

In order to state the Riesz-Thorin theorem in a general version, we will state and prove it in measurable spaces instead of $\mathbb{R}^n$ only.

Let (X, μ) be a measure space, μ always being a positive measure. We adopt the usual convention that two functions are considered equal if they agree except on a set of μ-measure zero. Then we denote by $L^p(X, d\mu)$ (or simply $L^p(d\mu)$, $L^p(X)$ or even L^p) the Lebesgue-space of (all equivalence classes of) scalar-valued μ-measurable functions f on X, such that

$$\|f\|_p = \left(\int_X |f(x)|^p d\mu \right)^{1/p}$$

is finite. Here we have $1 \leqslant p < \infty$. In the limiting case, $p = \infty$, L^p consists of all μ-measurable and bounded functions. Then we write

$$\|f\|_\infty = \sup_X |f(x)|.$$

In this section, scalars are supposed to be complex numbers.

Let T be a linear mapping from $L^p = L^p(X, d\mu)$ to $L^q(Y, d\nu)$. This means that $T(\alpha f + \beta g) = \alpha T(f) + \beta T(g)$. We shall write

$$T : L^p \to L^q$$

if in addition T is bounded, i.e., if

$$A = \sup_{f \neq 0} \frac{\|Tf\|_q}{\|f\|_p}$$

is finite. The number A is called the norm of the mapping T.

It will also be necessary to treat operators T defined on several L^p spaces simultaneously.

Definition 2.7. We define $L^{p_1} + L^{p_2}$ to be the space of all functions f, such that $f = f_1 + f_2$, with $f_1 \in L^{p_1}$ and $f_2 \in L^{p_2}$.

Suppose now $p_1 < p_2$. Then we observe that

$$L^p \subset L^{p_1} + L^{p_2}, \quad \forall p \in [p_1, p_2].$$

In fact, let $f \in L^p$ and let γ be a fixed positive constant. Set

$$f_1(x) = \begin{cases} f(x), & |f(x)| > \gamma, \\ 0, & |f(x)| \leqslant \gamma, \end{cases}$$

and $f_2(x) = f(x) - f_1(x)$. Then

$$\int |f_1(x)|^{p_1} dx = \int |f_1(x)|^p |f_1(x)|^{p_1 - p} dx \leqslant \gamma^{p_1 - p} \int |f(x)|^p dx,$$

since $p_1 - p \leqslant 0$. Similarly,

$$\int |f_2(x)|^{p_2} dx = \int |f_2(x)|^p |f_2(x)|^{p_2 - p} dx \leqslant \gamma^{p_2 - p} \int |f(x)|^p dx,$$

so $f_1 \in L^{p_1}$ and $f_2 \in L^{p_2}$, with $f = f_1 + f_2$.

Now, we have the following well-known theorem.

Theorem 2.8 (The Riesz-Thorin interpolation theorem). *Let T be a linear operator with domain $(L^{p_0} + L^{p_1})(X, d\mu)$, $p_0, p_1, q_0, q_1 \in [1, \infty]$. Assume that*

$$\|Tf\|_{L^{q_0}(Y, d\nu)} \leqslant A_0 \|f\|_{L^{p_0}(X, d\mu)}, \quad \text{if } f \in L^{p_0}(X, d\mu),$$

and

$$\|Tf\|_{L^{q_1}(Y, d\nu)} \leqslant A_1 \|f\|_{L^{p_1}(X, d\mu)}, \quad \text{if } f \in L^{p_1}(X, d\mu),$$

for some $p_0 \neq p_1$ and $q_0 \neq q_1$. Suppose that for a certain $0 < \theta < 1$

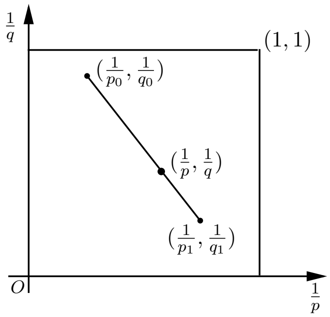
$$\frac{1}{p}=\frac{1-\theta}{p_0}+\frac{\theta}{p_1},\qquad \frac{1}{q}=\frac{1-\theta}{q_0}+\frac{\theta}{q_1}. \tag{2.3}$$

Then

$$\|Tf\|_{L^q(Y,d\nu)}\leqslant A_\theta\|f\|_{L^p(X,d\mu)},\quad \text{if } f\in L^p(X,d\mu),$$

with

$$A_\theta\leqslant A_0^{1-\theta}A_1^\theta. \tag{2.4}$$



Remark 2.9. 1) (2.4) means that A_θ is logarithmically convex, i.e., $\ln A_\theta$ is convex.
2) The geometrical meaning of (2.3) is that the points $(1/p,1/q)$ are the points on the line segment between $(1/p_0,1/q_0)$ and $(1/p_1,1/q_1)$.
3) The original proof of this theorem, published in 1926 by Marcel Riesz, was a long and difficult calculation. Riesz' student G. Olof Thorin subsequently discovered a far more elegant proof and published it in 1939, which contains the idea behind the complex interpolation method.

Proof. Denote

$$\langle h,g\rangle=\int_Y h(y)g(y)d\nu(y)$$

and $1/q'=1-1/q$. Then we have, by Hölder inequality,

$$\|h\|_q=\sup_{\|g\|_{q'}=1}|\langle h,g\rangle|,\text{ and } A_\theta=\sup_{\|f\|_p=\|g\|_{q'}=1}|\langle Tf,g\rangle|.$$

Noticing that $\mathcal{C}_c(X)$ is dense in $L^p(X,\mu)$ for $1\leqslant p<\infty$, we can assume that f and g are bounded with compact supports since $p,\,q'<\infty$.[1] Thus, we have $|f(x)|\leqslant M<\infty$ for all $x\in X$, and $\operatorname{supp} f=\overline{\{x\in X: f(x)\neq 0\}}$ is compact, i.e., $\mu(\operatorname{supp} f)<\infty$ which implies $\int_X|f(x)|^\ell d\mu(x)=\int_{\operatorname{supp} f}|f(x)|^\ell d\mu(x)\leqslant M^\ell\mu(\operatorname{supp} f)<\infty$ for any $\ell>0$. So g does.

For $0\leqslant \Re z\leqslant 1$, we put

$$\frac{1}{p(z)}=\frac{1-z}{p_0}+\frac{z}{p_1},\qquad \frac{1}{q'(z)}=\frac{1-z}{q_0'}+\frac{z}{q_1'},$$

and

$$\begin{aligned}\eta(z)=&\eta(x,z)=|f(x)|^{\frac{p}{p(z)}}\frac{f(x)}{|f(x)|},\ x\in X;\\ \zeta(z)=&\zeta(y,z)=|g(y)|^{\frac{q'}{q'(z)}}\frac{g(y)}{|g(y)|},\ y\in Y.\end{aligned}$$

Now, we prove $\eta(z),\,\eta'(z)\in L^{p_j}$ for $j=0,1$. Indeed, we have

$$\begin{aligned}|\eta(z)|=&\left||f(x)|^{\frac{p}{p(z)}}\right|=\left||f(x)|^{p(\frac{1-z}{p_0}+\frac{z}{p_1})}\right|=\left||f(x)|^{p(\frac{1-\Re z}{p_0}+\frac{\Re z}{p_1})+ip(\frac{\Im z}{p_1}-\frac{\Im z}{p_0})}\right|\\ =&|f(x)|^{p(\frac{1-\Re z}{p_0}+\frac{\Re z}{p_1})}=|f(x)|^{\frac{p}{p(\Re z)}}.\end{aligned}$$

Thus,

$$\|\eta(z)\|_{p_j}^{p_j}=\int_X|\eta(x,z)|^{p_j}d\mu(x)=\int_X|f(x)|^{\frac{pp_j}{p(\Re z)}}d\mu(x)<\infty.$$

We have

[1] Otherwise, it will be $p_0=p_1=\infty$ if $p=\infty$, or $\theta=\frac{1-1/q_0}{1/q_1-1/q_0}\geqslant 1$ if $q'=\infty$.

$$\eta'(z) = |f(x)|^{\frac{p}{p(z)}} \left[\frac{p}{p(z)}\right]' \frac{f(x)}{|f(x)|} \ln|f(x)| = p\left(\frac{1}{p_1} - \frac{1}{p_0}\right) |f(x)|^{\frac{p}{p(z)}} \frac{f(x)}{|f(x)|} \ln|f(x)|.$$

On one hand, we have $\lim_{|f(x)|\to 0_+} |f(x)|^\alpha \ln|f(x)| = 0$ for any $\alpha > 0$, that is, $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t. $||f(x)|^\alpha \ln|f(x)|| < \varepsilon$ if $|f(x)| < \delta$. On the other hand, if $|f(x)| > \delta$, then we have $||f(x)|^\alpha \ln|f(x)|| \leqslant M^\alpha |\ln|f(x)|| \leqslant M^\alpha \max(|\ln M|, |\ln\delta|) < \infty$. Thus, $||f(x)|^\alpha \ln|f(x)|| \leqslant C$. Hence,

$$\begin{aligned} |\eta'(z)| =& p\left|\frac{1}{p_1} - \frac{1}{p_0}\right| \left||f(x)|^{\frac{p}{p(z)}-\alpha}\right| |f(x)|^\alpha |\ln|f(x)|| \\ \leqslant& C\left||f(x)|^{\frac{p}{p(z)}-\alpha}\right| = C|f(x)|^{\frac{p}{p(\Re z)}-\alpha}, \end{aligned}$$

which yields

$$\|\eta'(z)\|_{p_j}^{p_j} \leqslant C\int_X |f(x)|^{(\frac{p}{p(\Re z)}-\alpha)p_j} d\mu(x) < \infty.$$

Therefore, $\eta(z)$, $\eta'(z) \in L^{p_j}$ for $j = 0, 1$. So $\zeta(z)$, $\zeta'(z) \in L^{q_j'}$ for $j = 0, 1$ in the same way. It follows that $T\eta(z) \in L^{q_j}$, and $(T\eta)'(z) \in L^{q_j}$ with $0 < \Re z < 1$, for $j = 0, 1$. This implies the existence of

$$F(z) = \langle T\eta(z), \zeta(z)\rangle, \quad 0 \leqslant \Re z \leqslant 1.$$

Since

$$\begin{aligned} \frac{dF(z)}{dz} =& \frac{d}{dz}\langle T\eta(z), \zeta(z)\rangle = \frac{d}{dz}\int_Y (T\eta)(y,z)\zeta(y,z)d\nu(y) \\ =& \int_Y (T\eta)_z(y,z)\zeta(y,z)d\nu(y) + \int_Y (T\eta)(y,z)\zeta_z(y,z)d\nu(y) \\ =& \langle (T\eta)'(z), \zeta(z)\rangle + \langle T\eta(z), \zeta'(z)\rangle < \infty, \end{aligned}$$

$F(z)$ is analytic on the open strip $0 < \Re z < 1$. Moreover it is easy to see that $F(z)$ is bounded and continuous on the closed strip $0 \leqslant \Re z \leqslant 1$.

Next, we note that for $j = 0, 1$

$$\|\eta(j + it)\|_{p_j} = \|f\|_p^{\frac{p}{p_j}} = 1.$$

Similarly, we also have $\|\zeta(j + it)\|_{q_j'} = 1$ for $j = 0, 1$. Thus, for $j = 0, 1$

$$\begin{aligned} |F(j + it)| =& |\langle T\eta(j + it), \zeta(j + it)\rangle| \leqslant \|T\eta(j + it)\|_{q_j} \|\zeta(j + it)\|_{q_j'} \\ \leqslant& A_j \|\eta(j + it)\|_{p_j} \|\zeta(j + it)\|_{q_j'} = A_j. \end{aligned}$$

Using Hadamard three line theorem, reproduced as Theorem 2.6, we get the conclusion

$$|F(\theta + it)| \leqslant A_0^{1-\theta} A_1^\theta, \quad \forall t \in \mathbb{R}.$$

Taking $t = 0$, we have $|F(\theta)| \leqslant A_0^{1-\theta} A_1^\theta$. We also note that $\eta(\theta) = f$ and $\zeta(\theta) = g$, thus $F(\theta) = \langle Tf, g\rangle$. That is, $|\langle Tf, g\rangle| \leqslant A_0^{1-\theta} A_1^\theta$. Therefore, $A_\theta \leqslant A_0^{1-\theta} A_1^\theta$. □

Now, we shall give two rather simple applications of the Riesz-Thorin interpolation theorem.

Theorem 2.10 (Hausdorff-Young inequality)**.** *Let* $1 \leqslant p \leqslant 2$ *and* $1/p + 1/p' = 1$. *Then the Fourier transform defined as in* (1.1) *satisfies*

$$\|\mathscr{F} f\|_{p'} \leqslant \left(\frac{|\omega|}{2\pi}\right)^{-n/p'} \|f\|_p.$$

Proof. It follows by interpolation between the L^1-L^∞ result $\|\mathscr{F} f\|_\infty \leqslant \|f\|_1$ (cf. Theorem 1.5) and Plancherel's theorem $\|\mathscr{F} f\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|f\|_2$ (cf. Theorem 1.26). □

Theorem 2.11 (Young's inequality for convolutions)**.** *If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, $1 \leqslant p, q, r \leqslant \infty$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, then*

$$\|f * g\|_r \leqslant \|f\|_p \|g\|_q.$$

Proof. We fix $f \in L^p$, $p \in [1, \infty]$ and then will apply the Riesz-Thorin interpolation theorem to the mapping $g \mapsto f * g$. Our endpoints are Hölder's inequality which gives

$$|f * g(x)| \leqslant \|f\|_p \|g\|_{p'}$$

and thus $g \mapsto f * g$ maps $L^{p'}(\mathbb{R}^n)$ to $L^\infty(\mathbb{R}^n)$ and the simpler version of Young's inequality (proved by Minkowski's inequality) which tells us that if $g \in L^1$, then

$$\|f * g\|_p \leqslant \|f\|_p \|g\|_1.$$

Thus $g \mapsto f * g$ also maps L^1 to L^p. Thus, this map also takes L^q to L^r where

$$\frac{1}{q} = \frac{1-\theta}{1} + \frac{\theta}{p'}, \text{ and } \frac{1}{r} = \frac{1-\theta}{p} + \frac{\theta}{\infty}.$$

Eliminating θ, we have $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$.

The condition $q \geqslant 1$ is equivalent with $\theta \geqslant 0$ and $r \geqslant 1$ is equivalent with the condition $\theta \leqslant 1$. Thus, we obtain the stated inequality for precisely the exponents p, q and r in the hypothesis. □

Remark 2.12. The sharp form of Young's inequality for convolutions can be found in [Bec75, Theorem 3], we just state it as follows. Under the assumption of Theorem 2.11, we have

$$\|f * g\|_r \leqslant (A_p A_q A_{r'})^n \|f\|_p \|g\|_q,$$

where $A_m = (m^{1/m}/m'^{1/m'})^{1/2}$ for $m \in (1, \infty)$, $A_1 = A_\infty = 1$ and primes always denote dual exponents, $1/m + 1/m' = 1$.

The Riesz-Thorin interpolation theorem can be extended to the case where the interpolated operators allowed to vary. In particular, if a family of operators depends analytically on a parameter z, then the proof of this theorem can be adapted to work in this setting.

We now describe the setup for this theorem. Suppose that for every z in the closed strip S there is an associated linear operator T_z defined on the space of simple functions on X and taking values in the space of measurable functions on Y such that

$$\int_Y |T_z(f)g| d\nu < \infty \tag{2.5}$$

whenever f and g are simple functions on X and Y, respectively. The family $\{T_z\}_z$ is said to be *analytic* if the function

$$z \to \int_Y T_z(f) g d\nu \tag{2.6}$$

is analytic in the open strip S° and continuous on its closure S. Finally, the analytic family is of *admissible growth* if there is a constant $0 < a < \pi$ and a constant $C_{f,g}$ such that

$$e^{-a|\Im z|} \ln \left| \int_Y T_z(f) g d\nu \right| \leqslant C_{f,g} < \infty \tag{2.7}$$

for all $z \in S$. The extension of the Riesz-Thorin interpolation theorem is now stated.

Theorem 2.13 (Stein interpolation theorem)**.** *Let T_z be an analytic family of linear operators of admissible growth. Let $1 \leqslant p_0, p_1, q_0, q_1 \leqslant \infty$ and suppose that M_0 and M_1 are real-valued functions such that*

$$\sup_{t \in \mathbb{R}} e^{-b|t|} \ln M_j(t) < \infty \tag{2.8}$$

for $j = 0, 1$ and some $0 < b < \pi$. Let $0 < \theta < 1$ satisfy

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \text{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{2.9}$$

Suppose that

$$\|T_{it}(f)\|_{q_0} \leqslant M_0(t)\|f\|_{p_0}, \quad \|T_{1+it}(f)\|_{q_1} \leqslant M_1(t)\|f\|_{p_1} \tag{2.10}$$

for all simple functions f on X. Then

$$\|T_\theta(f)\|_q \leqslant M(\theta)\|f\|_p, \quad \text{when } 0<\theta<1 \tag{2.11}$$

for all simple functions f on X, where

$$M(\theta) = \exp\left\{\frac{\sin \pi\theta}{2}\int_{\mathbb{R}}\left[\frac{\ln M_0(t)}{\cosh \pi t - \cos \pi\theta} + \frac{\ln M_1(t)}{\cosh \pi t + \cos \pi\theta}\right]dt\right\}.$$

By density, T_θ has a unique extension as a bounded operator from $L^p(X,\mu)$ into $L^q(Y,\nu)$ for all p and q as in (2.9).

The proof of the Stein interpolation theorem can be obtained from that of the Riesz-Thorin theorem simply "by adding a single letter of the alphabet". Indeed, the way the Riesz-Thorin theorem is proven is to study an expression of the form

$$F(z) = \langle T\eta(z), \zeta(z)\rangle,$$

the Stein interpolation theorem proceeds by instead studying the expression

$$F(z) = \langle T_z\eta(z), \zeta(z)\rangle.$$

One can then repeat the proof of the Riesz-Thorin theorem more or less verbatim to obtain the Stein interpolation theorem. Of course, the explicit expression of $M(\theta)$ need an extension of the three lines theorem. For the detailed proof, one can see [SW71, p. 205-209] or [Gra04, p.38-42].

2.2 The distribution function and weak L^p spaces

We shall now be interested in giving a concise expression for the relative size of a function. Thus we give the following concept.

Definition 2.14. Let $f(x)$ be a measurable function on $\mathbb{R}^n$. Then the function $f_* : [0,\infty) \mapsto [0,\infty]$ defined by

$$f_*(\alpha) = \mathfrak{m}(\{x : |f(x)| > \alpha\})$$

is called to be the *distribution function* of f.

The distribution function f_* provides information about the size of f but not about the behavior of f itself near any given point. For instance, a function on Rn and each of its translates have the same distribution function.

In particular, the decrease of $f_*(\alpha)$ as α grows describes the relative largeness of the function; this is the main concern locally. The increase of $f_*(\alpha)$ as α tends to zero describes the relative smallness of the function "at infinity"; this is its importance globally, and is of no interest if, for example, the function is supported on a bounded set.

Now, we give some properties of distribution functions.

Proposition 2.15. *For the distribution function, we have following fundamental properties.*

(i) *$f_*(\alpha)$ is decreasing and continuous on the right.*

(ii) *If $|f(x)| \leqslant |g(x)|$, then $f_*(\alpha) \leqslant g_*(\alpha)$.*

(iii) *If* $|f(x)| \leqslant \liminf_{k\to\infty} |f_k(x)|$ *for a.e.* x*, then* $f_*(\alpha) \leqslant \liminf_{k\to\infty}(f_k)_*(\alpha)$ *for any* $\alpha \geqslant 0$.

(iv) *If* $|f(x)| \leqslant |g(x)| + |h(x)|$*, then* $f_*(\alpha_1+\alpha_2) \leqslant g_*(\alpha_1) + h_*(\alpha_2)$ *for any* $\alpha_1, \alpha_2 \geqslant 0$.

(v) $(fg)_*(\alpha_1\alpha_2) \leqslant f_*(\alpha_1) + g_*(\alpha_2)$ *for any* $\alpha_1, \alpha_2 \geqslant 0$.

(vi) *For any* $p \in (0,\infty)$ *and* $\alpha > 0$*, it holds* $f_*(\alpha) \leqslant \alpha^{-p} \int_{\{x:|f(x)|>\alpha\}} |f(x)|^p dx$.

(vii) *If* $f \in L^p$*,* $p \in [1,\infty)$*, then* $\lim_{\alpha\to+\infty} \alpha^p f_*(\alpha) = 0 = \lim_{\alpha\to 0} \alpha^p f_*(\alpha)$.

(viii) *If* $\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha < \infty$*,* $p \in [1,\infty)$*, then* $\alpha^p f_*(\alpha) \to 0$ *as* $\alpha \to +\infty$ *and* $\alpha \to 0$*, respectively.*

Proof. For simplicity, denote $E_f(\alpha) = \{x : |f(x)| > \alpha\}$ for $\alpha > 0$.

(i) Let $\{\alpha_k\}$ is a decreasing positive sequence which tends to α, then we have $E_f(\alpha) = \cup_{k=1}^\infty E_f(\alpha_k)$. Since $\{E_f(\alpha_k)\}$ is a increasing sequence of sets, it follows $\lim_{k\to\infty} f_*(\alpha_k) = f_*(\alpha)$. This implies the continuity of $f_*(\alpha)$ on the right.

(iii) Let $E = \{x : |f(x)| > \alpha\}$ and $E_k = \{x : |f_k(x)| > \alpha\}$, $k \in \mathbb{N}$. By the assumption and the definition of inferior limit, i.e.,

$$|f(x)| \leqslant \liminf_{k\to\infty} |f_k(x)| = \sup_{\ell\in\mathbb{N}} \inf_{k>\ell} |f_k(x)|,$$

for $x \in E$, there exists an integer M such that for all $k > M$, $|f_k(x)| > \alpha$. Thus, $E \subset \bigcup_{M=1}^\infty \bigcap_{k=M}^\infty E_k$, and for any $\ell \geqslant 1$,

$$m\left(\bigcap_{k=\ell}^\infty E_k\right) \leqslant \inf_{k\geqslant\ell} \mathrm{m}(E_k) \leqslant \sup_\ell \inf_{k\geqslant\ell} \mathrm{m}(E_k) = \liminf_{k\to\infty} \mathrm{m}(E_k).$$

Since $\{\bigcap_{k=M}^\infty E_k\}_{M=1}^\infty$ is an increasing sequence of sets, we obtain

$$f_*(\alpha) = \mathrm{m}(E) \leqslant m\left(\bigcup_{M=1}^\infty \bigcap_{k=M}^\infty E_k\right) = \lim_{M\to\infty} m\left(\bigcap_{k=M}^\infty E_k\right) \leqslant \liminf_{k\to\infty}(f_k)_*(\alpha).$$

(v) Noticing that $\{x : |f(x)g(x)| > \alpha_1\alpha_2\} \subset \{x : |f(x)| > \alpha_1\} \cup \{x : |g(x)| > \alpha_2\}$, we have the desired result.

(vi) $f_*(\alpha) = \mathrm{m}(\{x : |f(x)| > \alpha\}) = \int_{\{x:|f(x)|>\alpha\}} dx \leqslant \int_{\{x:|f(x)|>\alpha\}} (\frac{|f(x)|}{\alpha})^p dx$ $= \alpha^{-p} \int_{\{x:|f(x)|>\alpha\}} |f(x)|^p dx$.

(vii) From (vi), it follows $\alpha^p f_*(\alpha) \leqslant \int_{\{x:|f(x)|>\alpha\}} |f(x)|^p dx \leqslant \int_{\mathbb{R}^n} |f(x)|^p dx$. Thus, $\mathrm{m}(\{x : |f(x)| > \alpha\}) \to 0$ as $\alpha \to +\infty$ and

$$\lim_{\alpha\to+\infty} \int_{\{x:|f(x)|>\alpha\}} |f(x)|^p dx = 0.$$

Hence, $\alpha^p f_*(\alpha) \to 0$ as $\alpha \to +\infty$ since $\alpha^p f_*(\alpha) \geqslant 0$.

For any $0 < \alpha < \beta$, we have, by noticing that $1 \leqslant p < \infty$, that

$$\begin{aligned}\lim_{\alpha\to 0} \alpha^p f_*(\alpha) = \lim_{\alpha\to 0} \alpha^p(f_*(\alpha) - f_*(\beta)) &= \lim_{\alpha\to 0} \alpha^p \mathrm{m}(\{x : \alpha < |f(x)| \leqslant \beta\}) \\ &\leqslant \int_{\{x:|f(x)|\leqslant\beta\}} |f(x)|^p dx.\end{aligned}$$

By the arbitrariness of β, it follows $\alpha^p f_*(\alpha) \to 0$ as $\alpha \to 0$.

(viii) Since $\int_{\alpha/2}^\alpha (t^p)' dt = \alpha^p - (\alpha/2)^p$ and $f_*(\alpha) \leqslant f_*(t)$ for $t \leqslant \alpha$, we have

$$f_*(\alpha)\alpha^p(1 - 2^{-p}) \leqslant p \int_{\alpha/2}^\alpha t^{p-1} f_*(t) dt$$

which implies the desired result.

For other ones, they are easy to verify. □

From this proposition, we can prove the following equivalent norm of L^p spaces.

Theorem 2.16 (The equivalent norm of L^p). *Let $f(x)$ be a measurable function in $\mathbb{R}^n$, then*

i) $\|f\|_p = \left(p\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha\right)^{1/p}$, *if* $1 \leqslant p < \infty$,

ii) $\|f\|_\infty = \inf\{\alpha : f_*(\alpha) = 0\}$.

Proof. In order to prove i), we first prove the following conclusion: If $f(x)$ is finite and $f_*(\alpha) < \infty$ for any $\alpha > 0$, then

$$\int_{\mathbb{R}^n} |f(x)|^p dx = -\int_0^\infty \alpha^p df_*(\alpha). \tag{2.12}$$

Indeed, the r.h.s. of the equality is well-defined from the conditions. For the integral in the l.h.s., we can split it into Lebesgue integral summation. Let $0 < \varepsilon < 2\varepsilon < \cdots < k\varepsilon < \cdots$ and

$$E_j = \{x \in \mathbb{R}^n : (j-1)\varepsilon < |f(x)| \leqslant j\varepsilon\}, \quad j = 1, 2, \cdots,$$

then, $\mathfrak{m}(E_j) = f_*((j-1)\varepsilon) - f_*(j\varepsilon)$, and

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x)|^p dx &= \lim_{\varepsilon\to 0}\sum_{j=1}^\infty (j\varepsilon)^p \mathfrak{m}(E_j) = -\lim_{\varepsilon\to 0}\sum_{j=1}^\infty (j\varepsilon)^p[f_*(j\varepsilon) - f_*((j-1)\varepsilon)] \\ &= -\int_0^\infty \alpha^p df_*(\alpha).\end{aligned}$$

Now we return to prove i). If the values of both sides are infinite, then it is clearly true. If one of the integral is finite, then it is clear that $f_*(\alpha) < +\infty$ and $f(x)$ is finite almost everywhere. Thus (2.12) is valid.

If either $f \in L^p(\mathbb{R}^n)$ or $\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha < \infty$ for $1 \leqslant p < \infty$, then we always have $\alpha^p f_*(\alpha) \to 0$ as $\alpha \to +\infty$ and $\alpha \to 0$ from the property (vii) and (viii) in Proposition 2.15.

Therefore, integrating by part, we have

$$-\int_0^\infty \alpha^p df_*(\alpha) = p\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha - \alpha^p f_*(\alpha)|_0^{+\infty} = p\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha.$$

Thus, i) is true.

For ii), we have

$$\begin{aligned}\inf\{\alpha : f_*(\alpha) = 0\} &= \inf\{\alpha : \mathfrak{m}(\{x : |f(x)| > \alpha\}) = 0\} \\ &= \inf\{\alpha : |f(x)| \leqslant \alpha, \ a.e.\} = \operatorname{ess\,sup}_{x\in\mathbb{R}^n}|f(x)| = \|f\|_{L^\infty}.\end{aligned}$$

We complete the proofs. □

Notice that the same argument yields the more general fact that for any increasing continuously differentiable function φ on $[0, \infty)$ with $\varphi(0) = 0$ we have

$$\int_X \varphi(|f|)d\mu = \int_0^\infty \varphi'(\alpha) f_*(\alpha)d\alpha. \tag{2.13}$$

Using the distribution function f_*, we now introduce the weak L^p-spaces denoted by L^p_*.

Definition 2.17. The space L^p_*, $1 \leqslant p < \infty$, consists of all f such that

$$\|f\|_{L^p_*} = \sup_\alpha \alpha f_*^{1/p}(\alpha) < \infty.$$

In the limiting case $p = \infty$, we put $L^\infty_* = L^\infty$.

By the part (iv) in Proposition 2.15 and the triangle inequality of L^p norms, we have

$$\|f + g\|_{L^p_*} \leqslant 2^{1/p}(\|f\|_{L^p_*} + \|g\|_{L^p_*}).$$

Thus, one can verify that L^p_* is a quasi-normed vector space. The weak L^p spaces are larger than the usual L^p spaces. We have the following:

Theorem 2.18. *For any $1 \leqslant p < \infty$, and any $f \in L^p$, we have $\|f\|_{L^p_*} \leqslant \|f\|_p$, hence $L^p \subset L^p_*$.*

Proof. From the part (vi) in Proposition 2.15, we have

$$\alpha f_*^{1/p}(\alpha) \leqslant \left(\int_{\{x:|f(x)|>\alpha\}} |f(x)|^p dx \right)^{1/p}$$

which yields the desired result. □

The inclusion $L^p \subset L^p_*$ is strict for $1 \leqslant p < \infty$. For example, let $h(x) = |x|^{-n/p}$. Obviously, h is not in $L^p(\mathbb{R}^n)$ but h is in $L^p_*(\mathbb{R}^n)$ and we may check easily that

$$\begin{aligned}
\|h\|_{L^p_*} =& \sup_\alpha \alpha h_*^{1/p}(\alpha) = \sup_\alpha \alpha(\mathfrak{m}(\{x : |x|^{-n/p} > \alpha\}))^{1/p} \\
=& \sup_\alpha \alpha(\mathfrak{m}(\{x : |x| < \alpha^{-p/n}\}))^{1/p} = \sup_\alpha \alpha(\alpha^{-p} V_n)^{1/p} \\
=& V_n^{1/p},
\end{aligned}$$

where $V_n = \pi^{n/2}/\Gamma(1 + n/2)$ is the volume of the unit ball in $\mathbb{R}^n$ and Γ-function $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ for $\Re z > 0$.

It is not immediate from their definition that the weak L^p spaces are complete with respect to the quasi-norm $\|\cdot\|_{L^p_*}$. For the completeness, we will state it later as a special case of Lorentz spaces.

2.3 The decreasing rearrangement and Lorentz spaces

The spaces L^p_* are special cases of the more general Lorentz spaces $L^{p,q}$. In their definition, we use yet another concept, i.e., the decreasing rearrangement of functions.

Definition 2.19. If f is a measurable function on $\mathbb{R}^n$, the *decreasing rearrangement* of f is the function $f^* : [0, \infty) \mapsto [0, \infty]$ defined by

$$f^*(t) = \inf \{\alpha \geqslant 0 : f_*(\alpha) \leqslant t\},$$

where we use the convention that $\inf \varnothing = \infty$.

Now, we first give some examples of distribution function and decreasing rearrangement. The first example establish some important relations between a simple function, its distribution function and decreasing rearrangement.

Example 2.20 (Decreasing rearrangement of a simple function). Let f be a simple function of the following form

$$f(x) = \sum_{j=1}^k a_j \chi_{A_j}(x)$$

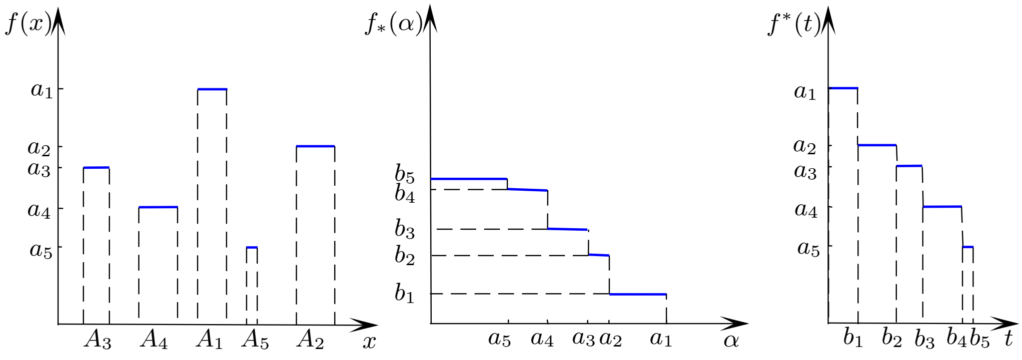
where $a_1 > a_2 > \cdots > a_k > 0$, $A_j = \{x \in \mathbb{R} : f(x) = a_j\}$ and χ_A is the characteristic function of the set A (see Figure (a)). Then

$$f_*(\alpha) = \mathfrak{m}(\{x : |f(x)| > \alpha\}) = \mathfrak{m}(\{x : \sum_{j=1}^k a_j \chi_{A_j}(x) > \alpha\}) = \sum_{j=1}^k b_j \chi_{B_j}(\alpha),$$

where $b_j = \sum_{i=1}^j \mathfrak{m}(A_i)$, $B_j = [a_{j+1}, a_j)$ for $j = 1, 2, \cdots, k$ and $a_{k+1} = 0$ which shows that the distribution function of a simple function is a simple function (see Figure (b)). We can also find the decreasing rearrangement

$$f^*(t) = \inf\{\alpha \geqslant 0 : f_*(\alpha) \leqslant t\} = \inf\{\alpha \geqslant 0 : \sum_{j=1}^{k} b_j \chi_{B_j}(\alpha) \leqslant t\} = \sum_{j=1}^{k} a_j \chi_{[b_{j-1}, b_j)}(t)$$

which is also a simple function (see Figure (c)).



(a) (b) (c)

Example 2.21. Let $f : [0, \infty) \mapsto [0, \infty)$ be

$$f(x) = \begin{cases} 1 - (x-1)^2, & 0 \leqslant x \leqslant 2, \\ 0, & x > 2. \end{cases}$$

It is clear that $f_*(\alpha) = 0$ for $\alpha > 1$ since $|f(x)| \leqslant 1$. For $\alpha \in [0, 1]$, we have

$$\begin{aligned} f_*(\alpha) =& \mathfrak{m}(\{x \in [0, \infty) : 1 - (x-1)^2 > \alpha\}) \\ =& \mathfrak{m}(\{x \in [0, \infty) : 1 - \sqrt{1-\alpha} < x < 1 + \sqrt{1-\alpha}\}) = 2\sqrt{1-\alpha}. \end{aligned}$$

That is,

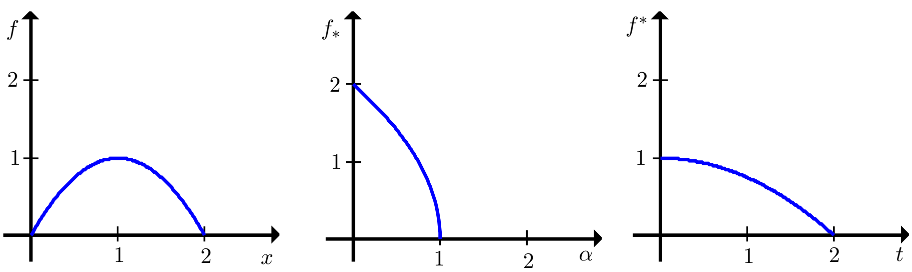
$$f_*(\alpha) = \begin{cases} 2\sqrt{1-\alpha}, & 0 \leqslant \alpha \leqslant 1, \\ 0, & \alpha > 1. \end{cases}$$

The decreasing rearrangement $f^*(t) = 0$ for $t > 2$ since $f_*(\alpha) \leqslant 2$ for any $\alpha \geqslant 0$. For $t \leqslant 2$, we have

$$\begin{aligned} f^*(t) =& \inf\{\alpha \geqslant 0 : 2\sqrt{1-\alpha} \leqslant t\} \\ =& \inf\{\alpha \geqslant 0 : \alpha \geqslant 1 - t^2/4\} = 1 - t^2/4. \end{aligned}$$

Thus,

$$f^*(t) = \begin{cases} 1 - t^2/4, & 0 \leqslant t \leqslant 2, \\ 0, & t > 2. \end{cases}$$



(a) (b) (c)

Observe that the integral over f, f_* and f^* are all the same, i.e.,

$$\int_0^\infty f(x)dx = \int_0^2 [1 - (x-1)^2]dx = \int_0^1 2\sqrt{1-\alpha}d\alpha = \int_0^2 (1 - t^2/4)dt = 4/3.$$

Example 2.22. We define an extended function $f : [0, \infty) \mapsto [0, \infty]$ as

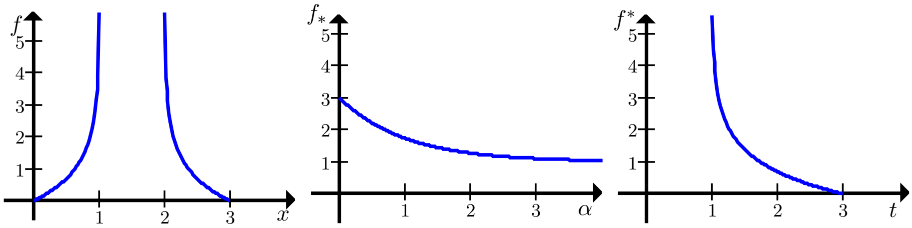
$$f(x) = \begin{cases} 0, & x = 0, \\ \ln(\frac{1}{1-x}), & 0 < x < 1, \\ \infty, & 1 \leqslant x \leqslant 2, \\ \ln(\frac{1}{x-2}), & 2 < x < 3, \\ 0, & x \geqslant 3. \end{cases}$$

Even if f is infinite over some interval the distribution function and the decreasing rearrangement are still defined and can be calculated, for any $\alpha \geqslant 0$

$$\begin{aligned} f_*(\alpha) =& \mathrm{m}(\{x \in [1,2] : \infty > \alpha\} \cup \{x \in (0,1) : \ln(\frac{1}{1-x}) > \alpha\} \\ & \cup \{x \in (2,3) : \ln(\frac{1}{x-2}) > \alpha\}) \\ =& 1 + \mathrm{m}((1 - e^{-\alpha}, 1)) + \mathrm{m}((2, e^{-\alpha} + 2)) \\ =& 1 + 2e^{-\alpha}, \end{aligned}$$

and

$$f^*(t) = \begin{cases} \infty, & 0 \leqslant t \leqslant 1, \\ \ln(\frac{2}{t-1}), & 1 < t < 3, \\ 0, & t \geqslant 3. \end{cases}$$



(a) (b) (c)

Example 2.23. Consider the function $f(x) = x$ for all $x \in [0, \infty)$. Then $f_*(\alpha) = \mathrm{m}(\{x \in [0, \infty) : x > \alpha\}) = \infty$ for all $\alpha \geqslant 0$, which implies that $f^*(t) = \inf\{\alpha \geqslant 0 : \infty \leqslant t\} = \infty$ for all $t \geqslant 0$.

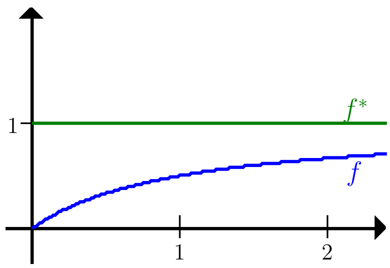
Example 2.24. Consider $f(x) = \frac{x}{1+x}$ for $x \geqslant 0$. It is clear that $f_*(\alpha) = 0$ for $\alpha \geqslant 1$ since $|f(x)| < 1$. For $\alpha \in [0, 1)$, we have

$$\begin{aligned} f_*(\alpha) =& \mathrm{m}(\{x \in [0, \infty) : \frac{x}{1+x} > \alpha\}) \\ =& \mathrm{m}(\{x \in [0, \infty) : x > \frac{\alpha}{1-\alpha}\}) = \infty. \end{aligned}$$

That is,

$$f_*(t) = \begin{cases} \infty, & 0 \leqslant \alpha < 1, \\ 0, & \alpha \geqslant 1. \end{cases}$$

Thus, $f^*(t) = \inf\{\alpha \geqslant 0 : f_*(\alpha) \leqslant t\} = 1$.



Proposition 2.25. *The decreasing rearrangement f^* of the measurable function f on $\mathbb{R}^n$ has the following properties:*

(i) *$f^*(t)$ is a non-negative and non-increasing function on $[0, \infty)$.*
(ii) *$f^*(t)$ is right continuous on $[0, \infty)$.*
(iii) *$(kf)^* = |k| f^*$ for $k \in \mathbb{C}$.*
(iv) *$|f| \leqslant |g|$ a.e. implies that $f^* \leqslant g^*$.*

(v) $(f+g)^*(t_1+t_2) \leqslant f^*(t_1)+g^*(t_2)$.
(vi) $(fg)^*(t_1+t_2) \leqslant f^*(t_1)g^*(t_2)$.
(vii) $|f| \leqslant \liminf_{k\to\infty} |f_k|$ *a.e. implies that* $f^* \leqslant \liminf_{k\to\infty} f_k^*$.
(viii) $|f_k| \uparrow |f|$ *a.e. implies that* $f_k^* \uparrow f^*$.
(ix) $f^*(f_*(\alpha)) \leqslant \alpha$ *whenever* $f_*(\alpha) < \infty$.
(x) $f_*(f^*(t)) = \mathrm{m}(\{|f| > f^*(t)\}) \leqslant t \leqslant \mathrm{m}(\{|f| \geqslant f^*(t)\})$ *if* $f^*(t) < \infty$.
(xi) $f^*(t) > \alpha$ *if and only if* $f_*(\alpha) > t$.
(xii) f^* *is equimeasurable with* f*, that is,* $(f^*)_*(\alpha) = f_*(\alpha)$ *for any* $\alpha \geqslant 0$.
(xiii) $(|f|^p)^*(t) = (f^*(t))^p$ *for* $1 \leqslant p < \infty$.
(xiv) $\|f^*\|_p = \|f\|_p$ *for* $1 \leqslant p < \infty$.
(xv) $\|f\|_\infty = f^*(0)$.
(xvi) $\sup_{t>0} t^s f^*(t) = \sup_{\alpha>0} \alpha(f_*(\alpha))^s$ *for* $0 < s < \infty$.

Proof. (v) Assume that $f^*(t_1)+g^*(t_2) < \infty$, otherwise, there is nothing to prove. Then for $\alpha_1 = f^*(t_1)$ and $\alpha_2 = g^*(t_2)$, by (x), we have $f_*(\alpha_1) \leqslant t_1$ and $g_*(\alpha_2) \leqslant t_2$. From (iv) in Proposition 2.15, it holds

$$(f+g)_*(\alpha_1+\alpha_2) \leqslant f_*(\alpha_1)+g_*(\alpha_2) \leqslant t_1+t_2.$$

Using the definition of the decreasing rearrangement, we have

$$(f+g)^*(t_1+t_2) = \inf\{\alpha : (f+g)_*(\alpha) \leqslant t_1+t_2\} \leqslant \alpha_1+\alpha_2 = f^*(t_1)+g^*(t_2).$$

(vi) Similar to (v), by (v) in Proposition 2.15, it holds that $(fg)_*(\alpha_1\alpha_2) \leqslant f_*(\alpha_1) + g_*(\alpha_2) \leqslant t_1+t_2$. Then, we have

$$(fg)^*(t_1+t_2) = \inf\{\alpha : (fg)_*(\alpha) \leqslant t_1+t_2\} \leqslant \alpha_1\alpha_2 = f^*(t_1)g^*(t_2).$$

(xi) If $f_*(\alpha) > t$, then by the decreasing of f_*, we have $\alpha < \inf\{\beta : f_*(\beta) \leqslant t\} = f^*(t)$. Conversely, if $f^*(t) > \alpha$, i.e., $\inf\{\beta : f_*(\beta) \leqslant t\} > \alpha$, we get $f_*(\alpha) > t$ by the decreasing of f_* again.

(xii) By the definition and (xi), we have

$$(f^*)_*(\alpha) = \mathrm{m}(\{t \geqslant 0 : f^*(t) > \alpha\}) = \mathrm{m}(\{t \geqslant 0 : f_*(\alpha) > t\}) = f_*(\alpha).$$

(xiii) For $\alpha \in [0,\infty)$, we have

$$\begin{aligned}(|f|^p)^*(t) &= \inf\{\alpha \geqslant 0 : \mathrm{m}(\{x : |f(x)|^p > \alpha\}) \leqslant t\}\\ &= \inf\{\sigma^p \geqslant 0 : \mathrm{m}(\{x : |f(x)| > \sigma\}) \leqslant t\} = (f^*(t))^p,\end{aligned}$$

where $\sigma = \alpha^{1/p}$.

(xiv) From Theorem 2.16, we have

$$\|f^*(t)\|_p^p = \int_0^\infty |f^*(t)|^p dt = p\int_0^\infty \alpha^{p-1}(f^*)_*(\alpha)d\alpha = p\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha = \|f\|_p^p.$$

We remain the proofs of others to interested readers. □

Having disposed of the basic properties of the decreasing rearrangement of functions, we proceed with the definition of the Lorentz spaces.

Definition 2.26. Given f a measurable function on $\mathbb{R}^n$ and $1 \leqslant p, q \leqslant \infty$, define

$$\|f\|_{L^{p,q}} = \begin{cases} \left(\displaystyle\int_0^\infty \left(t^{\frac{1}{p}} f^*(t)\right)^q \frac{dt}{t}\right)^{\frac{1}{q}}, & q < \infty,\\ \displaystyle\sup_{t>0} t^{\frac{1}{p}} f^*(t), & q = \infty.\end{cases}$$

The set of all f with $\|f\|_{L^{p,q}} < \infty$ is denoted by $L^{p,q}(\mathbb{R}^n)$ and is called the *Lorentz space* with indices p and q.

As in L^p and in weak L^p, two functions in $L^{p,q}$ will be considered equal if they are equal almost everywhere. Observe that the previous definition implies that $L^{p,\infty} = L^p_*$ in view of (xvi) in Proposition 2.25 and $L^{p,p} = L^p$ in view of (xiv) in Proposition 2.25 for $1 \leqslant p < \infty$. By (i) and (xv) in Proposition 2.25, we have $\|f\|_{L^{\infty,\infty}} = \sup_{t>0} f^*(t) = f^*(0) = \|f\|_\infty$ which implies that $L^{\infty,\infty} = L^\infty = L^\infty_*$. Thus, we have

Theorem 2.27. *Let $1 \leqslant p \leqslant \infty$. Then it holds, with equality of norms, that*

$$L^{p,p} = L^p, \quad L^{p,\infty} = L^p_*.$$

Remark 2.28. For the Lorentz space $L^{p,q}$, the case when $p = \infty$ and $1 \leqslant q < \infty$ is not of any interest. The reason is that $\|f\|_{L^{\infty,q}} < \infty$ implies that $f = 0$ a.e. on $\mathbb{R}^n$. In fact, assume that $L^{\infty,q}$ is a non-trivial space, there exists a nonzero function $f \in L^{\infty,q}$ on a nonzero measurable set, that is, there exists a constant $c > 0$ and a set E of positive measure such that $|f(x)| > c$ for all $x \in E$. Then, by (iv) in Proposition 2.25, we have

$$\|f\|^q_{L^{\infty,q}} = \int_0^\infty (f^*(t))^q \frac{dt}{t} \geqslant \int_0^\infty [(f\chi_E)^*(t)]^q \frac{dt}{t} \geqslant \int_0^{\mathfrak{m}(E)} c^q \frac{dt}{t} = \infty,$$

since $(f\chi_E)^*(t) = 0$ for $t > \mathfrak{m}(E)$. Hence, we have a contradiction. Thus, $f = 0$ a.e. on $\mathbb{R}^n$.

The next result shows that for any fixed p, the Lorentz spaces $L^{p,q}$ increase as the exponent q increases.

Theorem 2.29. *Let $1 \leqslant p \leqslant \infty$ and $1 \leqslant q < r \leqslant \infty$. Then, there exists some constant $C_{p,q,r}$ such that*

$$\|f\|_{L^{p,r}} \leqslant C_{p,q,r} \|f\|_{L^{p,q}}, \tag{2.14}$$

where $C_{p,q,r} = (q/p)^{1/q-1/r}$. In other words, $L^{p,q} \subset L^{p,r}$.

Proof. We may assume $p < \infty$ since the case $p = \infty$ is trivial. Since f^* is non-creasing, we have

$$\begin{aligned} t^{1/p} f^*(t) &= \left[\frac{q}{p}\int_0^t s^{q/p-1} ds\right]^{1/q} f^*(t) = \left\{\frac{q}{p}\int_0^t [s^{1/p} f^*(t)]^q \frac{ds}{s}\right\}^{1/q} \\ &\leqslant \left\{\frac{q}{p}\int_0^t [s^{1/p} f^*(s)]^q \frac{ds}{s}\right\}^{1/q} \leqslant \left(\frac{q}{p}\right)^{1/q} \|f\|_{L^{p,q}}. \end{aligned}$$

Hence, taking the supremum over all $t > 0$, we obtain

$$\|f\|_{L^{p,\infty}} \leqslant \left(\frac{q}{p}\right)^{1/q} \|f\|_{L^{p,q}}. \tag{2.15}$$

This establishes (2.14) in the case $r = \infty$. Finally, when $r < \infty$, we have by (2.15)

$$\begin{aligned} \|f\|_{L^{p,r}} &= \left\{\int_0^\infty [t^{1/p} f^*(t)]^{r-q+q} \frac{dt}{t}\right\}^{1/r} \\ &\leqslant \sup_{t>0} [t^{1/p} f^*(t)]^{(r-q)/r} \left\{\int_0^\infty [t^{1/p} f^*(t)]^q \frac{dt}{t}\right\}^{\frac{1}{q}\cdot\frac{q}{r}} \\ &= \|f\|^{(r-q)/r}_{L^{p,\infty}} \|f\|^{q/r}_{L^{p,q}} \leqslant \left(\frac{q}{p}\right)^{\frac{r-q}{rq}} \|f\|_{L^{p,q}}. \end{aligned}$$

This completes the proof. □

In general, $L^{p,q}$ is a quasi-normed space, since the functional $\|\cdot\|_{L^{p,q}}$ satisfies the conditions of normed spaces except the triangle inequality. In fact, by (v) in Proposition 2.25, it holds

$$\|f + g\|_{L^{p,q}} \leqslant 2^{1/p+1/q+1}(\|f\|_{L^{p,q}} + \|g\|_{L^{p,q}}). \tag{2.16}$$

However, is this space complete with respect to its quasi-norm? The next theorem answers this question.

Theorem 2.30. *Let $1 \leqslant p, q \leqslant \infty$. Then the spaces $L^{p,q}(\mathbb{R}^n)$ are complete with respect to their quasi-norms and they are therefore quasi-Banach spaces.*

Proof. See [Gra04, p. 50, Theorem 1.4.11]. □

For the duals of Lorentz spaces, we have

Theorem 2.31. *Let $1 < p, q < \infty$, $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$. Then we have*
$$(L^{1,1})' = (L^1)' = L^\infty, \quad (L^{1,q})' = \{0\}, \quad (L^{p,q})' = L^{p',q'}.$$

Proof. See [Gra04, p. 52-55, Theorem 1.4.17]. □

For more results, one can see [Gra04, Kri02].

2.4 Marcinkiewicz' interpolation theorem

We first introduce the definition of quasi-linear operators.

Definition 2.32. An operator T mapping functions on a measure space into functions on another measure space is called *quasi-linear* if $T(f+g)$ is defined whenever Tf and Tg are defined and if $|T(\lambda f)(x)| \leqslant \kappa|\lambda||Tf(x)|$ and $|T(f+g)(x)| \leqslant K(|Tf(x)| + |Tg(x)|)$ for a.e. x, where κ and K is a positive constant independent of f and g.

The idea we have used, in Definition 2.7, of splitting f into two parts according to their respective size, is the main idea of the proof of the theorem that follows. There, we will also use two easily proved inequalities, which are well-known results of Hardy's (see [HLP88, p. 245–246]):

Lemma 2.33 (Hardy inequalities)**.** *If $q \geqslant 1$, $r > 0$ and g is a measurable, non-negative function on $(0, \infty)$, then*
$$\left(\int_0^\infty \left(\int_0^t g(y)dy\right)^q t^{-r}\frac{dt}{t}\right)^{1/q} \leqslant \frac{q}{r}\left(\int_0^\infty (yg(y))^q y^{-r}\frac{dy}{y}\right)^{1/q}, \tag{2.17}$$
$$\left(\int_0^\infty \left(\int_t^\infty g(y)dy\right)^q t^{r}\frac{dt}{t}\right)^{1/q} \leqslant \frac{q}{r}\left(\int_0^\infty (yg(y))^q y^{r}\frac{dy}{y}\right)^{1/q}. \tag{2.18}$$

Proof. To prove (2.17), we use Jensen's inequality[2] with the convex function $\varphi(x) = x^q$ on $(0, \infty)$. Then
$$\begin{aligned}\left(\int_0^t g(y)dy\right)^q &= \left(\frac{1}{\int_0^t y^{r/q-1}dy}\int_0^t g(y)y^{1-r/q}y^{r/q-1}dy\right)^q \left(\int_0^t y^{r/q-1}dy\right)^q \\ &\leqslant \left(\int_0^t y^{r/q-1}dy\right)^{q-1}\int_0^t \left(g(y)y^{1-r/q}\right)^q y^{r/q-1}dy\end{aligned}$$

[2] **Jensen's inequality:** If f is any real-valued measurable function on a set Ω and φ is convex over the range of f, then
$$\varphi\left(\frac{1}{G}\int_\Omega f(x)g(x)dx\right) \leqslant \frac{1}{G}\int_\Omega \varphi(f(x))g(x)dx,$$
where $g(x) \geqslant 0$ satisfies $G = \int_\Omega g(x)dx > 0$.

$$= \left(\frac{q}{r} t^{r/q}\right)^{q-1} \int_0^t (yg(y))^q \, y^{r/q-1-r} dy.$$

By integrating both sides over $(0, \infty)$ and use the Fubini theorem, we get that

$$\begin{aligned}\int_0^\infty \left(\int_0^t g(y)dy\right)^q t^{-r-1} dt \leqslant& \left(\frac{q}{r}\right)^{q-1} \int_0^\infty t^{-1-r/q} \left(\int_0^t (yg(y))^q \, y^{r/q-1-r} dy\right) dt \\ =& \left(\frac{q}{r}\right)^{q-1} \int_0^\infty (yg(y))^q \, y^{r/q-1-r} \left(\int_y^\infty t^{-1-r/q} dt\right) dy \\ =& \left(\frac{q}{r}\right)^q \int_0^\infty (yg(y))^q \, y^{-1-r} dy,\end{aligned}$$

which yields (2.17) immediately.

To prove (2.18), we denote $f(x) = g(1/x)/x^2$. Then by taking $t = 1/s$ and $y = 1/x$, and then applying (2.17) and changing variable again by $x = 1/y$, we obtain

$$\begin{aligned}&\left(\int_0^\infty \left(\int_t^\infty g(y)dy\right)^q t^{r-1} dt\right)^{1/q} = \left(\int_0^\infty \left(\int_{1/s}^\infty g(y)dy\right)^q s^{-r-1} ds\right)^{1/q} \\ =&\left(\int_0^\infty \left(\int_0^s g(1/x)/x^2 dx\right)^q s^{-r-1} ds\right)^{1/q} = \left(\int_0^\infty \left(\int_0^s f(x)dx\right)^q s^{-r-1} ds\right)^{1/q} \\ \leqslant&\frac{q}{r} \left(\int_0^\infty (xf(x))^q x^{-r-1} dx\right)^{1/q} = \frac{q}{r} \left(\int_0^\infty (g(1/x)/x)^q x^{-r-1} dx\right)^{1/q} \\ =&\frac{q}{r} \left(\int_0^\infty (g(y)y)^q y^{r-1} dy\right)^{1/q}.\end{aligned}$$

Thus, we complete the proofs. □

Now, we give the Marcinkiewicz[3] interpolation theorem[4] and its proof due to Hunt and Weiss in [HW64].

Theorem 2.34 (Marcinkiewicz interpolation theorem)**.** *Assume that $1 \leqslant p_i \leqslant q_i \leqslant \infty$, $p_0 < p_1$, $q_0 \neq q_1$ and T is a quasi-linear mapping, defined on $L^{p_0} + L^{p_1}$, which is simultaneously of weak types (p_0, q_0) and (p_1, q_1), i.e.,*

$$\|Tf\|_{L^{q_0,\infty}} \leqslant A_0 \|f\|_{p_0}, \quad \|Tf\|_{L^{q_1,\infty}} \leqslant A_1 \|f\|_{p_1}. \tag{2.19}$$

If $0 < \theta < 1$, and

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

then T is of type (p, q), namely

$$\|Tf\|_q \leqslant A \|f\|_p, \quad f \in L^p.$$

Here $A = A(A_i, p_i, q_i, \theta)$, but it does not otherwise depend on either T or f.

Proof. Let σ be the slope of the line segment in $\mathbb{R}^2$ joining $(1/p_0, 1/q_0)$ with $(1/p_1, 1/q_1)$. Since $(1/p, 1/q)$ lies on this segment, we can denote the slope of this segment by

$$\sigma = \frac{1/q_0 - 1/q}{1/p_0 - 1/p} = \frac{1/q - 1/q_1}{1/p - 1/p_1},$$

[3] Józef Marcinkiewicz (1910–1940) was a Polish mathematician. He was a student of Antoni Zygmund; and later worked with Juliusz Schauder, and Stefan Kaczmarz.

[4] The theorem was first announced by Marcinkiewicz (1939), who showed this result to Antoni Zygmund shortly before he died in World War II. The theorem was almost forgotten by Zygmund, and was absent from his original works on the theory of singular integral operators. Later Zygmund (1956) realized that Marcinkiewicz's result could greatly simplify his work, at which time he published his former student's theorem together with a generalization of his own.

which may be positive or negative, but is not either 0 or ∞ since $q_0 \neq q_1$ and $p_0 < p_1$.

For any $t > 0$, we split an arbitrary function $f \in L^p$ as follows:

$$f = f^t + f_t$$

where

$$f^t(x) = \begin{cases} f(x), & |f(x)| > f^*(t^\sigma), \\ 0, & \text{otherwise}, \end{cases}$$

and $f_t = f - f^t$.

Then we can verify that

$$\begin{aligned} (f^t)^*(y) &\begin{cases} \leqslant f^*(y), & 0 \leqslant y \leqslant t^\sigma, \\ = 0, & y > t^\sigma, \end{cases} \\ (f_t)^*(y) &\leqslant \begin{cases} f^*(t^\sigma), & 0 \leqslant y \leqslant t^\sigma, \\ f^*(y), & y > t^\sigma. \end{cases} \end{aligned} \tag{2.20}$$

In fact, by (iv) in Proposition 2.25, $|f^t| \leqslant |f|$ implies $(f^t)^*(y) \leqslant f^*(y)$ for all $y \geqslant 0$. Moreover, by the definition of f^t and (x) in Proposition 2.25, we have $(f^t)_*(\alpha) \leqslant (f^t)_*(f^*(t^\sigma)) = f_*(f^*(t^\sigma)) \leqslant t^\sigma$ for any $\alpha \geqslant 0$. Thus, for $y > t^\sigma$, we get $(f^t)^*(y) = 0$. Similarly, by (iv) in Proposition 2.25, we have $(f_t)^*(y) \leqslant f^*(y)$ for any $y \geqslant 0$ since $|f_t| \leqslant |f|$. On the other hand, for $y \geqslant 0$, we have $(f_t)^*(y) \leqslant (f_t)^*(0) = \|f_t\|_\infty \leqslant f^*(t^\sigma)$ with the help of the non-increasing of $(f_t)^*(y)$ and (xv) in Proposition 2.25. Thus, $(f_t)^*(y) \leqslant \min(f^*(y), f^*(t^\sigma))$ for any $y \geqslant 0$ which implies (2.20).

Suppose $p_1 < \infty$. Notice that $p \leqslant q$, because $p_i \leqslant q_i$. By $|Tf| \leqslant K(|Tf^t| + |Tf_t|)$, Theorems 2.27 and 2.29, (2.19), and then by a change of variables and Hardy's inequalities (2.17) and (2.18), we get

$$\begin{aligned} &\|Tf\|_q \leqslant K(\|Tf^t\|_q + \|Tf_t\|_q) = K(\|Tf^t\|_{L^{q,q}} + \|Tf_t\|_{L^{q,q}}) \\ \leqslant& K(p/q)^{1/p-1/q}(\|Tf^t\|_{L^{q,p}} + \|Tf_t\|_{L^{q,p}}) \\ =& K\left(\frac{p}{q}\right)^{1/p-1/q}\left\{\left(\int_0^\infty \left[t^{1/q}(Tf^t)^*(t)\right]^p \frac{dt}{t}\right)^{1/p} + \left(\int_0^\infty \left[t^{1/q}(Tf_t)^*(t)\right]^p \frac{dt}{t}\right)^{1/p}\right\} \\ \leqslant& K\left(\frac{p}{q}\right)^{1/p-1/q}\left\{A_0\left(\int_0^\infty \left[t^{1/q-1/q_0}\|f^t\|_{p_0}\right]^p \frac{dt}{t}\right)^{1/p}\right. \\ &\qquad\left. + A_1\left(\int_0^\infty \left[t^{1/q-1/q_1}\|f_t\|_{p_1}\right]^p \frac{dt}{t}\right)^{1/p}\right\} \\ \leqslant& K\left(\frac{p}{q}\right)^{1/p-1/q}\left\{A_0\left(\int_0^\infty \left[t^{1/q-1/q_0}\left(\frac{1}{p_0}\right)^{1-1/p_0}\|f^t\|_{L^{p_0,1}}\right]^p \frac{dt}{t}\right)^{1/p}\right. \\ &\qquad\left. + A_1\left(\int_0^\infty \left[t^{1/q-1/q_1}\left(\frac{1}{p_1}\right)^{1-1/p_1}\|f_t\|_{L^{p_1,1}}\right]^p \frac{dt}{t}\right)^{1/p}\right\} \\ =& K\left(\frac{p}{q}\right)^{1/p-1/q}\left\{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}\left(\int_0^\infty \left[t^{1/q-1/q_0}\left(\int_0^{t^\sigma} y^{1/p_0} f^*(y)\frac{dy}{y}\right)\right]^p \frac{dt}{t}\right)^{1/p}\right. \\ &\qquad + A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty \left[t^{1/q-1/q_1}\left(\int_{t^\sigma}^\infty y^{1/p_1} f^*(y)\frac{dy}{y}\right)\right]^p \frac{dt}{t}\right)^{1/p} \\ &\qquad\left. + A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty \left[t^{1/q-1/q_1}\left(\int_0^{t^\sigma} y^{1/p_1} f^*(t^\sigma)\frac{dy}{y}\right)\right]^p \frac{dt}{t}\right)^{1/p}\right\} \end{aligned}$$

$$
\begin{aligned}
=&K\left(\frac{p}{q}\right)^{1/p-1/q}|\sigma|^{-\frac{1}{p}}\left\{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}\left(\int_0^\infty s^{-p(1/p_0-1/p)}\left(\int_0^s y^{1/p_0}f^*(y)\frac{dy}{y}\right)^p\frac{ds}{s}\right)^{1/p}\right.\\
&+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty s^{p(1/p-1/p_1)}\left(\int_s^\infty y^{1/p_1}f^*(y)\frac{dy}{y}\right)^p\frac{ds}{s}\right)^{1/p}\\
&\left.+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty s^{p(1/p-1/p_1)}\left(\int_0^s y^{1/p_1}f^*(s)\frac{dy}{y}\right)^p\frac{ds}{s}\right)^{1/p}\right\}\\
\leqslant&K\left(\frac{p}{q}\right)^{1/p-1/q}|\sigma|^{-\frac{1}{p}}\left\{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}\frac{1}{(1/p_0-1/p)}\left(\int_0^\infty\left(y^{1/p}f^*(y)\right)^p\frac{dy}{y}\right)^{1/p}\right.\\
&+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\frac{1}{(1/p-1/p_1)}\left(\int_0^\infty\left(y^{1/p}f^*(y)\right)^p\frac{dy}{y}\right)^{1/p}\\
&\left.+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty s^{1-p/p_1}(p_1s^{1/p_1}f^*(s))^p\frac{ds}{s}\right)^{1/p}\right\}\\
=&K\left(\frac{p}{q}\right)^{1/p-1/q}|\sigma|^{-1/p}\left\{\frac{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}}{\frac{1}{p_0}-\frac{1}{p}}+\frac{A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}}{\frac{1}{p}-\frac{1}{p_1}}+A_1p_1^{1/p_1}\right\}\|f\|_p\\
=&A\|f\|_p.
\end{aligned}
$$

In case $p_1=\infty$ the proof is the same except for the use of the estimate $\|f_t\|_\infty\leqslant f^*(t^\sigma)$, we can get

$$
A=K\left(\frac{p}{q}\right)^{1/p-1/q}|\sigma|^{-1/p}\left\{\frac{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}}{\frac{1}{p_0}-\frac{1}{p}}+A_1\right\}.
$$

Thus, we complete the proof. $\square$

From the proof given above it is easy to see that the theorem can be extended to the following situation: The underlying measure space $\mathbb{R}^n$ of the $L^{p_i}(\mathbb{R}^n)$ can be replaced by a general measurable space (and the measurable space occurring in the domain of T need not be the same as the one entering in the range of T). A less superficial generalization of the theorem can be given in terms of the notation of Lorentz spaces, which unify and generalize the usual L^p spaces and the weak-type spaces. For a discussion of this more general form of the Marcinkiewicz interpolation theorem see [SW71, Chapter V] and [BL76, Chapter 5].

As an application of this powerful tool, we present a generalization of the Hausdorff-Young inequality due to Paley. The main difference between the theorems being that Payley introduced a weight function into his inequality and resorted to the theorem of Marcinkiewicz. In what follows, we consider the measure space $(\mathbb{R}^n,\mu)$ where μ denotes the Lebesgue measure. Let w be a weihgt function on $\mathbb{R}^n$, i.e., a positive and measurable function on $\mathbb{R}^n$. Then we denote by $L^p(w)$ the L^p-space with respect to wdx. The norm on $L^p(w)$ is

$$
\|f\|_{L^p(w)}=\left(\int_{\mathbb{R}^n}|f(x)|^pw(x)dx\right)^{1/p}.
$$

With this notation we have the following theorem.

Theorem 2.35 (Hardy-Littlewood-Paley theorem on $\mathbb{R}^n$). *Assume that $1\leqslant p\leqslant 2$. Then*

$$
\|\mathscr{F}f\|_{L^p(|\xi|^{-n(2-p)})}\leqslant C_p\|f\|_p.
$$

Proof. We considering the mapping $(Tf)(\xi) = |\xi|^n \hat{f}(\xi)$. By Plancherel theorem, we have
$$\|Tf\|_{L^2_*(|\xi|^{-2n})} \leqslant \|Tf\|_{L^2(|\xi|^{-2n})} = \|\hat{f}\|_2 \leqslant C\|f\|_2,$$
which implies that T is of weak type $(2,2)$. We now work towards showing that T is of weak type $(1,1)$. Thus, the Marcinkiewicz interpolation theorem implies the theorem.

Now, consider the set $E_\alpha = \{\xi : |\xi|^n \hat{f}(\xi) > \alpha\}$. For simplicity, we let ν denote the measure $|\xi|^{-2n} d\xi$ and assume that $\|f\|_1 = 1$. Then, $|\hat{f}(\xi)| \leqslant 1$. For $\xi \in E_\alpha$, we therefore have $\alpha \leqslant |\xi|^n$. Consequently,
$$(Tf)_*(\alpha) = \nu(E_\alpha) = \int_{E_\alpha} |\xi|^{-2n} d\xi \leqslant \int_{|\xi|^n \geqslant \alpha} |\xi|^{-2n} d\xi \leqslant C\alpha^{-1}.$$
Thus, we proves that
$$\alpha \cdot (Tf)_*(\alpha) \leqslant C\|f\|_1,$$
which implies T is of weak type $(1,1)$. Therefore, we complete the proof. □

Chapter 3
The Maximal Function and Calderón-Zygmund Decomposition

3.1 Two covering lemmas

Lemma 3.1 (Finite version of Vitali covering lemma)**.** *Suppose $\mathcal{B} = \{B_1, B_2, \cdots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^n$. Then, there exists a disjoint sub-collection B_{j_1}, B_{j_2}, $\cdots$, B_{j_k} of $\mathcal{B}$ such that*

$$m\left(\bigcup_{\ell=1}^{N} B_\ell\right) \leqslant 3^n \sum_{i=1}^{k} \mathrm{m}(B_{j_i}).$$

Proof. The argument we give is constructive and relies on the following simple observation: *Suppose B and B' are a pair of balls that intersect, with the radius of B' being not greater than that of B. Then B' is contained in the ball $\tilde{B}$ that is concentric with B but with 3 times its radius.* (See Fig 3.1.)

As a first step, we pick a ball B_{j_1} in $\mathcal{B}$ with maximal (i.e., largest) radius, and then delete from $\mathcal{B}$ the ball B_{j_1} as well as any balls that intersect B_{j_1}. Thus all the balls that are deleted are contained in the ball $\tilde{B}_{j_1}$ concentric with B_{j_1}, but with 3 times its radius.

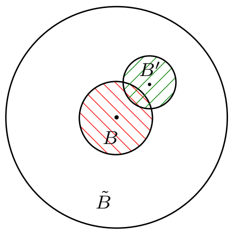


Fig. 3.1 The balls B and $\tilde{B}$

The remaining balls yield a new collection $\mathcal{B}'$, for which we repeat the procedure. We pick B_{j_2} and any ball that intersects B_{j_2}. Continuing this way, we find, after at most N steps, a collection of disjoint balls B_{j_1}, B_{j_2}, $\cdots$, B_{j_k}.

Finally, to prove that this disjoint collection of balls satisfies the inequality in the lemma, we use the observation made at the beginning of the proof. Let $\tilde{B}_{j_i}$ denote the ball concentric with B_{j_i}, but with 3 times its radius. Since any ball B in $\mathcal{B}$ must intersect a ball B_{j_i} and have equal or smaller radius than B_{j_i}, we must have $\cup_{B\cap B_{j_i}\neq\varnothing} B \subset \tilde{B}_{j_i}$, thus

$$m\left(\bigcup_{\ell=1}^{N} B_\ell\right) \leqslant m\left(\bigcup_{i=1}^{k} \tilde{B}_{j_i}\right) \leqslant \sum_{i=1}^{k} \mathrm{m}(\tilde{B}_{j_i}) = 3^n \sum_{i=1}^{k} \mathrm{m}(B_{j_i}).$$

In the last step, we have used the fact that in $\mathbb{R}^n$ a dilation of a set by $\delta > 0$ results in the multiplication by δ^n of the Lebesgue measure of this set. $\square$

For the infinite version of Vitali covering lemma, one can see the textbook [Ste70, the lemma on p.9].

The decomposition of a given set into a disjoint union of cubes (or balls) is a fundamental tool in the theory described in this chapter. By cubes we mean closed cubes; by disjoint we mean that their interiors are disjoint. We have in mind the idea first introduced by Whitney and formulated as follows.

Theorem 3.2 (Whitney covering lemma)**.** *Let F be a non-empty closed set in $\mathbb{R}^n$ and Ω be its complement. Then there exists a collection of cubes $\mathscr{F} = \{Q_k\}$ whose sides are parallel to the axes, such that*

(i) $\bigcup_{k=1}^{\infty} Q_k = \Omega = F^c$,

(ii) $Q_j^\circ \cap Q_k^\circ = \varnothing$ *if* $j \neq k$*, where* Q° *denotes the interior of* Q*,*

(iii) *there exist two constants* $c_1, c_2 > 0$ *independent of* F *(In fact we may take* $c_1 = 1$ *and* $c_2 = 4$*.), such that*

$$c_1 \operatorname{diam}(Q_k) \leqslant \operatorname{dist}(Q_k, F) \leqslant c_2 \operatorname{diam}(Q_k).$$

Proof. Consider the lattice of points in $\mathbb{R}^n$ whose coordinates are integral. This lattice determines a mesh $\mathscr{M}_0$, which is a collection of cubes: namely all cubes of unit length, whose vertices are points of the above lattice. The mesh $\mathscr{M}_0$ leads to a two-way infinite chain of such meshes $\{\mathscr{M}_k\}_{-\infty}^{\infty}$, with $\mathscr{M}_k = 2^{-k}\mathscr{M}_0$.

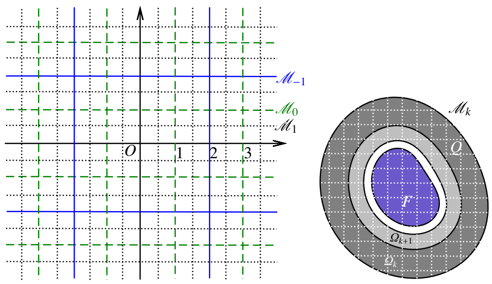


Fig. 3.2 Meshes and layers: $\mathscr{M}_0$ with dashed (green) lines; $\mathscr{M}_1$ with dotted lines; $\mathscr{M}_{-1}$ with solid (blue) lines

Thus each cube in the mesh $\mathscr{M}_k$ gives rise to 2^n cubes in the mesh $\mathscr{M}_{k+1}$ by bisecting the sides. The cubes in the mesh $\mathscr{M}_k$ each have sides of length 2^{-k} and are thus of diameter $\sqrt{n}2^{-k}$.

In addition to the meshes $\mathscr{M}_k$, we consider the layers Ω_k, defined by

$$\Omega_k = \left\{x : c2^{-k} < \operatorname{dist}(x, F) \leqslant c2^{-k+1}\right\},$$

where c is a positive constant which we shall fix momentarily. Obviously, $\Omega = \bigcup_{k=-\infty}^{\infty} \Omega_k$.

Now we make an initial choice of cubes, and denote the resulting collection by $\mathscr{F}_0$. Our choice is made as follows. We consider the cubes of the mesh $\mathscr{M}_k$, (each such cube is of size approximately 2^{-k}), and include a cube of this mesh in $\mathscr{F}_0$ if it intersects Ω_k, (the points of the latter are all approximately at a distance 2^{-k} from F). Namely,

$$\mathscr{F}_0 = \bigcup_k \{Q \in \mathscr{M}_k : Q \cap \Omega_k \neq \varnothing\}.$$

We then have

$$\bigcup_{Q \in \mathscr{F}_0} Q = \Omega.$$

For appropriate choice of c, we claim that

$$\operatorname{diam}(Q) \leqslant \operatorname{dist}(Q, F) \leqslant 4 \operatorname{diam}(Q), \quad Q \in \mathscr{F}_0. \tag{3.1}$$

Let us prove (3.1) first. Suppose $Q \in \mathscr{M}_k$; then $\operatorname{diam}(Q) = \sqrt{n}2^{-k}$. Since $Q \in \mathscr{F}_0$, there exists an $x \in Q \cap \Omega_k$. Thus $\operatorname{dist}(Q, F) \leqslant \operatorname{dist}(x, F) \leqslant c2^{-k+1}$, and $\operatorname{dist}(Q, F) \geqslant \operatorname{dist}(x, F) - \operatorname{diam}(Q) > c2^{-k} - \sqrt{n}2^{-k}$. If we choose $c = 2\sqrt{n}$ we get (3.1).

Then by (3.1) the cubes $Q \in \mathscr{F}_0$ are disjoint from F and clearly cover Ω. Therefore, (i) is also proved.

Notice that the collection $\mathscr{F}_0$ has all our required properties, except that the cubes in it are not necessarily disjoint. To finish the proof of the theorem, we need to refine our choice leading to $\mathscr{F}_0$, eliminating those cubes which were really unnecessary.

We require the following simple observation. Suppose Q_1 and Q_2 are two cubes (taken respectively from the mesh $\mathscr{M}_{k_1}$ and $\mathscr{M}_{k_2}$). Then if Q_1 and Q_2 are not disjoint, one of the two must be contained in the other. (In particular, $Q_1 \subset Q_2$, if $k_1 \geqslant k_2$.)

Start now with any cube $Q \in \mathscr{F}_0$, and consider the maximal cube in $\mathscr{F}_0$ which contains it. In view of the inequality (3.1), for any cube $Q' \in \mathscr{F}_0$ which contains $Q \in \mathscr{F}_0$, we have $\operatorname{diam}(Q') \leqslant \operatorname{dist}(Q', F) \leqslant \operatorname{dist}(Q, F) \leqslant 4 \operatorname{diam}(Q)$. Moreover, any two cubes Q' and Q'' which contain Q have obviously a non-trivial intersection. Thus by the observation made above each cube $Q \in \mathscr{F}_0$ has a unique maximal cube in $\mathscr{F}_0$ which contains it. By the same taken these maximal cubes are also disjoint. We let $\mathscr{F}$ denote the collection of maximal cubes of $\mathscr{F}_0$. Then obviously

(i) $\bigcup_{Q \in \mathscr{F}} Q = \Omega$,

(ii) The cubes of $\mathscr{F}$ are disjoint,

(iii) $\operatorname{diam}(Q) \leqslant \operatorname{dist}(Q, F) \leqslant 4 \operatorname{diam}(Q)$, $Q \in \mathscr{F}$.

Therefore, we complete the proof. □

3.2 Hardy-Littlewood maximal function

Maximal functions appear in many forms in harmonic analysis. One of the most important of these is the Hardy-Littlewood maximal function. They play an important role in understanding, for example, the differentiability properties of functions, singular integrals and partial differential equations. They often provide a deeper and more simplified approach to understanding problems in these areas than other methods.

First, we consider the differentiation of the integral for one-dimensional functions. If f is given on $[a, b]$ and integrable on that interval, we let

$$F(x) = \int_a^x f(y)dy, \quad x \in [a, b].$$

To deal with $F'(x)$, we recall the definition of the derivative as the limit of the quotient $\frac{F(x+h)-F(x)}{h}$ when h tends to 0, i.e.,

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h}.$$

We note that this quotient takes the form (say in the case $h > 0$)

$$\frac{1}{h} \int_x^{x+h} f(y)dy = \frac{1}{|I|} \int_I f(y)dy,$$

where we use the notation $I = (x, x+h)$ and $|I|$ for the length of this interval.

At this point, we pause to observe that the above expression in the "average" value of f over I, and that in the limit as $|I| \to 0$, *we might expect that these averages tend to* $f(x)$. Reformulating the question slightly, we may ask whether

$$\lim_{\substack{|I| \to 0 \\ x \in I}} \frac{1}{|I|} \int_I f(y)dy = f(x)$$

holds for suitable points x. In higher dimensions we can pose a similar question, where the averages of f are taken over appropriate sets that generalize the intervals in one dimension.

In particular, we can take the sets involved as the ball $B(x, r)$ of radius r, centered at x, and denote its measure by $\mathfrak{m}(B(x, r))$. It follows

$$\lim_{r \to 0} \frac{1}{\mathfrak{m}(B(x, r))} \int_{B(x,r)} f(y)dy = f(x), \quad \text{for a.e. } x? \tag{3.2}$$

Let us first consider a simple case, *when f is continuous at x, the limit does converge to $f(x)$*. Indeed, given $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$. Since

$$f(x) - \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} f(y)dy = \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} (f(x) - f(y))dy,$$

we find that whenever $B(x,r)$ is a ball of radius $r < \delta$, then

$$\left| f(x) - \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} f(y)dy \right| \leqslant \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} |f(x) - f(y)|dy < \varepsilon,$$

as desired.

In general, for this "averaging problem" (3.2), we shall have an affirmative answer. In order to study the limit (3.2), we consider its quantitative analogue, where "$\lim_{r\to 0}$" is replaced by "$\sup_{r>0}$", this is the *(centered) maximal function*. Since the properties of this maximal function are expressed in term of relative size and do not involve any cancelation of positive and negative values, we replace f by $|f|$.

Definition 3.3. If f is locally integrable[1] on $\mathbb{R}^n$, we define its *maximal function* $Mf : \mathbb{R}^n \to [0, \infty]$ by

$$Mf(x) = \sup_{r>0} \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} |f(y)|dy, \quad x \in \mathbb{R}^n. \tag{3.3}$$

Moreover, M is also called as the *Hardy-Littlewood maximal operator*.

The maximal function that we consider arose first in the one-dimensional situation treated by Hardy and Littlewood.[2] It is to be noticed that nothing excludes the possibility that $Mf(x)$ is infinite for any given x.

It is immediate from the definition that

Theorem 3.4. *If $f \in L^\infty(\mathbb{R}^n)$, then $Mf \in L^\infty(\mathbb{R}^n)$ and*

$$\|Mf\|_\infty \leqslant \|f\|_\infty.$$

By the previous statements, if f is continuous at x, then we have

$$\begin{aligned}|f(x)| &= \lim_{r\to 0} \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} |f(y)|dy \\ &\leqslant \sup_{r>0} \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} |f(y)|dy = Mf(x).\end{aligned}$$

Thus, we have proved

Theorem 3.5. *If $f \in C(\mathbb{R}^n)$, then*

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

for all $x \in \mathbb{R}^n$.

Sometimes, we will define the maximal function with cubes in place of balls. If $Q(x,r)$ is the cubes $[x - r, x + r]^n$, define

---

[1] A measurable function f on $\mathbb{R}^n$ is called to be *locally integrable*, if for every ball B the function $f(x)\chi_B(x)$ is integrable. We shall denote by $L^1_{loc}(\mathbb{R}^n)$ the space of all locally integrable functions. Loosely speaking, the behavior at infinity does not affact the local integrability of a function. For example, the functions $e^{|x|}$ and $|x|^{-1/2}$ are both locally integrable, but not integrable on $\mathbb{R}^n$.

[2] The Hardy-Littlewood maximal operator appears in many places but some of its most notable uses are in the proofs of the Lebesgue differentiation theorem and Fatou's theorem and in the theory of singular integral operators.

$$M'f(x) = \sup_{r>0} \frac{1}{(2r)^n} \int_{Q(x,r)} |f(y)|dy, \quad x \in \mathbb{R}^n. \tag{3.4}$$

When $n = 1$, M and M' coincide. If $n > 1$, then there exist constants c_n and C_n, depending only on n, such that

$$c_n M'f(x) \leqslant Mf(x) \leqslant C_n M'f(x). \tag{3.5}$$

Thus, the two operators M and M' are essentially interchangeable, and we will use whichever is more appropriate, depending on the circumstances. In addition, we can define a more general maximal function

$$M''f(x) = \sup_{Q \ni x} \frac{1}{\mathrm{m}(Q)} \int_Q |f(y)|dy, \tag{3.6}$$

where the supremum is taken over all cubes containing x. Again, M'' is pointwise equivalent to M. One sometimes distinguishes between M' and M'' by referring to the former as the centered and the latter as the non-centered maximal operator. Alternatively, we could define the non-centered maximal function with balls instead of cubes:

$$\tilde{M}f(x) = \sup_{B \ni x} \frac{1}{\mathrm{m}(B)} \int_B |f(y)|dy$$

at each $x \in \mathbb{R}^n$. Here, the supremum is taken over balls B in $\mathbb{R}^n$ which contain the point x and $\mathrm{m}(B)$ denotes the measure of B (in this case a multiple of the radius of the ball raised to the power n).

Example 3.6. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \chi_{(0,1)}(x)$. Then

$$\begin{aligned} Mf(x) =& M'f(x) = \begin{cases} \frac{1}{2x}, & x > 1, \\ 1, & 0 \leqslant x \leqslant 1, \\ \frac{1}{2(1-x)}, & x < 0, \end{cases} \\ \tilde{M}f(x) =& M''f(x) = \begin{cases} \frac{1}{x}, & x > 1, \\ 1, & 0 \leqslant x \leqslant 1, \\ \frac{1}{1-x}, & x < 0. \end{cases} \end{aligned}$$

In fact, for $x > 1$, we get

$$\begin{aligned} Mf(x) = M'f(x) =& \sup_{h>0} \frac{1}{2h} \int_{x-h}^{x+h} \chi_{(0,1)}(y)dy \\ =& \max\left(\sup_{x-h>0} \frac{1-x+h}{2h}, \sup_{x-h\leqslant 0} \frac{1}{2h} \right) = \frac{1}{2x}, \\ \tilde{M}f(x) = M''f(x) =& \sup_{h_1,h_2>0} \frac{1}{h_1+h_2} \int_{x-h_1}^{x+h_2} \chi_{(0,1)}(y)dy \\ =& \max\left(\sup_{0<x-h_1<1} \frac{1-x+h_1}{h_1}, \sup_{x-h_1\leqslant 0} \frac{1}{h_1} \right) = \frac{1}{x}. \end{aligned}$$

For $0 \leqslant x \leqslant 1$, it follows

$$\begin{aligned} Mf(x) = M'f(x) =& \sup_{h>0} \frac{1}{2h} \int_{x-h}^{x+h} \chi_{(0,1)}(y)dy \\ =& \max\left(\sup_{0<x-h<x+h<1} \frac{2h}{2h}, \sup_{0<x-h<1\leqslant x+h} \frac{1-x+h}{2h}, \right. \\ & \left. \sup_{x-h\leqslant 0<x+h<1} \frac{x+h}{2h}, \sup_{x-h\leqslant 0<1\leqslant x+h} \frac{1}{2h} \right) \\ =& \max\left(1, 1, 1, \frac{1}{2}\min\left(\frac{1}{x}, \frac{1}{1-x}\right)\right) = 1, \end{aligned}$$

$$\begin{aligned}\tilde{M}f(x) = M''f(x) =& \sup_{h_1,h_2>0} \frac{1}{h_1+h_2}\int_{x-h_1}^{x+h_2} \chi_{(0,1)}(y)dy\\ =&\max\Bigg(\sup_{0<x-h_1<x+h_2<1} \frac{h_1+h_2}{h_1+h_2}, \sup_{x-h_1<0<x+h_2<1} \frac{x+h_2}{h_1+h_2},\\ &\sup_{0<x-h_1<1<x+h_2} \frac{1-x+h_1}{h_1+h_2}, \sup_{x-h_1<0<1<x+h_2} \frac{1}{h_1+h_2}\Bigg)\\ =&1.\end{aligned}$$

For $x<0$, we have

$$\begin{aligned}Mf(x) = M'f(x) =& \max\left(\sup_{0<x+h<1, h>0} \frac{x+h}{2h}, \sup_{x+h\geqslant 1} \frac{1}{2h}\right) = \frac{1}{2(1-x)},\\ \tilde{M}f(x) = M''f(x) =& \max\left(\sup_{h_1,h_2>0, 0<x+h_2<1} \frac{x+h_2}{h_1+h_2}, \sup_{h_1>0, x+h_2\geqslant 1} \frac{1}{h_1+h_2}\right)\\ =&\frac{1}{1-x}.\end{aligned}$$

Observe that $f\in L^1(\mathbb{R})$, but $Mf, M'f, M''f, \tilde{M}f \notin L^1(\mathbb{R})$.

Remark 3.7. (i) Mf is defined at every point $x\in\mathbb{R}^n$ and if $f=g$ a.e., then $Mf(x) = Mg(x)$ at every $x\in\mathbb{R}^n$.

(ii) It may be well that $Mf=\infty$ for every $x\in\mathbb{R}^n$. For example, let $n=1$ and $f(x)=x^2$.

(iii) There are several definitions in the literature which are often equivalent.

Next, we state some immediate properties of the maximal function. The proofs are left to interested readers.

Proposition 3.8. *Let $f,g\in L^1_{loc}(\mathbb{R}^n)$. Then*

(i) *Positivity: $Mf(x)\geqslant 0$ for all $x\in\mathbb{R}^n$.*

(ii) *Sub-linearity: $M(f+g)(x)\leqslant Mf(x)+Mg(x)$.*

(iii) *Homogeneity: $M(\alpha f)(x) = |\alpha|Mf(x)$, $\alpha\in\mathbb{R}$.*

(iv) *Translation invariance: $M(\tau_y f) = (\tau_y Mf)(x) = Mf(x-y)$.*

With the Vitali covering lemma, we can state and prove the main results for the maximal function.

Theorem 3.9 (The maximal function theorem)**.** *Let f be a given function defined on $\mathbb{R}^n$.*

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

(ii) *If $f\in L^1(\mathbb{R}^n)$, then for every $\alpha>0$, M is of weak type $(1,1)$, i.e.,*

$$\mathfrak{m}(\{x : Mf(x)>\alpha\}) \leqslant \frac{3^n}{\alpha}\|f\|_1.$$

(iii) *If $f\in L^p(\mathbb{R}^n)$, $p\in(1,\infty]$, then $Mf\in L^p(\mathbb{R}^n)$ and*

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

where $A_p = 3^n p/(p-1)+1$ for $p\in(1,\infty)$ and $A_\infty=1$.

Proof. We first prove the second one, i.e., (ii). Denote

$$E_\alpha = \{x : Mf(x)>\alpha\},$$

then from the definitions of Mf and the supremum, for each $x\in E_\alpha$ and $0<\varepsilon< Mf(x)-\alpha$, there exists a $r>0$ such that

$$\frac{1}{\mathfrak{m}(B(x,r))}\int_{B(x,r)} |f(y)|dy > Mf(x)-\varepsilon>\alpha.$$

We denote that ball $B(x,r)$ by B_x that contains x. Therefore, for each B_x, we have

$$\mathfrak{m}(B_x) < \frac{1}{\alpha}\int_{B_x} |f(y)|dy. \tag{3.7}$$

Fix a compact subset K of E_α. Since K is covered by $\cup_{x\in E_\alpha} B_x$, by Heine-Borel theorem,[3] we may select a finite subcover of K, say $K \subset \bigcup_{\ell=1}^N B_\ell$. Lemma 3.1 guarantees the existence of a sub-collection $B_{j_1}, \cdots, B_{j_k}$ of disjoint balls with

$$\mathfrak{m}(\bigcup_{\ell=1}^N B_\ell) \leqslant 3^n \sum_{i=1}^k \mathfrak{m}(B_{j_i}). \tag{3.8}$$

Since the balls $B_{j_1}, \cdots, B_{j_k}$ are disjoint and satisfy (3.7) as well as (3.8), we find that

$$\begin{aligned}\mathfrak{m}(K) \leqslant &\mathfrak{m}(\bigcup_{\ell=1}^N B_\ell) \leqslant 3^n \sum_{i=1}^k \mathfrak{m}(B_{j_i}) \leqslant \frac{3^n}{\alpha}\sum_{i=1}^k \int_{B_{j_i}} |f(y)|dy \\ =&\frac{3^n}{\alpha}\int_{\bigcup_{i=1}^k B_{j_i}} |f(y)|dy \leqslant \frac{3^n}{\alpha}\int_{\mathbb{R}^n} |f(y)|dy.\end{aligned}$$

Since this inequality is true for all compact subsets K of E_α, the proof of the weak type inequality (ii) for the maximal operator is complete.

The above proof also gives the proof of (i) for the case when $p = 1$. For the case $p = \infty$, by Theorem 3.4, (i) and (iii) is true with $A_\infty = 1$.

Now, by using the Marcinkiewicz interpolation theorem between $L^1 \to L^{1,\infty}$ and $L^\infty \to L^\infty$, we can obtain simultaneously (i) and (iii) for the case $p \in (1,\infty)$. □

Now, we make some clarifying comments.

Remark 3.10. (1) The weak type estimate (ii) is the *best possible* for the distribution function of Mf, where f is an arbitrary function in $L^1(\mathbb{R}^n)$.

Indeed, we replace $|f(y)|dy$ in the definition of (3.3) by a Dirac measure $d\mu$ whose total measure of one is concentrated at the origin. The integral $\int_{B(x,r)} d\mu = 1$ only if the ball $B(x,r)$ contains the origin; otherwise, it will be zeros. Thus,

$$M(d\mu)(x) = \sup_{r>0,\ 0\in B(x,r)} \frac{1}{\mathfrak{m}(B(x,r))} = (V_n|x|^n)^{-1},$$

i.e., it reaches the supremum when $r = |x|$. Hence, the distribution function of $M(d\mu)$ is

$$\begin{aligned}(M(d\mu))_*(\alpha) =&\mathfrak{m}(\{x : |M(d\mu)(x)| > \alpha\}) = \mathfrak{m}(\{x : (V_n|x|^n)^{-1} > \alpha\}) \\ =&\mathfrak{m}(\{x : V_n|x|^n < \alpha^{-1}\}) = \mathfrak{m}(B(0,(V_n\alpha)^{-1/n})) \\ =&V_n(V_n\alpha)^{-1} = 1/\alpha.\end{aligned}$$

But we can always find a sequence $\{f_m(x)\}$ of positive integrable functions, whose L^1 norm is each 1, and which converges weakly to the measure $d\mu$. So we cannot expect an estimate essentially stronger than the estimate (ii) in Theorem 3.9, since, in the limit, a similar stronger version would have to hold for $M(d\mu)(x)$.

(2) It is useful, for certain applications, to observe that

$$A_p = O\left(\frac{1}{p-1}\right), \quad \text{as } p \to 1.$$

In contrast with the case $p > 1$, when $p = 1$ the mapping $f \mapsto Mf$ is not bounded on $L^1(\mathbb{R}^n)$. So the proof of the weak bound (ii) for Mf requires a less elementary arguments of geometric measure theory, like the Vitali covering lemma. In fact, we have

[3] The *Heine-Borel theorem* reads as follows: A set $K \subset \mathbb{R}^n$ is closed and bounded if and only if K is a compact set (i.e., every open cover of K has a finite subcover). In words, *any* covering of a compact set by a collection of open sets contains a *finite* sub-covering. For the proof, one can see the wiki: http://en.wikipedia.org/wiki/Heine%E2%80%93Borel_theorem.

Theorem 3.11. *If $f \in L^1(\mathbb{R}^n)$ is not identically zero, then Mf is never integrable on the whole of $\mathbb{R}^n$, i.e., $Mf \notin L^1(\mathbb{R}^n)$.*

Proof. We can choose an N large enough such that

$$\int_{B(0,N)} |f(x)|dx \geqslant \frac{1}{2}\|f\|_1.$$

Then, we take an $x \in \mathbb{R}^n$ such that $|x| \geqslant N$. Let $r = 2(|x|+N)$, we have

$$\begin{aligned} Mf(x) \geqslant & \frac{1}{\mathfrak{m}(B(x,r))}\int_{B(x,r)} |f(y)|dy = \frac{1}{V_n(2(|x|+N))^n}\int_{B(x,r)} |f(y)|dy \\ \geqslant & \frac{1}{V_n(2(|x|+N))^n}\int_{B(0,N)} |f(y)|dy \geqslant \frac{1}{2V_n(2(|x|+N))^n}\|f\|_1 \\ \geqslant & \frac{1}{2V_n(4|x|)^n}\|f\|_1. \end{aligned}$$

It follows that for sufficiently large $|x|$, we have

$$Mf(x) \geqslant c|x|^{-n}, \quad c = (2V_n 4^n)^{-1}\|f\|_1.$$

This implies that $Mf \notin L^1(\mathbb{R}^n)$. □

Moreover, even if we limit our consideration to any bounded subset of $\mathbb{R}^n$, then the integrability of Mf holds only if stronger conditions than the integrability of f are required. In fact, we have

Theorem 3.12. *Let E be a bounded subset of $\mathbb{R}^n$. If $f \ln^+ |f| \in L^1(\mathbb{R}^n)$ and $\operatorname{supp} f \subset E$, then*

$$\int_E Mf(x)dx \leqslant 2\mathfrak{m}(E) + C\int_E |f(x)|\ln^+ |f(x)|dx,$$

where $\ln^+ t = \max(\ln t, 0)$.

Proof. By Theorem 2.16, it follows that

$$\begin{aligned} \int_E Mf(x)dx =& 2\int_0^\infty \mathfrak{m}(\{x \in E : Mf(x) > 2\alpha\})d\alpha \\ =& 2\left(\int_0^1 + \int_1^\infty\right)\mathfrak{m}(\{x \in E : Mf(x) > 2\alpha\})d\alpha \\ \leqslant & 2\mathfrak{m}(E) + 2\int_1^\infty \mathfrak{m}(\{x \in E : Mf(x) > 2\alpha\})d\alpha. \end{aligned}$$

Decompose f as $f_1 + f_2$, where $f_1 = f\chi_{\{x:|f(x)|>\alpha\}}$ and $f_2 = f - f_1$. Then, by Theorem 3.4, it follows that

$$Mf_2(x) \leqslant \|Mf_2\|_\infty \leqslant \|f_2\|_\infty \leqslant \alpha,$$

which yields

$$\{x \in E : Mf(x) > 2\alpha\} \subset \{x \in E : Mf_1(x) > \alpha\}.$$

Hence, by Theorem 3.9, we have

$$\begin{aligned} &\int_1^\infty \mathfrak{m}(\{x \in E : Mf(x) > 2\alpha\})d\alpha \leqslant \int_1^\infty \mathfrak{m}(\{x \in E : Mf_1(x) > \alpha\})d\alpha \\ \leqslant & C\int_1^\infty \frac{1}{\alpha}\int_{\{x\in E:|f(x)|>\alpha\}} |f(x)|dxd\alpha \leqslant C\int_E |f(x)|\int_1^{\max(1,|f(x)|)} \frac{d\alpha}{\alpha}dx \\ =& C\int_E |f(x)|\ln^+ |f(x)|dx. \end{aligned}$$

This completes the proof. □

As a corollary of Theorem 3.9, we have the differentiability almost everywhere of the integral, expressed in (3.2).

Theorem 3.13 (Lebesgue differentiation theorem)**.** *If $f \in L^p(\mathbb{R}^n)$, $p \in [1,\infty]$, or more generally if f is locally integrable (i.e., $f \in L^1_{loc}(\mathbb{R}^n)$), then*

$$\lim_{r\to 0} \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} f(y)dy = f(x), \quad \text{for a.e. } x. \tag{3.9}$$

Proof. We first consider the case $p = 1$. It suffices to show that for each $\alpha > 0$, the set

$$E_\alpha = \left\{ x : \limsup_{r\to 0} \left| \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} f(y)dy - f(x) \right| > 2\alpha \right\}$$

has measure zero, because this assertion then guarantees that the set $E = \bigcup_{k=1}^\infty E_{1/k}$ has measure zero, and the limit in (3.9) holds at all points of E^c.

Fix α, since the continuous functions of compact support are dense in $L^1(\mathbb{R}^n)$, for each $\varepsilon > 0$ we may select a continuous function g of compact support with $\|f - g\|_1 < \varepsilon$. As we remarked earlier, the continuity of g implies that

$$\lim_{r\to 0} \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} g(y)dy = g(x), \quad \text{for all } x.$$

Since we may write the difference $\frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} f(y)dy - f(x)$ as

$$\frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} (f(y) - g(y))dy + \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} g(y)dy - g(x) + g(x) - f(x),$$

we find that

$$\limsup_{r\to 0} \left| \frac{1}{\mathrm{m}(B(x,r))} \int_{B(x,r)} f(y)dy - f(x) \right| \leqslant M(f-g)(x) + |g(x) - f(x)|.$$

Consequently, if

$$F_\alpha = \{x : M(f-g)(x) > \alpha\} \quad \text{and} \quad G_\alpha = \{x : |f(x) - g(x)| > \alpha\},$$

then $E_\alpha \subset F_\alpha \cup G_\alpha$, because if u_1 and u_2 are positive, then $u_1 + u_2 > 2\alpha$ only if $u_i > \alpha$ for at least one u_i.

On the one hand, Tchebychev's inequality[4] yields

$$\mathrm{m}(G_\alpha) \leqslant \frac{1}{\alpha} \|f - g\|_1,$$

and on the other hand, the weak type estimate for the maximal function gives

$$\mathrm{m}(F_\alpha) \leqslant \frac{3^n}{\alpha} \|f - g\|_1.$$

Since the function g was selected so that $\|f - g\|_1 < \varepsilon$, we get

$$\mathrm{m}(E_\alpha) \leqslant \frac{3^n}{\alpha}\varepsilon + \frac{1}{\alpha}\varepsilon = \frac{3^n + 1}{\alpha}\varepsilon.$$

Since ε is arbitrary, we must have $\mathrm{m}(E_\alpha) = 0$, and the proof for $p = 1$ is completed.

Indeed, the limit in the theorem is taken over balls that shrink to the point x, so the behavior of f far from x is irrelevant. Thus, we expect the result to remain valid if we simply assume integrability of f on every ball. Clearly, the conclusion holds under the weaker assumption that f is locally integrable.

---

[4] *Tchebychev inequality* (also spelled as Chebyshev's inequality): Suppose $f \geqslant 0$, and f is integrable. If $\alpha > 0$ and $E_\alpha = \{x \in \mathbb{R}^n : f(x) > \alpha\}$, then

$$\mathrm{m}(E_\alpha) \leqslant \frac{1}{\alpha} \int_{\mathbb{R}^n} f dx.$$

For the remained cases $p \in (1, \infty]$, we have by Hölder inequality, for any ball B,

$$\int_B |f(x)|dx \leqslant \|f\|_{L^p(B)} \|1\|_{L^{p'}(B)} \leqslant \mathfrak{m}(B)^{1/p'} \|f\|_p.$$

Thus, $f \in L^1_{loc}(\mathbb{R}^n)$ and then the conclusion is valid for $p \in (1, \infty]$. Therefore, we complete the proof of the theorem. □

By the Lebesgue differentiation theorem, we have

Theorem 3.14. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then*

$$|f(x)| \leqslant Mf(x), \quad a.e.\ x \in \mathbb{R}^n.$$

Combining with the maximal function theorem (i.e., Theorem 3.9), we get

Corollary 3.15. *If $f \in L^p(\mathbb{R}^n)$, $p \in (1, \infty]$, then we have*

$$\|f\|_p \leqslant \|Mf\|_p \leqslant A_p \|f\|_p.$$

As an application, we prove the (Gagliardo-Nirenberg-) Sobolev inequality by using the maximal function theorem for the case $1 < p < n$. We note that the inequality also holds for the case $p = 1$ and one can see [Eva98, p.263-264] for the proof.

Theorem 3.16 ((Gagliardo-Nirenberg-) Sobolev inequality)**.** *Let $p \in (1, n)$ and its Sobolev conjugate $p^* = np/(n-p)$. Then for $f \in \mathscr{D}(\mathbb{R}^n)$, we have*

$$\|f\|_{p^*} \leqslant C \|\nabla f\|_p,$$

where C depends only on n and p.

Proof. Since $f \in \mathscr{D}(\mathbb{R}^n)$, we have

$$f(x) = -\int_0^\infty \frac{\partial}{\partial r} f(x + rz) dr,$$

where $z \in S^{n-1}$. Integrating this over the whole unit sphere surface S^{n-1} yields

$$\begin{aligned}
\omega_{n-1} f(x) = \int_{S^{n-1}} f(x) d\sigma(z) &= -\int_{S^{n-1}} \int_0^\infty \frac{\partial}{\partial r} f(x + rz) dr d\sigma(z) \\
&= -\int_{S^{n-1}} \int_0^\infty \nabla f(x + rz) \cdot z dr d\sigma(z) \\
&= -\int_0^\infty \int_{S^{n-1}} \nabla f(x + rz) \cdot z d\sigma(z) dr.
\end{aligned}$$

Changing variables $y = x + rz$, $d\sigma(z) = r^{-(n-1)} d\sigma(y)$, $z = (y - x)/|y - x|$ and $r = |y - x|$, we get

$$\omega_{n-1} f(x) = -\int_0^\infty \int_{\partial B(x,r)} \nabla f(y) \cdot \frac{y - x}{|y - x|^n} d\sigma(y) dr = -\int_{\mathbb{R}^n} \nabla f(y) \cdot \frac{y - x}{|y - x|^n} dy,$$

which implies that

$$|f(x)| \leqslant \frac{1}{\omega_{n-1}} \int_{\mathbb{R}^n} \frac{|\nabla f(y)|}{|y - x|^{n-1}} dy.$$

We split this integral into two parts as $\int_{\mathbb{R}^n} = \int_{B(x,r)} + \int_{\mathbb{R}^n \setminus B(x,r)}$. For the first part, we have

$$\begin{aligned}
\frac{1}{\omega_{n-1}} \int_{B(x,r)} \frac{|\nabla f(y)|}{|x - y|^{n-1}} dy =& \frac{1}{\omega_{n-1}} \sum_{k=0}^\infty \int_{B(x, 2^{-k} r) \setminus B(x, 2^{-k-1} r)} \frac{|\nabla f(y)|}{|x - y|^{n-1}} dy \\
\leqslant& \frac{1}{\omega_{n-1}} \sum_{k=0}^\infty \int_{B(x, 2^{-k} r) \setminus B(x, 2^{-k-1} r)} \frac{|\nabla f(y)|}{(2^{-k-1} r)^{n-1}} dy
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{k=0}^{\infty} \frac{2^{-k}r}{nV_n 2^{-k}r} \int_{B(x,2^{-k}r)} 2^{n-1} \frac{|\nabla f(y)|}{(2^{-k}r)^{n-1}} dy \\
&\leqslant \frac{1}{n} \sum_{k=0}^{\infty} 2^{-k+n-1} r \frac{1}{\mathrm{m}(B(x,2^{-k}r))} \int_{B(x,2^{-k}r)} |\nabla f(y)| dy \\
&\leqslant \frac{2^{n-1}}{n} r M(\nabla f)(x) \sum_{k=0}^{\infty} 2^{-k} = \frac{2^n}{n} r M(\nabla f)(x).
\end{aligned}
$$

For the second part, by Hölder inequality, we get for $1 < p < n$

$$
\begin{aligned}
&\int_{\mathbb{R}^n \setminus B(x,r)} \frac{|\nabla f(y)|}{|x-y|^{n-1}} dy \\
\leqslant &\left(\int_{\mathbb{R}^n \setminus B(x,r)} |\nabla f(y)|^p dy \right)^{1/p} \left(\int_{\mathbb{R}^n \setminus B(x,r)} |x-y|^{(1-n)p'} dy \right)^{1/p'} \\
\leqslant &\left(\omega_{n-1} \int_r^{\infty} \rho^{(1-n)p'} \rho^{n-1} d\rho \right)^{1/p'} \|\nabla f\|_p \\
= &\left(\frac{(p-1)\omega_{n-1}}{n-p} \right)^{1/p'} r^{1-n/p} \|\nabla f\|_p.
\end{aligned}
$$

Choosing $r = \frac{(p-1)^{(p-1)/n}}{(n-p)^{(p-1)/n} \omega_{n-1}^{1/n} 2^p} \left(\frac{n\|\nabla f\|_p}{M(\nabla f)(x)} \right)^{p/n}$ satisfying

$$
2^n r M(\nabla f)(x) = \frac{n}{\omega_{n-1}} \left(\frac{(p-1)\omega_{n-1}}{n-p} \right)^{1/p'} r^{1-n/p} \|\nabla f\|_p,
$$

then we get

$$
|f(x)| \leqslant C \|\nabla f\|_p^{p/n} (M(\nabla f)(x))^{1-p/n}.
$$

Thus, by part (iii) in Theorem 3.9, we obtain for $1 < p < n$

$$
\|f\|_{p^*} \leqslant C \|\nabla f\|_p^{p/n} \|M(\nabla f)\|_{p^*(1-p/n)}^{1-p/n} = C \|\nabla f\|_p^{p/n} \|M(\nabla f)\|_p^{1-p/n} \leqslant C \|\nabla f\|_p.
$$

This completes the proof. □

3.3 Calderón-Zygmund decomposition

Applying Lebesgue differentiation theorem, we give a decomposition of $\mathbb{R}^n$, called Calderón-Zygmund decomposition, which is extremely useful in harmonic analysis.

Theorem 3.17 (Calderón-Zygmund decomposition of $\mathbb{R}^n$)**.** *Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. Then there exists a decomposition of $\mathbb{R}^n$ such that*

(i) *$\mathbb{R}^n = F \cup \Omega$, $F \cap \Omega = \varnothing$.*

(ii) *$|f(x)| \leqslant \alpha$ for a.e. $x \in F$.*

(iii) *Ω is the union of cubes, $\Omega = \bigcup_k Q_k$, whose interiors are disjoint and edges parallel to the coordinate axes, and such that for each Q_k*

$$
\alpha < \frac{1}{\mathrm{m}(Q_k)} \int_{Q_k} |f(x)| dx \leqslant 2^n \alpha. \tag{3.10}
$$

Proof. We decompose $\mathbb{R}^n$ into a mesh of equal cubes $Q_k^{(0)}$ $(k = 1, 2, \cdots)$, whose interiors are disjoint and edges parallel to the coordinate axes, and whose common diameter is so large that

$$\frac{1}{\mathfrak{m}(Q_k^{(0)})}\int_{Q_k^{(0)}}|f(x)|dx \leqslant \alpha, \tag{3.11}$$

since $f \in L^1$.

Split each $Q_k^{(0)}$ into 2^n congruent cubes. These we denote by $Q_k^{(1)}$, $k = 1, 2, \cdots$. There are two possibilities:

$$\text{either } \frac{1}{\mathfrak{m}(Q_k^{(1)})}\int_{Q_k^{(1)}}|f(x)|dx \leqslant \alpha, \text{ or } \frac{1}{\mathfrak{m}(Q_k^{(1)})}\int_{Q_k^{(1)}}|f(x)|dx > \alpha.$$

In the first case, we split $Q_k^{(1)}$ again into 2^n congruent cubes to get $Q_k^{(2)}$ $(k = 1, 2, \cdots)$. In the second case, we have

$$\alpha < \frac{1}{\mathfrak{m}(Q_k^{(1)})}\int_{Q_k^{(1)}}|f(x)|dx \leqslant \frac{1}{2^{-n}\mathfrak{m}(Q_{\tilde{k}}^{(0)})}\int_{Q_{\tilde{k}}^{(0)}}|f(x)|dx \leqslant 2^n\alpha$$

in view of (3.11) where $Q_k^{(1)}$ is split from $Q_{\tilde{k}}^{(0)}$, and then we take $Q_k^{(1)}$ as one of the cubes Q_k.

A repetition of this argument shows that if $x \notin \Omega =: \bigcup_{k=1}^\infty Q_k$ then $x \in Q_{k_j}^{(j)}$ $(j = 0, 1, 2, \cdots)$ for which

$$\mathfrak{m}(Q_{k_j}^{(j)}) \to 0 \text{ as } j \to \infty, \quad \text{and } \frac{1}{\mathfrak{m}(Q_{k_j}^{(j)})}\int_{Q_{k_j}^{(j)}}|f(x)|dx \leqslant \alpha \quad (j = 0, 1, \cdots).$$

Thus $|f(x)| \leqslant \alpha$ a.e. $x \in F = \Omega^c$ by a variation of the Lebesgue differentiation theorem. Thus, we complete the proof. □

We now state an immediate corollary.

Corollary 3.18. *Suppose f, α, F, Ω and Q_k have the same meaning as in Theorem 3.17. Then there exists two constants A and B (depending only on the dimension n), such that (i) and (ii) of Theorem 3.17 hold and*

(a) $\mathfrak{m}(\Omega) \leqslant \dfrac{A}{\alpha}\|f\|_1$,

(b) $\dfrac{1}{\mathfrak{m}(Q_k)}\displaystyle\int_{Q_k}|f|dx \leqslant B\alpha$.

Proof. In fact, by (3.10) we can take $B = 2^n$, and also because of (3.10)

$$\mathfrak{m}(\Omega) = \sum_k \mathfrak{m}(Q_k) < \frac{1}{\alpha}\int_\Omega |f(x)|dx \leqslant \frac{1}{\alpha}\|f\|_1.$$

This proves the corollary with $A = 1$ and $B = 2^n$. □

It is possible however to give another proof of this corollary without using Theorem 3.17 from which it was deduced, but by using the maximal function theorem (Theorem 3.9) and also the theorem about the decomposition of an arbitrary open set as a union of disjoint cubes. This more indirect method of proof has the advantage of *clarifying the roles of the sets F and Ω into which $\mathbb{R}^n$ was divided.*

Another proof of the corollary. We know that in F, $|f(x)| \leqslant \alpha$, but this fact does not determine F. The set F is however determined, in effect, by the fact that the maximal function satisfies $Mf(x) \leqslant \alpha$ on it. So we choose $F = \{x : Mf(x) \leqslant \alpha\}$ and $\Omega = E_\alpha = \{x : Mf(x) > \alpha\}$. Then by Theorem 3.9, part (b) we know that $\mathfrak{m}(\Omega) \leqslant \frac{3^n}{\alpha}\|f\|_1$. Thus, we can take $A = 3^n$.

Since by definition F is closed, we can choose cubes Q_k according to Theorem 3.2, such that $\Omega = \bigcup_k Q_k$, and whose diameters are approximately proportional to their distances from F. Let Q_k then be one of these cubes, and p_k a point of F such that

$$\operatorname{dist}(F, Q_k) = \operatorname{dist}(p_k, Q_k).$$

Let B_k be the smallest ball whose center is p_k and which contains the interior of Q_k. Let us set

$$\gamma_k = \frac{\mathrm{m}(B_k)}{\mathrm{m}(Q_k)}.$$

We have, because $p_k \in \{x : Mf(x) \leqslant \alpha\}$, that

$$\alpha \geqslant Mf(p_k) \geqslant \frac{1}{\mathrm{m}(B_k)} \int_{B_k} |f(x)|dx \geqslant \frac{1}{\gamma_k \mathrm{m}(Q_k)} \int_{Q_k} |f(x)|dx.$$

Thus, we can take a upper bound of γ_k as the value of B.

The elementary geometry and the inequality (iii) of Theorem 3.2 then show that

$$\begin{aligned}\operatorname{radius}(B_k) \leqslant & \operatorname{dist}(p_k, Q_k) + \operatorname{diam}(Q_k) = \operatorname{dist}(F, Q_k) + \operatorname{diam}(Q_k) \\ \leqslant & (c_2 + 1) \operatorname{diam}(Q_k),\end{aligned}$$

and so

$$\mathrm{m}(B_k) = V_n(\operatorname{radius}(B_k))^n \leqslant V_n(c_2+1)^n(\operatorname{diam}(Q_k))^n = V_n(c_2+1)^n n^{n/2} \mathrm{m}(Q_k),$$

since $\mathrm{m}(Q_k) = (\operatorname{diam}(Q_k)/\sqrt{n})^n$. Thus, $\gamma_k \leqslant V_n(c_2+1)^n n^{n/2}$ for all k. Thus, we complete the proof with $A = 3^n$ and $B = V_n(c_2+1)^n n^{n/2}$. □

Remark 3.19. (1) Notice that this second proof of the lemma also rewarded us with an unexpected benefit: the cubes Q_k are at a distance from F comparable to their diameters.

(2) Theorem 3.17 may be used to give another proof of the fundamental inequality for the maximal function in part (ii) of Theorem 3.9. (See [Ste70, §5.1, p.22–23] for more details.)

The Calderón-Zygmund decomposition is a key step in the real-variable analysis of singular integrals. The idea behind this decomposition is that it is often useful to split an arbitrary integrable function into its “small” and “large” parts, and then use different techniques to analyze each part.

The scheme is roughly as follows. Given a function f and an altitude α, we write $f = g + b$, where g is called the good function of the decomposition since it is both integrable and bounded; hence the letter g. The function b is called the bad function since it contains the singular part of f (hence the letter b), but it is carefully chosen to have mean value zero. To obtain the decomposition $f = g + b$, one might be tempted to “cut” f at the height α; however, this is not what works. Instead, one bases the decomposition on the set where the maximal function of f has height α.

Indeed, the Calderón-Zygmund decomposition on $\mathbb{R}^n$ may be used to deduce the Calderón-Zygmund decomposition on functions. The later is a very important tool in harmonic analysis.

Theorem 3.20 (Calderón-Zygmund decomposition for functions)**.** *Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. Then there exist functions g and b on $\mathbb{R}^n$ such that $f = g + b$ and*

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

(ii) $b = \sum_j b_j$, *where each b_j is supported in a dyadic cube Q_j satisfying $\int_{Q_j} b_j(x)dx = 0$ and $\|b_j\|_1 \leqslant 2^{n+1}\alpha \mathrm{m}(Q_j)$. Furthermore, the cubes Q_j and Q_k have disjoint interiors when $j \neq k$.*

(iii) $\sum_j \mathrm{m}(Q_j) \leqslant \alpha^{-1}\|f\|_1$.

Proof. Applying Corollary 3.18 (with $A = 1$ and $B = 2^n$), we have

1) $\mathbb{R}^n = F \cup \Omega$, $F \cap \Omega = \varnothing$;

2) $|f(x)| \leqslant \alpha$, a.e. $x \in F$;

3) $\Omega = \bigcup_{j=1}^{\infty} Q_j$, with the interiors of the Q_j mutually disjoint;

4) $m(\Omega) \leqslant \alpha^{-1} \int_{\mathbb{R}^n} |f(x)|dx$, and $\alpha < \frac{1}{m(Q_j)} \int_{Q_j} |f(x)|dx \leqslant 2^n\alpha$.

Now define

$$b_j = \left(f - \frac{1}{\mathfrak{m}(Q_j)} \int_{Q_j} f dx \right) \chi_{Q_j},$$

$b = \sum_j b_j$ and $g = f - b$. Consequently,

$$\begin{aligned}\int_{Q_j} |b_j| dx \leqslant & \int_{Q_j} |f(x)|dx + \mathfrak{m}(Q_j) \left| \frac{1}{\mathfrak{m}(Q_j)} \int_{Q_j} f(x) dx \right| \\ \leqslant & 2 \int_{Q_j} |f(x)|dx \leqslant 2^{n+1}\alpha \mathfrak{m}(Q_j),\end{aligned}$$

which proves $\|b_j\|_1 \leqslant 2^{n+1}\alpha\mathfrak{m}(Q_j)$.

Next, we need to obtain the estimates on g. Write $\mathbb{R}^n = \cup_j Q_j \cup F$, where F is the closed set obtained by Corollary 3.18. Since $b = 0$ on F and $f - b_j = \frac{1}{\mathfrak{m}(Q_j)} \int_{Q_j} f(x)dx$, we have

$$g = \begin{cases} f, & \text{on } F, \\ \dfrac{1}{\mathfrak{m}(Q_j)} \displaystyle\int_{Q_j} f(x)dx, & \text{on } Q_j. \end{cases} \tag{3.12}$$

On the cube Q_j, g is equal to the constant $\frac{1}{\mathfrak{m}(Q_j)} \int_{Q_j} f(x)dx$, and this is bounded by $2^n\alpha$ by 4). Then by 2), we can get $\|g\|_\infty \leqslant 2^n\alpha$. Finally, it follows from (3.12) that $\|g\|_1 \leqslant \|f\|_1$. This completes the proof. □

As an application of Calderón-Zygmund decomposition and Marcinkiewicz interpolation theorem, we now prove the Fefferman-Stein inequality on the Hardy-Littlewood maximal operator M.

Theorem 3.21 (Fefferman-Stein inequality). *For $p \in (1, \infty)$, there exists a constant $C = C_{n,p}$ such that, for any measurable function on $\mathbb{R}^n$, $\varphi(x) \geqslant 0$ and f, we have the inequality*

$$\int_{\mathbb{R}^n} (Mf(x))^p \varphi(x) dx \leqslant C \int_{\mathbb{R}^n} |f(x)|^p M\varphi(x) dx. \tag{3.13}$$

Proof. Except when $M\varphi(x) = \infty$ a.e., in which case (3.13) holds trivially, $M\varphi$ is the density of a positive measure μ. Thus, we may assume that $M\varphi(x) < \infty$ a.e. $x \in \mathbb{R}^n$ and $M\varphi(x) > 0$. If we denote

$$d\mu(x) = M\varphi(x)dx \quad \text{and} \quad d\nu(x) = \varphi(x)dx,$$

then by the Marcinkiewicz interpolation theorem in order to get (3.13), it suffices to prove that M is both of type $(L^\infty(\mu), L^\infty(\nu))$ and of weak type $(L^1(\mu), L^1(\nu))$.

Let us first show that M is of type $(L^\infty(\mu), L^\infty(\nu))$. In fact, if $\|f\|_{L^\infty(\mu)} \leqslant \alpha$, then

$$\int_{\{x\in\mathbb{R}^n : |f(x)|>\alpha\}} M\varphi(x)dx = \mu(\{x \in \mathbb{R}^n : |f(x)| > \alpha\}) = 0.$$

Since $M\varphi(x) > 0$ for any $x \in \mathbb{R}^n$, we have $\mathfrak{m}(\{x \in \mathbb{R}^n : |f(x)| > \alpha\}) = 0$, equivalently, $|f(x)| \leqslant \alpha$ a.e. $x \in \mathbb{R}^n$. Thus, $Mf(x) \leqslant \alpha$ a.e. $x \in \mathbb{R}^n$ and this follows $\|Mf\|_{L^\infty(\nu)} \leqslant \alpha$. Therefore, $\|Mf\|_{L^\infty(\nu)} \leqslant \|f\|_{L^\infty(\mu)}$.

Before proving that M is also of weak type $(L^1(\mu), L^1(\nu))$, we give the following lemma.

Lemma 3.22. *Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. If the sequence $\{Q_k\}$ of cubes is chosen from the Calderón-Zygmund decomposition of $\mathbb{R}^n$ for f and $\alpha > 0$, then*

$$\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha\} \subset \bigcup_k Q_k^*,$$

where $Q_k^* = 2Q_k$. *Then we have*

$$\mathfrak{m}(\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha\}) \leqslant 2^n \sum_k \mathfrak{m}(Q_k).$$

Proof. Suppose that $x \notin \bigcup_k Q_k^*$. Then there are two cases for any cube Q with the center x. If $Q \subset F := \mathbb{R}^n \setminus \bigcup_k Q_k$, then

$$\frac{1}{\mathfrak{m}(Q)} \int_Q |f(x)| dx \leqslant \alpha.$$

If $Q \cap Q_k \neq \varnothing$ for some k, then it is easy to check that $Q_k \subset 3Q$, and

$$\bigcup_k \{Q_k : Q_k \cap Q \neq \varnothing\} \subset 3Q.$$

Hence, we have

$$\begin{aligned}
\int_Q |f(x)| dx \leqslant & \int_{Q \cap F} |f(x)| dx + \sum_{Q_k \cap Q \neq \varnothing} \int_{Q_k} |f(x)| dx \\
\leqslant & \alpha \mathfrak{m}(Q) + \sum_{Q_k \cap Q \neq \varnothing} 2^n \alpha \mathfrak{m}(Q_k) \\
\leqslant & \alpha \mathfrak{m}(Q) + 2^n \alpha \mathfrak{m}(3Q) \\
\leqslant & 7^n \alpha \mathfrak{m}(Q).
\end{aligned}$$

Thus we know that $M'f(x) \leqslant 7^n\alpha$ for any $x \notin \bigcup_k Q_k^*$, and it yields that

$$\mathfrak{m}(\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha\}) \leqslant \mathfrak{m}\left(\bigcup_k Q_k^*\right) = 2^n \sum_k \mathfrak{m}(Q_k).$$

We complete the proof of the lemma. $\square$

Let us return to the proof of weak type $(L^1(\mu), L^1(\nu))$. We need to prove that there exists a constant C such that for any $\alpha > 0$ and $f \in L^1(\mu)$

$$\begin{aligned}
\int_{\{x \in \mathbb{R}^n : Mf(x) > \alpha\}} \varphi(x) dx = & \nu(\{x \in \mathbb{R}^n : Mf(x) > \alpha\}) \\
\leqslant & \frac{C}{\alpha} \int_{\mathbb{R}^n} |f(x)| M\varphi(x) dx.
\end{aligned} \tag{3.14}$$

We may assume that $f \in L^1(\mathbb{R}^n)$. In fact, if we take $f_k = |f|\chi_{B(0,k)}$, then $f_k \in L^1(\mathbb{R}^n)$, $0 \leqslant f_k(x) \leqslant f_{k+1}(x)$ for $x \in \mathbb{R}^n$ and $k = 1, 2, \cdots$. Moreover, $\lim_{k\to\infty} f_k(x) = |f(x)|$.

By the pointwise equivalence of M and M', there exists $c_n > 0$ such that $Mf(x) \leqslant c_n M'f(x)$ for all $x \in \mathbb{R}^n$. Applying the Calderón-Zygmund decomposition on $\mathbb{R}^n$ for f and $\alpha' = \alpha/(c_n 7^n)$, we get a sequence $\{Q_k\}$ of cubes satisfying

$$\alpha' < \frac{1}{\mathfrak{m}(Q_k)} \int_{Q_k} |f(x)| dx \leqslant 2^n \alpha'.$$

By Lemma 3.22 and the pointwise equivalence of M and M'', we have that

$$\begin{aligned}
& \int_{\{x \in \mathbb{R}^n : Mf(x) > \alpha\}} \varphi(x) dx \\
\leqslant & \int_{\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha'\}} \varphi(x) dx \\
\leqslant & \int_{\bigcup_k Q_k^*} \varphi(x) dx \leqslant \sum_k \int_{Q_k^*} \varphi(x) dx \\
\leqslant & \sum_k \left(\frac{1}{\mathfrak{m}(Q_k)} \int_{Q_k^*} \varphi(x) dx\right) \left(\frac{1}{\alpha'} \int_{Q_k} |f(y)| dy\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{c_n 7^n}{\alpha} \sum_k \int_{Q_k} |f(y)| \left(\frac{2^n}{\mathrm{m}(Q_k^*)} \int_{Q_k^*} \varphi(x) dx \right) dy \\
&\leqslant \frac{c_n 14^n}{\alpha} \sum_k \int_{Q_k} |f(y)| M'' \varphi(y) dy \\
&\leqslant \frac{C}{\alpha} \int_{\mathbb{R}^n} |f(y)| M \varphi(y) dy.
\end{aligned}
$$

Thus, M is of weak type $(L^1(\mu), L^1(\nu))$, and the Fefferman-Stein inequality can be obtained by applying the Marcinkiewicz interpolation theorem. $\square$

Chapter 4
Singular Integrals

4.1 Harmonic functions and Poisson equation

Among the most important of all PDEs are undoubtedly *Laplace equation*

$$\Delta u = 0 \tag{4.1}$$

and *Poisson equation*

$$-\Delta u = f. \tag{4.2}$$

In both (4.1) and (4.2), $x \in \Omega$ and the unknown is $u : \bar{\Omega} \to \mathbb{R}$, $u = u(x)$, where $\Omega \subset \mathbb{R}^n$ is a given open set. In (4.2), the function $f : \Omega \to \mathbb{R}$ is also given. Remember that the Laplacian of u is $\Delta u = \sum_{k=1}^n \partial_{x_k}^2 u$.

Definition 4.1. A C^2 function u satisfying (4.1) is called a *harmonic function*.

Now, we derive a fundamental solution of Laplace's equation. One good strategy for investigating any PDEs is first to identify some explicit solutions and then, provided the PDE is linear, to assemble more complicated solutions out of the specific ones previously noted. Furthermore, in looking for explicit solutions it is often wise to restrict attention to classes of functions with certain symmetry properties. Since Laplace equation is invariant under rotations, it consequently seems advisable to search first for radial solutions, that is, functions of $r = |x|$. Let us therefore attempt to find a solution u of Laplace equation (4.1) in $\Omega = \mathbb{R}^n$, having the form

$$u(x) = v(r),$$

where $r = |x|$ and v is to be selected (if possible) so that $\Delta u = 0$ holds. First note for $k = 1, \cdots, n$ that

$$\frac{\partial r}{\partial x_k} = \frac{x_k}{r}, \quad x \neq 0.$$

We thus have

$$\partial_{x_k} u = v'(r)\frac{x_k}{r}, \quad \partial_{x_k}^2 u = v''(r)\frac{x_k^2}{r^2} + v'(r)\left(\frac{1}{r} - \frac{x_k^2}{r^3}\right)$$

for $k = 1, \cdots, n$, and so

$$\Delta u = v''(r) + \frac{n-1}{r} v'(r).$$

Hence $\Delta u = 0$ if and only if

$$v'' + \frac{n-1}{r} v' = 0. \tag{4.3}$$

If $v' \neq 0$, we deduce

$$(\ln v')' = \frac{v''}{v'} = \frac{1-n}{r},$$

and hence $v'(r) = \frac{a}{r^{n-1}}$ for some constant a. Consequently, if $r > 0$, we have

$$v(r) = \begin{cases} b \ln r + c, & n = 2, \\ \dfrac{b}{r^{n-2}} + c, & n \geqslant 3, \end{cases}$$

where b and c are constants.

These considerations motivate the following

Definition 4.2. The function

$$\Phi(x) := \begin{cases} -\dfrac{1}{2\pi} \ln |x|, & n = 2, \\ \dfrac{1}{n(n-2)V_n} \dfrac{1}{|x|^{n-2}}, & n \geqslant 3, \end{cases} \tag{4.4}$$

defined for $x \in \mathbb{R}^n$, $x \neq 0$, is the fundamental solution of Laplace equation.

The reason for the particular choices of the constants in (4.4) will be apparent in a moment.

We will sometimes slightly abuse notation and write $\Phi(x) = \Phi(|x|)$ to emphasize that the fundamental solution is radial. Observe also that we have the estimates

$$|\nabla \Phi(x)| \leqslant \frac{C}{|x|^{n-1}}, \quad |\Delta \Phi(x)| \leqslant \frac{C}{|x|^n}, \quad (x \neq 0) \tag{4.5}$$

for some constant $C > 0$.

By construction, the function $x \mapsto \Phi(x)$ is harmonic for $x \neq 0$. If we shift the origin to a new point y, the PDE (4.1) is unchanged; and so $x \mapsto \Phi(x-y)$ is also harmonic as a function of x for $x \neq y$. Let us now take $f : \mathbb{R}^n \to \mathbb{R}$ and note that the mapping $x \mapsto \Phi(x-y)f(y)$ $(x \neq y)$ is harmonic for each point $y \in \mathbb{R}^n$, and thus so is the sum of finitely many such expression built for different points y. This reasoning might suggest that the convolution

$$u(x) = \int_{\mathbb{R}^n} \Phi(x-y) f(y) dy = \begin{cases} -\dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}^2} (\ln |x-y|) f(y) dy, & n = 2, \\ \dfrac{1}{n(n-2)V_n} \displaystyle\int_{\mathbb{R}^n} \dfrac{f(y)}{|x-y|^{n-2}} dy, & n \geqslant 3 \end{cases} \tag{4.6}$$

would solve Laplace equation (4.1). *However, this is wrong*: we *cannot* just compute

$$\Delta u(x) = \int_{\mathbb{R}^n} \Delta_x \Phi(x-y) f(y) dy = 0. \tag{4.7}$$

Indeed, as intimated by estimate (4.5), $\Delta \Phi(x-y)$ is *not* summable near the singularity at $y = x$, and so the differentiation under the integral sign above is unjustified (and incorrect). We must proceed more carefully in calculating Δu.

Let us for simplicity now assume $f \in C_c^2(\mathbb{R}^n)$, that is, f is twice continuously differentiable, with compact support.

Theorem 4.3 (Solving Poisson equation)**.** *Let $f \in C_c^2(\mathbb{R}^n)$, define u by (4.6). Then $u \in C^2(\mathbb{R}^n)$ and $-\Delta u = f$ in $\mathbb{R}^n$.*

We consequently see that (4.6) provides us with a formula for a solution of Poisson's equation (4.2) in $\mathbb{R}^n$.

Proof. *Step 1: To show $u \in C^2(\mathbb{R}^n)$.* We have

$$u(x) = \int_{\mathbb{R}^n} \Phi(x-y) f(y) dy = \int_{\mathbb{R}^n} \Phi(y) f(x-y) dy,$$

hence

$$\frac{u(x+he_k)-u(x)}{h}=\int_{\mathbb{R}^n}\Phi(y)\left[\frac{f(x+he_k-y)-f(x-y)}{h}\right]dy,$$

where $h\neq 0$ and $e_k=(0,\cdots,1,\cdots,0)$, the 1 in the k^{th}-slot. But

$$\frac{f(x+he_k-y)-f(x-y)}{h}\to\frac{\partial f}{\partial x_k}(x-y)$$

uniformly on $\mathbb{R}^n$ as $h\to 0$, and thus

$$\frac{\partial u}{\partial x_k}(x)=\int_{\mathbb{R}^n}\Phi(y)\frac{\partial f}{\partial x_k}(x-y)dy,\quad k=1,\cdots,n.$$

Similarly,

$$\frac{\partial^2 u}{\partial x_k\partial x_j}(x)=\int_{\mathbb{R}^n}\Phi(y)\frac{\partial^2 f}{\partial x_k\partial x_j}(x-y)dy,\quad k,j=1,\cdots,n. \tag{4.8}$$

As the expression on the r.h.s. of (4.8) is continuous in the variable x, we see that $u\in C^2(\mathbb{R}^n)$.

Step 2: To prove the second part. Since Φ blows up at 0, we will need for subsequent calculations to isolate this singularity inside a small ball. So fix $\varepsilon>0$. Then

$$\Delta u(x)=\int_{B(0,\varepsilon)}\Phi(y)\Delta_x f(x-y)dy+\int_{\mathbb{R}^n\setminus B(0,\varepsilon)}\Phi(y)\Delta_x f(x-y)dy=:I_\varepsilon+J_\varepsilon. \tag{4.9}$$

Now

$$|I_\varepsilon|\leqslant C\|\Delta f\|_\infty\int_{B(0,\varepsilon)}|\Phi(y)|dy\leqslant\begin{cases}C\varepsilon^2(1+|\ln\varepsilon|), & n=2,\\ C\varepsilon^2, & n\geqslant 3,\end{cases} \tag{4.10}$$

since

$$\begin{aligned}\int_{B(0,\varepsilon)}|\ln|y||dy=&-2\pi\int_0^\varepsilon r\ln r dr=-\pi\left(r^2\ln r|_0^\varepsilon-\int_0^\varepsilon r dr\right)\\ =&-\pi(\varepsilon^2\ln\varepsilon-\varepsilon^2/2)\\ =&\pi\varepsilon^2|\ln\varepsilon|+\frac{\pi}{2}\varepsilon^2,\end{aligned}$$

for $\varepsilon\in(0,1]$ and $n=2$ by an integration by parts.

An integration by parts yields

$$\begin{aligned}J_\varepsilon=&\int_{\mathbb{R}^n\setminus B(0,\varepsilon)}\Phi(y)\Delta_x f(x-y)dy\\ =&\int_{\partial B(0,\varepsilon)}\Phi(y)\frac{\partial f}{\partial\nu}(x-y)d\sigma(y)-\int_{\mathbb{R}^n\setminus B(0,\varepsilon)}\nabla\Phi(y)\cdot\nabla_y f(x-y)dy\\ =:&K_\varepsilon+L_\varepsilon,\end{aligned} \tag{4.11}$$

where ν denotes the inward pointing unit normal along $\partial B(0,\varepsilon)$. We readily check

$$\begin{aligned}|K_\varepsilon|\leqslant&\|\nabla f\|_\infty\int_{\partial B(0,\varepsilon)}|\Phi(y)|d\sigma(y)\leqslant C|\Phi(\varepsilon)|\int_{\partial B(0,\varepsilon)}d\sigma(y)=C|\Phi(\varepsilon)|\varepsilon^{n-1}\\ \leqslant&\begin{cases}C\varepsilon|\ln\varepsilon|, & n=2,\\ C\varepsilon, & n\geqslant 3,\end{cases}\end{aligned} \tag{4.12}$$

since $\Phi(y)=\Phi(|y|)=\Phi(\varepsilon)$ on $\partial B(0,\varepsilon)=\{y\in\mathbb{R}^n:\ |y|=\varepsilon\}$.

We continue by integrating by parts once again in the term L_ε, to discover

$$\begin{aligned}L_\varepsilon=&-\int_{\partial B(0,\varepsilon)}\frac{\partial\Phi}{\partial\nu}(y)f(x-y)d\sigma(y)+\int_{\mathbb{R}^n\setminus B(0,\varepsilon)}\Delta\Phi(y)f(x-y)dy\\ =&-\int_{\partial B(0,\varepsilon)}\frac{\partial\Phi}{\partial\nu}(y)f(x-y)d\sigma(y),\end{aligned}$$

since Φ is harmonic away from the origin. Now, $\nabla\Phi(y) = -\frac{1}{nV_n}\frac{y}{|y|^n}$ for $y \neq 0$ and $\nu = \frac{-y}{|y|} = -\frac{y}{\varepsilon}$ on $\partial B(0,\varepsilon)$. Consequently, $\frac{\partial\Phi}{\partial\nu}(y) = \nu\cdot\nabla\Phi(y) = \frac{1}{nV_n\varepsilon^{n-1}}$ on $\partial B(0,\varepsilon)$. Since $nV_n\varepsilon^{n-1}$ is the surface area of the sphere $\partial B(0,\varepsilon)$, we have

$$\begin{aligned} L_\varepsilon &= -\frac{1}{nV_n\varepsilon^{n-1}}\int_{\partial B(0,\varepsilon)} f(x-y)d\sigma(y) \\ &= -\frac{1}{\mathrm{m}(\partial B(x,\varepsilon))}\int_{\partial B(x,\varepsilon)} f(y)d\sigma(y) \to -f(x) \quad \text{as } \varepsilon\to 0. \end{aligned} \tag{4.13}$$

by Lebesgue differentiation theorem.

Combining now (4.9)-(4.13) and letting $\varepsilon\to 0$, we find that $-\Delta u(x) = f(x)$, as asserted. □

Remark 4.4. (i) We sometimes write

$$-\Delta\Phi = \delta_0 \quad \text{in } \mathbb{R}^n,$$

where δ_0 denotes the Dirac measure on $\mathbb{R}^n$ giving unit mass to the point 0. Adopting this notation, we may formally compute

$$-\Delta u(x) = \int_{\mathbb{R}^n} -\Delta_x\Phi(x-y)f(y)dy = \int_{\mathbb{R}^n}\delta_x f(y)dy = f(x), \quad x\in\mathbb{R}^n,$$

in accordance with Theorem 4.3. This corrects the erroneous calculation (4.7).

(ii) Theorem 4.3 is in fact valid under far less stringent smoothness requirements for f: see [GT01].

Consider now an open set $\Omega\subset\mathbb{R}^n$ and suppose u is a harmonic function within Ω. We next derive the important mean-value formulas, which declare that $u(x)$ equals both the average of u over the sphere $\partial B(x,r)$ and the average of u over the entire ball $B(x,r)$, provided $B(x,r)\subset\Omega$.

Theorem 4.5 (Mean-value formula for harmonic functions)**.** *If $u\in C^2(\Omega)$ is harmonic, then for each ball $B(x,r)\subset\Omega$,*

$$u(x) = \frac{1}{\mathrm{m}(\partial B(x,r))}\int_{\partial B(x,r)} u(y)d\sigma(y) = \frac{1}{\mathrm{m}(B(x,r))}\int_{B(x,r)} u(y)dy.$$

Proof. Denote

$$f(r) = \frac{1}{\mathrm{m}(\partial B(x,r))}\int_{\partial B(x,r)} u(y)d\sigma(y) = \frac{1}{\omega_{n-1}}\int_{S^{n-1}} u(x+rz)d\sigma(z).$$

Obviously,

$$f'(r) = \frac{1}{\omega_{n-1}}\int_{S^{n-1}}\sum_{j=1}^n \partial_{x_j}u(x+rz)z_jd\sigma(z) = \frac{1}{\omega_{n-1}}\int_{S^{n-1}}\frac{\partial u}{\partial\nu}(x+rz)d\sigma(z),$$

where $\frac{\partial}{\partial\nu}$ denotes the differentiation w.r.t. the outward normal. Thus, by changes of variable

$$f'(r) = \frac{1}{\omega_{n-1}}\int_{\partial B(x,r)}\frac{\partial u}{\partial\nu}(y)d\sigma(y).$$

By Stokes theorem, we get

$$f'(r) = \frac{1}{\omega_{n-1}}\int_{B(x,r)}\Delta u(y)dy = 0.$$

Thus $f(r) = const$. Since $\lim_{r\to 0} f(r) = u(x)$, hence, $f(r) = u(x)$.

Next, observe that our employing polar coordinates gives

$$\int_{B(x,r)} u(y)dy = \int_0^r\left(\int_{\partial B(x,s)} u(y)d\sigma(y)\right)ds = \int_0^r \mathrm{m}(\partial B(x,s))u(x)ds$$

$$=u(x)\int_0^r nV_n s^{n-1}ds = V_n r^n u(x).$$

This completes the proof. □

Theorem 4.6 (Converse to mean-value property)**.** *If $u \in C^2(\Omega)$ satisfies*

$$u(x) = \frac{1}{\mathrm{m}(\partial B(x,r))}\int_{\partial B(x,r)} u(y)d\sigma(y)$$

for each ball $B(x,r) \subset \Omega$, then u is harmonic.

Proof. If $\Delta u \not\equiv 0$, then there exists some ball $B(x,r) \subset \Omega$ such that, say, $\Delta u > 0$ within $B(x,r)$. But then for f as above,

$$0 = f'(r) = \frac{1}{r^{n-1}\omega_{n-1}}\int_{B(x,r)} \Delta u(y)dy > 0,$$

is a contradiction. □

4.2 Poisson kernel and Hilbert transform

We shall now introduce a notation that will be indispensable in much of our further work. Indeed, we have shown some properties of Poisson kernel in Chapter 1. The setting for the application of this theory will be as follows. We shall think of $\mathbb{R}^n$ as the boundary hyperplane of the $(n+1)$ dimensional upper-half space $\mathbb{R}^{n+1}$. In coordinate notation,

$$\mathbb{R}^{n+1}_+ = \{(x,y) : x \in \mathbb{R}^n, y > 0\}.$$

We shall consider the *Poisson integral* of a function f given on $\mathbb{R}^n$. This Poisson integral is effectively the solution to the Dirichlet Problem for $\mathbb{R}^{n+1}_+$: find a *harmonic function* $u(x,y)$ on $\mathbb{R}^{n+1}_+$, whose boundary values on $\mathbb{R}^n$ (in the appropriate sense) are $f(x)$, that is

$$\begin{cases} \Delta_{x,y}u(x,y) = 0, \quad (x,y) \in \mathbb{R}^{n+1}_+, \\ u(x,0) = f, \quad x \in \mathbb{R}^n. \end{cases} \tag{4.14}$$

The formal solution of this problem can be given neatly in the context of the L^2 theory. In fact, let $f \in L^2(\mathbb{R}^n)$, and consider

$$u(x,y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}\hat{f}(\xi)d\xi, \quad y > 0. \tag{4.15}$$

This integral converges absolutely (cf. Theorem 1.15), because $\hat{f} \in L^2(\mathbb{R}^n)$, and $e^{-|\omega\xi|y}$ is rapidly decreasing in $|\xi|$ for $y > 0$. For the same reason, the integral above may be differentiated w.r.t. x and y any number of times by carrying out the operation under the sign of integration. This gives

$$\Delta_{x,y}u = \frac{\partial^2 u}{\partial y^2} + \sum_{k=1}^n \frac{\partial^2 u}{\partial x_k^2} = 0,$$

because the factor $e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}$ satisfies this property for each fixed ξ. Thus, $u(x,y)$ is a harmonic function on $\mathbb{R}^{n+1}_+$.

By Theorem 1.15, we get that $u(x,y) \to f(x)$ in $L^2(\mathbb{R}^n)$ norm, as $y \to 0$. That is, $u(x,y)$ satisfies the boundary condition and so $u(x,y)$ structured above is a solution for the above Dirichlet problem.

This solution of the problem can also be written without explicit use of the Fourier transform. For this purpose, we define the *Poisson kernel* $P_y(x) := P(x,y)$ by

$$P_y(x) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i\xi\cdot x} e^{-|\omega\xi|y} d\xi = (\mathscr{F}^{-1} e^{-|\omega\xi|y})(x), \quad y>0. \tag{4.16}$$

Then the function $u(x,y)$ obtained above can be written as a convolution

$$u(x,y) = \int_{\mathbb{R}^n} P_y(z) f(x-z) dz, \tag{4.17}$$

as the same as in Theorem 1.15. We shall say that u is the *Poisson integral* of f.

For convenience, we recall (1.12) and (1.10) as follows.

Proposition 4.7. *The Poisson kernel has the following explicit expression:*

$$P_y(x) = \frac{c_n y}{(|x|^2+y^2)^{\frac{n+1}{2}}}, \quad c_n = \frac{\Gamma((n+1)/2)}{\pi^{\frac{n+1}{2}}}. \tag{4.18}$$

Remark 4.8. We list the properties of the Poisson kernel that are now more or less evident:

(i) $P_y(x) > 0$ for $y > 0$.

(ii) $\int_{\mathbb{R}^n} P_y(x)dx = \widehat{P_y}(0) = 1$, $y > 0$; more generally, $\widehat{P_y}(\xi) = e^{-|\omega\xi|y}$ by Lemma 1.14 and Corollary 1.23, respectively.

(iii) $P_y(x)$ is homogeneous of degree $-n$: $P_y(x) = y^{-n}P_1(x/y)$, $y>0$.

(iv) $P_y(x)$ is a decreasing function of $|x|$, and $P_y(x) \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$. Indeed, by changes of variables, we have for $1 \leqslant p < \infty$

$$\begin{aligned}
\|P_y(x)\|_p^p =& c_n^p \int_{\mathbb{R}^n} \left(\frac{y}{(|x|^2+y^2)^{(n+1)/2}}\right)^p dx \\
\overset{x=yz}{=\!=} & c_n^p y^{-n(p-1)} \int_{\mathbb{R}^n} \frac{1}{(1+|z|^2)^{p(n+1)/2}} dz \\
\overset{z=rz'}{=\!=} & c_n^p y^{-n(p-1)} \omega_{n-1} \int_0^\infty \frac{1}{(1+r^2)^{p(n+1)/2}} r^{n-1} dr \\
\leqslant & c_n^p y^{-n(p-1)} \omega_{n-1} \left(\int_0^1 dr + \int_1^\infty r^{n-1-p(n+1)} dr\right) \\
\leqslant & c_n^p y^{-n(p-1)} \omega_{n-1} \left(1 + \frac{1}{p(n+1)-n}\right).
\end{aligned}$$

For $p=\infty$, it is clear that $\|P_y(x)\|_\infty \leqslant c_n y^{-n}$.

(v) Suppose $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, then its Poisson integral u, given by (4.17), is harmonic in $\mathbb{R}^{n+1}_+$. This is a simple consequence of the fact that $P_y(x)$ is harmonic in $\mathbb{R}^{n+1}_+$; the latter is immediately derived from (4.16).

(vi) We have the "semi-group property" $P_{y_1} * P_{y_2} = P_{y_1+y_2}$ if $y_1, y_2 > 0$ in view of Corollary 1.24.

The boundary behavior of Poisson integrals is already described to a significant extent by the following theorem.

Theorem 4.9. *Suppose $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, and let $u(x,y)$ be its Poisson integral. Then*

(a) $\sup_{y>0} |u(x,y)| \leqslant Mf(x)$, *where Mf is the maximal function.*

(b) $\lim_{y\to 0} u(x,y) = f(x)$, *for almost every x.*

(c) *If $p<\infty$, $u(x,y)$ converges to $f(x)$ in $L^p(\mathbb{R}^n)$ norm, as $y \to 0$.*

The theorem will now be proved in a more general setting, valid for a large class of approximations to the identity.

Let φ be an integrable function on $\mathbb{R}^n$, and set $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$, $\varepsilon > 0$.

Theorem 4.10. *Suppose that the least decreasing radial majorant of φ is integrable; i.e., let $\psi(x) = \sup_{|y|\geqslant|x|} |\varphi(y)|$, and we suppose $\int_{\mathbb{R}^n} \psi(x)dx = A < \infty$. Then with the same A,*

(a) $\sup_{\varepsilon>0} |(f * \varphi_\varepsilon)(x)| \leqslant AMf(x)$, $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$.
(b) *If in addition* $\int_{\mathbb{R}^n} \varphi(x)dx = 1$, *then* $\lim_{\varepsilon\to 0}(f * \varphi_\varepsilon)(x) = f(x)$ *almost everywhere.*
(c) *If* $p < \infty$, *then* $\|f * \varphi_\varepsilon - f\|_p \to 0$, *as* $\varepsilon \to 0$.

Proof. For the part (c), we have shown in Theorem 1.15.

Next, we prove assertion (a). We have already considered a special case of (a) in Chapter 3, with $\varphi(x) = \frac{1}{m(B)}\chi_B$. The point of the theorem is to reduce matters to this fundamental special case.

With a slight abuse of notation, let us write $\psi(r) = \psi(x)$, if $|x| = r$; it should cause no confusion since $\psi(x)$ is anyway radial. Now observe that $\psi(r)$ is decreasing and then $\int_{r/2\leqslant|x|\leqslant r} \psi(x)dx \geqslant \psi(r)\int_{r/2\leqslant|x|\leqslant r} dx = c\psi(r)r^n$. Therefore the assumption $\psi \in L^1$ proves that $r^n\psi(r) \to 0$ as $r \to 0$ or $r \to \infty$. To prove (a), we need to show that

$$(f * \psi_\varepsilon)(x) \leqslant AMf(x), \tag{4.19}$$

where $f \geqslant 0$, $f \in L^p(\mathbb{R}^n)$, $\varepsilon > 0$ and $A = \int_{\mathbb{R}^n} \psi(x)dx$.

Since (4.19) is clearly translation invariant w.r.t f and also dilation invariant w.r.t. ψ, it suffices to show that

$$(f * \psi)(0) \leqslant AMf(0). \tag{4.20}$$

In proving (4.20), we may clearly assume that $Mf(0) < \infty$. Let us write $\lambda(r) = \int_{S^{n-1}} f(rx')d\sigma(x')$, and $\Lambda(r) = \int_{|x|\leqslant r} f(x)dx$, so

$$\Lambda(r) = \int_0^r \int_{S^{n-1}} f(tx')d\sigma(x')t^{n-1}dt = \int_0^r \lambda(t)t^{n-1}dt, \quad \text{i.e., } \Lambda'(r) = \lambda(r)r^{n-1}.$$

We have

$$\begin{aligned}(f * \psi)(0) =& \int_{\mathbb{R}^n} f(x)\psi(x)dx = \int_0^\infty r^{n-1}\int_{S^{n-1}} f(rx')\psi(r)d\sigma(x')dr\\ =& \int_0^\infty r^{n-1}\lambda(r)\psi(r)dr = \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N \lambda(r)\psi(r)r^{n-1}dr\\ =& \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N \Lambda'(r)\psi(r)dr = \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \left\{[\Lambda(r)\psi(r)]_\varepsilon^N - \int_\varepsilon^N \Lambda(r)d\psi(r)\right\}.\end{aligned}$$

Since $\Lambda(r) = \int_{|x|\leqslant r} f(x)dx \leqslant V_n r^n Mf(0)$, and the fact $r^n\psi(r) \to 0$ as $r \to 0$ or $r \to \infty$, we have

$$0 \leqslant \lim_{N\to\infty} \Lambda(N)\psi(N) \leqslant V_n Mf(0) \lim_{N\to\infty} N^n\psi(N) = 0,$$

which implies $\lim_{N\to\infty} \Lambda(N)\psi(N) = 0$ and similarly $\lim_{\varepsilon\to 0} \Lambda(\varepsilon)\psi(\varepsilon) = 0$. Thus, by integration by parts, we have

$$\begin{aligned}(f * \psi)(0) =& \int_0^\infty \Lambda(r)d(-\psi(r)) \leqslant V_n Mf(0)\int_0^\infty r^n d(-\psi(r))\\ =& nV_n Mf(0)\int_0^\infty \psi(r)r^{n-1}dr = Mf(0)\int_{\mathbb{R}^n} \psi(x)dx,\end{aligned}$$

since $\psi(r)$ is decreasing which implies $\psi'(r) \leqslant 0$, and $nV_n = \omega_{n-1}$. This proves (4.20) and then (4.19).

Finally, we prove (b) in a familiar way as follows. First, we can verify that if $f_1 \in \mathcal{C}_c^0$, then $(f_1 * \varphi_\varepsilon)(x) \to f_1(x)$ uniformly as $\varepsilon \to 0$ (cf. Theorem 1.15). Next we can deal with the case $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, by writing $f = f_1 + f_2$ with f_1 as described and with $\|f_2\|_p$ small. The argument then follows closely that given in the proof of Corollary 3.13 (the Lebesgue differentiation theorem). Thus we get that $\lim_{\varepsilon\to 0} f_\varepsilon(x)$ exists almost everywhere and equals $f(x)$.

To deal with the remaining case, that of bounded f, we fix any ball B, and set ourselves the task of showing that

$$\lim_{\varepsilon\to 0}(f*\varphi_\varepsilon)(x)=f(x),\ \text{for almost every } x\in B.$$

Let B_1 be any other ball which strictly contains B, and let $\delta=\operatorname{dist}(B,B_1^c)$ be the distance from B to the complement of B_1. Let $f_1(x)=\begin{cases} f(x), & x\in B_1,\\ 0, & x\notin B_1,\end{cases}$; $f(x)=f_1(x)+f_2(x)$. Then, $f_1\in L^1(\mathbb{R}^n)$, and so the appropriate conclusion holds for it. However, for $x\in B$,

$$\begin{aligned}|(f_2*\varphi_\varepsilon)(x)|=&\left|\int_{\mathbb{R}^n} f_2(x-y)\varphi_\varepsilon(y)dy\right|\leqslant\int_{|y|\geqslant\delta>0}|f_2(x-y)||\varphi_\varepsilon(y)|dy\\ \leqslant&\|f\|_\infty\int_{|y|\geqslant\delta/\varepsilon}|\varphi(y)|dy\to 0,\ \text{as } \varepsilon\to 0.\end{aligned}$$

Thus, we complete the proof. □

Proof of Theorem 4.9. Theorem 4.10 then applies directly to prove Theorem 4.9, because of properties (i)–(iv) of the Poisson kernel in the case $\varphi(x)=\psi(x)=P_1(x)$. □

There are also some variants of the result of Theorem 4.10, which apply equally well to Poisson integrals. The first is an easy adaptation of the argument already given, and is stated without proof.

Corollary 4.11. *Suppose f is continuous and bounded on $\mathbb{R}^n$. Then $(f*\varphi_\varepsilon)(x)\to f(x)$ uniformly on compact subsets of $\mathbb{R}^n$.*

The second variant is somewhat more difficult. It is the analogue for finite Borel measures in place of integrable functions, and is outlined in further result of [Ste70, §4.1, p.77–78].

Now, we give the definition of harmonic conjugate functions as follows.

Definition 4.12. The *harmonic conjugate* to a given function $u(x,y)$ is a function $v(x,y)$ such that

$$f(x,y)=u(x,y)+iv(x,y)$$

is analytic, i.e., satisfies the *Cauchy-Riemann equations*

$$u_x=v_y,\quad u_y=-v_x,$$

where $u_x\equiv\partial u/\partial x$, $u_y\equiv\partial u/\partial y$. It is given by

$$v(x,y)=\int_{(x_0,y_0)}^{(x,y)}u_xdy-u_ydx+C,$$

along any path connecting (x_0,y_0) and (x,y) in the domain, where C is a constant of integration.

Given a function f in $\mathscr{S}(\mathbb{R})$, its harmonic extension to the upper half-plane is given by $u(x,y)=P_y*f(x)$, where P_y is the Poisson kernel. We can also write, in view of (4.15),

$$\begin{aligned}u(z)=&u(x,y)=\frac{|\omega|}{2\pi}\int_{\mathbb{R}}e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}\hat{f}(\xi)d\xi\\ =&\frac{|\omega|}{2\pi}\left[\int_0^\infty e^{\omega i\xi\cdot x}e^{-|\omega|\xi y}\hat{f}(\xi)d\xi+\int_{-\infty}^0 e^{\omega i\xi\cdot x}e^{|\omega|\xi y}\hat{f}(\xi)d\xi\right]\\ =&\frac{|\omega|}{2\pi}\left[\int_0^\infty e^{\omega i\xi\cdot(x+i\operatorname{sgn}(\omega)y)}\hat{f}(\xi)d\xi+\int_{-\infty}^0 e^{\omega i\xi\cdot(x-i\operatorname{sgn}(\omega)y)}\hat{f}(\xi)d\xi\right],\end{aligned}$$

where $z = x + iy$. If we now define

$$i\,\mathrm{sgn}\,(\omega)v(z) = \frac{|\omega|}{2\pi}\left[\int_0^\infty e^{\omega i\xi\cdot(x+i\,\mathrm{sgn}\,(\omega)y)}\hat{f}(\xi)d\xi - \int_{-\infty}^0 e^{\omega i\xi\cdot(x-i\,\mathrm{sgn}\,(\omega)y)}\hat{f}(\xi)d\xi\right],$$

then v is also harmonic in $\mathbb{R}^2_+$ and both u and v are real if f is. Furthermore, $u+iv$ is analytic since it satisfies the Cauchy-Riemann equations $u_x = v_y = \omega i\xi u(z)$ and $u_y = -v_x = -\omega i\xi v(z)$, so v is the harmonic conjugate of u.

Clearly, v can also be written as, by Theorem 1.12, Proposition 1.3 and Theorem 1.28,

$$\begin{aligned}
v(z) =&\frac{|\omega|}{2\pi}\int_{\mathbb{R}} -i\,\mathrm{sgn}\,(\omega)\,\mathrm{sgn}\,(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}\hat{f}(\xi)d\xi\\
=&\frac{|\omega|}{2\pi}\int_{\mathbb{R}} -i\,\mathrm{sgn}\,(\omega)\mathscr{F}_\xi[\mathrm{sgn}\,(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}](\eta)f(\eta)d\eta\\
=&\frac{|\omega|}{2\pi}\int_{\mathbb{R}} -i\,\mathrm{sgn}\,(\omega)\mathscr{F}_\xi[\mathrm{sgn}\,(\xi)e^{-|\omega\xi|y}](\eta-x)f(\eta)d\eta\\
=&\int_{\mathbb{R}} -i\,\mathrm{sgn}\,(\omega)\mathscr{F}_\xi^{-1}[\mathrm{sgn}\,(\xi)e^{-|\omega\xi|y}](x-\eta)f(\eta)d\eta,
\end{aligned}$$

which is equivalent to

$$v(x,y) = Q_y * f(x), \tag{4.21}$$

where

$$\hat{Q}_y(\xi) = -i\,\mathrm{sgn}\,(\omega)\,\mathrm{sgn}\,(\xi)e^{-|\omega\xi|y}. \tag{4.22}$$

Now we invert the Fourier transform, we get, by a change of variables and integration by parts,

$$\begin{aligned}
Q_y(x) =& -i\,\mathrm{sgn}\,(\omega)\frac{|\omega|}{2\pi}\int_{\mathbb{R}} e^{\omega ix\cdot\xi}\,\mathrm{sgn}\,(\xi)e^{-|\omega\xi|y}d\xi\\
=& -i\,\mathrm{sgn}\,(\omega)\frac{|\omega|}{2\pi}\left[\int_0^\infty e^{\omega ix\cdot\xi}e^{-|\omega|\xi y}d\xi - \int_{-\infty}^0 e^{\omega ix\cdot\xi}e^{|\omega|\xi y}d\xi\right]\\
=& -i\,\mathrm{sgn}\,(\omega)\frac{|\omega|}{2\pi}\left[\int_0^\infty e^{\omega ix\cdot\xi}e^{-|\omega|\xi y}d\xi - \int_0^\infty e^{-\omega ix\cdot\xi}e^{-|\omega|\xi y}d\xi\right]\\
=& -i\,\mathrm{sgn}\,(\omega)\frac{|\omega|}{2\pi}\int_0^\infty \left(e^{\omega ix\cdot\xi} - e^{-\omega ix\cdot\xi}\right)\frac{\partial_\xi e^{-|\omega|\xi y}}{-|\omega|y}d\xi\\
=&i\,\mathrm{sgn}\,(\omega)\frac{1}{2\pi y}\left[\left(e^{\omega ix\cdot\xi} - e^{-\omega ix\cdot\xi}\right)e^{-|\omega|\xi y}\Big|_0^\infty - \int_0^\infty \omega ix\left(e^{\omega ix\cdot\xi} + e^{-\omega ix\cdot\xi}\right)e^{-|\omega|\xi y}d\xi\right]\\
=&\frac{|\omega|x}{2\pi y}\int_0^\infty \left(e^{\omega ix\cdot\xi} + e^{-\omega ix\cdot\xi}\right)e^{-|\omega|\xi y}d\xi\\
=&\frac{|\omega|x}{2\pi y}\int_{\mathbb{R}} e^{-\omega ix\cdot\xi}e^{-|\omega\xi|y}d\xi = \frac{x}{y}\mathscr{F}\left(\frac{|\omega|}{2\pi}e^{-|\omega\xi|y}\right)\\
=&\frac{x}{y}P_y(x) = \frac{x}{y}\frac{c_1 y}{y^2+x^2} = \frac{c_1 x}{y^2+x^2},
\end{aligned}$$

where $c_1 = \Gamma(1)/\pi = 1/\pi$. That is,

$$Q_y(x) = \frac{1}{\pi}\frac{x}{y^2+x^2}.$$

One can immediately verify that $Q(x,y) = Q_y(x)$ is a harmonic function in the upper half-plane and the conjugate of the Poisson kernel $P_y(x) = P(x,y)$. More precisely, they satisfy Cauchy-Riemann equations

$$\partial_x P = \partial_y Q = -\frac{1}{\pi}\frac{2xy}{(y^2+x^2)^2}, \quad \partial_y P = -\partial_x Q = \frac{1}{\pi}\frac{x^2-y^2}{(y^2+x^2)^2}.$$

In Theorem 4.9, we studied the limit of $u(x,t)$ as $y \to 0$ using the fact that $\{P_y\}$ is an approximation of the identity. We would like to do the same for $v(x,y)$, but we immediately run into an obstacle: $\{Q_y\}$ is not an approximation of the identity and, in fact, Q_y is not integrable for any $y > 0$. Formally,

$$\lim_{y\to 0} Q_y(x) = \frac{1}{\pi x},$$

this is not even locally integrable, so we cannot define its convolution with smooth functions.

We define a tempered distribution called the *principal value of* $1/x$, abbreviated p.v. $1/x$, by

$$\left\langle \text{p.v.}\,\frac{1}{x}, \phi \right\rangle = \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \frac{\phi(x)}{x} dx, \quad \phi \in \mathscr{S}.$$

To see that this expression defines a tempered distribution, we rewrite it as

$$\left\langle \text{p.v.}\,\frac{1}{x}, \phi \right\rangle = \int_{|x|<1} \frac{\phi(x)-\phi(0)}{x} dx + \int_{|x|\geqslant 1} \frac{\phi(x)}{x} dx,$$

this holds since the integral of $1/x$ on $\varepsilon < |x| < 1$ is zero. It is now immediate that

$$\left| \left\langle \text{p.v.}\,\frac{1}{x}, \phi \right\rangle \right| \leqslant C(\|\phi'\|_\infty + \|x\phi\|_\infty).$$

Proposition 4.13. *In* $\mathscr{S}'(\mathbb{R})$*, we have* $\lim_{y\to 0} Q_y(x) = \frac{1}{\pi}\,\text{p.v.}\,\frac{1}{x}$.

Proof. For each $\varepsilon > 0$, the functions $\psi_\varepsilon(x) = x^{-1}\chi_{|x|>\varepsilon}$ are bounded and define tempered distributions. It follows at once from the definition that in $\mathscr{S}'$,

$$\lim_{\varepsilon\to 0} \psi_\varepsilon(x) = \text{p.v.}\,\frac{1}{x}.$$

Therefore, it will suffice to prove that in $\mathscr{S}'$

$$\lim_{y\to 0}\left(Q_y - \frac{1}{\pi}\psi_y\right) = 0.$$

Fix $\phi \in \mathscr{S}$, then by a change of variables, we have

$$\begin{aligned}
\langle \pi Q_y - \psi_y, \phi\rangle &= \int_{\mathbb{R}} \frac{x\phi(x)}{y^2+x^2} dx - \int_{|x|\geqslant y} \frac{\phi(x)}{x} dx \\
&= \int_{|x|<y} \frac{x\phi(x)}{y^2+x^2} dx + \int_{|x|\geqslant y} \left(\frac{x}{y^2+x^2} - \frac{1}{x}\right)\phi(x) dx \\
&= \int_{|x|<1} \frac{x\phi(yx)}{1+x^2} dx - \int_{|x|\geqslant 1} \frac{\phi(yx)}{x(1+x^2)} dx.
\end{aligned}$$

If we take the limit as $y \to 0$ and apply the dominated convergence theorem, we get two integrals of odd functions on symmetric domains. Hence, the limit equals 0. $\square$

As a consequence of this proposition, we get that

$$\lim_{y\to 0} Q_y * f(x) = \frac{1}{\pi} \lim_{\varepsilon\to 0} \int_{|t|>\varepsilon} \frac{f(x-t)}{t} dt,$$

and by the continuity of the Fourier transform on $\mathscr{S}'$ and by (4.22), we get

$$\mathscr{F}\left(\frac{1}{\pi}\,\text{p.v.}\,\frac{1}{x}\right)(\xi) = -i\,\text{sgn}\,(\omega)\,\text{sgn}\,(\xi).$$

Given a function $f \in \mathscr{S}$, we can define its *Hilbert transform* by any one of the following equivalent expressions:

$$Hf = \lim_{y\to 0} Q_y * f,$$

$$Hf = \frac{1}{\pi} \text{ p.v. } \frac{1}{x} * f,$$
$$Hf = \mathscr{F}^{-1}(-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \hat{f}(\xi)).$$

The third expression also allows us to define the Hilbert transform of functions in $L^2(\mathbb{R})$, which satisfies, with the help of Theorem 1.26,

$$\|Hf\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{1/2} \|\mathscr{F}(Hf)\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{1/2} \|\hat{f}\|_2 = \|f\|_2, \tag{4.23}$$

that is, H is an isometry on $L^2(\mathbb{R})$. Moreover, H satisfies

$$H^2 f = H(Hf) = \mathscr{F}^{-1}((-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi))^2 \hat{f}(\xi)) = -f, \tag{4.24}$$

By Theorem 1.28, we have

$$\begin{aligned}
\langle Hf, g \rangle = \int_{\mathbb{R}} Hf \cdot g dx &= \int_{\mathbb{R}} \mathscr{F}^{-1}(-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \hat{f}(\xi)) \cdot g dx \\
&= \int_{\mathbb{R}} -i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \hat{f}(\xi) \cdot \check{g}(\xi) d\xi \\
&= \int_{\mathbb{R}} f(x) \cdot \mathscr{F}[-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \check{g}(\xi)](x) dx \\
&= \int_{\mathbb{R}} f(x) \cdot \mathscr{F}[-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \frac{|\omega|}{2\pi} \hat{g}(-\xi)](x) dx \\
&= \int_{\mathbb{R}} f(x) \cdot \mathscr{F}^{-1}[i \operatorname{sgn}(\omega) \operatorname{sgn}(\eta) \hat{g}(\eta)](x) dx \\
&= -\int_{\mathbb{R}} f \cdot Hg dx = \langle f, -Hg \rangle,
\end{aligned} \tag{4.25}$$

namely, the dual/conjugate operator of H is $H' = -H$. Similarly, the adjoint operator H^* of H is uniquely defined via the identity

$$(f, Hg) = \int_{\mathbb{R}} f \cdot \overline{Hg} dx = -\int_{\mathbb{R}} Hf \bar{g} dx = (-Hf, g) =: (H^* f, g),$$

that is, $H^* = -H$.

Note that for given $x \in \mathbb{R}$, $Hf(x)$ is defined for all integrable functions f on $\mathbb{R}$ that satisfy a Hölder condition near the point x, that is,

$$|f(x) - f(y)| \leqslant C_x |x - y|^{\varepsilon_x}$$

for some $C_x > 0$ and $\varepsilon_x > 0$ whenever $|y - x| < \delta_x$. Indeed, suppose that this is the case, then

$$\begin{aligned}
Q_y * f(x) = &\frac{1}{\pi} \int_{\varepsilon < |x-y| < \delta_x} \frac{f(y)}{x - y} dy + \frac{1}{\pi} \int_{|x-y| \geqslant \delta_x} \frac{f(y)}{x - y} dy \\
= &\frac{1}{\pi} \int_{\varepsilon < |x-y| < \delta_x} \frac{f(y) - f(x)}{x - y} dy + \frac{1}{\pi} \int_{|x-y| \geqslant \delta_x} \frac{f(y)}{x - y} dy.
\end{aligned}$$

Both integrals converge absolutely, and hence the limit of $Q_y * f(x)$ exists as $\varepsilon \to 0$. Therefore, the Hilbert transform of a piecewise smooth integrable function is well defined at all points of Hölder-Lipschitz continuity of the function. On the other hand, observe that $Q_y * f$ is well defined for all $f \in L^p$, $1 \leqslant p < \infty$, as it follows from the Hölder inequality, since $1/x$ is in $L^{p'}$ on the set $|x| \geqslant \varepsilon$.

Example 4.14. Consider the characteristic function $\chi_{[a,b]}$ of an interval $[a, b]$. It is a simple calculation to show that

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \ln \frac{|x - a|}{|x - b|}. \tag{4.26}$$

Let us verify this identity. By the definition, we have

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{|y|>\varepsilon} \frac{\chi_{[a,b]}(x-y)}{y} dy = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\substack{|y|>\varepsilon \\ x-b \leqslant y \leqslant x-a}} \frac{1}{y} dy.$$

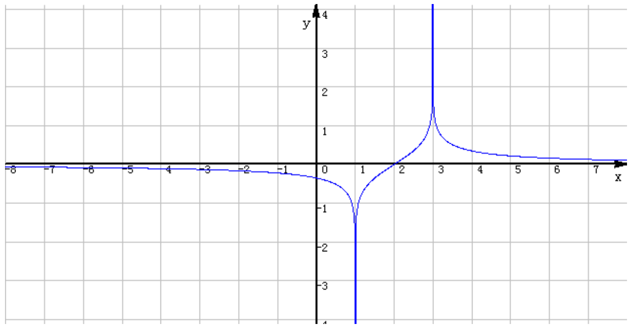
Thus, we only need to consider three cases: $x-b>0$, $x-a<0$ and $x-b<0<x-a$. For the first two cases, we have

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \int_{x-b}^{x-a} \frac{1}{y} dy = \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}.$$

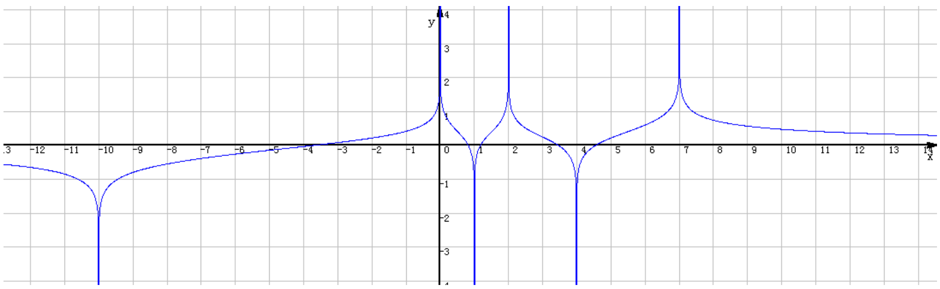
For the third case we get (without loss of generality, we can assume $\varepsilon < \min(|x-a|, |x-b|)$)

$$\begin{aligned} H(\chi_{[a,b]})(x) =& \frac{1}{\pi} \lim_{\varepsilon \to 0} \left(\int_{x-b}^{-\varepsilon} \frac{1}{y} dy + \int_{\varepsilon}^{x-a} \frac{1}{y} dy \right) = \frac{1}{\pi} \lim_{\varepsilon \to 0} \left(\ln \frac{|x-a|}{\varepsilon} + \ln \frac{\varepsilon}{|x-b|} \right) \\ =& \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}, \end{aligned}$$

where it is crucial to observe how the cancellation of the odd kernel $1/x$ is manifested. Note that $H(\chi_{[a,b]})(x)$ blows up logarithmically for x near the points a and b and decays like x^{-1} as $x \to \pm\infty$. See the following graph with $a=1$ and $b=3$:



The following is a graph of the function $H(\chi_{[-10,0]\cup[1,2]\cup[4,7]})$:



It is obvious, for the dilation operator δ_ε with $\varepsilon > 0$, by changes of variables ($\varepsilon y \to y$), that

$$(H\delta_\varepsilon) f(x) = \lim_{\sigma \to 0} \frac{1}{\pi} \int_{|y| \geqslant \sigma} \frac{f(\varepsilon x - \varepsilon y)}{y} dy = \lim_{\sigma \to 0} \int_{|y| \geqslant \varepsilon\sigma} \frac{f(\varepsilon x - y)}{y} dy = (\delta_\varepsilon H) f(x),$$

so $H\delta_\varepsilon = \delta_\varepsilon H$; and it is equally obvious that $H\delta_\varepsilon = -\delta_\varepsilon H$, if $\varepsilon < 0$.

These simple considerations of dilation "invariance" and the obvious translation invariance in fact characterize the Hilbert transform.

Proposition 4.15 (Characterization of Hilbert transform)**.** *Suppose T is a bounded linear operator on $L^2(\mathbb{R})$ which satisfies the following properties:*

(a) *T commutes with translations;*

(b) *T commutes with positive dilations;*

(c) *T anticommutes with the reflection $f(x) \to f(-x)$.*

Then, T is a constant multiple of the Hilbert transform.

Proof. Since T commutes with translations and maps $L^2(\mathbb{R})$ to itself, according to Theorem 1.62, there is a bounded function $m(\xi)$ such that $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$. The assumptions (b) and (c) may be written as $T\delta_\varepsilon f = \operatorname{sgn}(\varepsilon)\delta_\varepsilon Tf$ for all $f \in L^2(\mathbb{R})$. By part (iv) in Proposition 1.3, we have

$$\begin{aligned}\mathscr{F}(T\delta_\varepsilon f)(\xi) =& m(\xi)\mathscr{F}(\delta_\varepsilon f)(\xi) = m(\xi)|\varepsilon|^{-1}\hat{f}(\xi/\varepsilon),\\ \operatorname{sgn}(\varepsilon)\mathscr{F}(\delta_\varepsilon Tf)(\xi) =& \operatorname{sgn}(\varepsilon)|\varepsilon|^{-1}\widehat{Tf}(\xi/\varepsilon) = \operatorname{sgn}(\varepsilon)|\varepsilon|^{-1}m(\xi/\varepsilon)\hat{f}(\xi/\varepsilon),\end{aligned}$$

which means $m(\varepsilon\xi) = \operatorname{sgn}(\varepsilon)m(\xi)$, if $\varepsilon \neq 0$. This shows that $m(\xi) = c\operatorname{sgn}(\xi)$, and the proposition is proved. □

The next theorem shows that the Hilbert transform, now defined for functions in $\mathscr{S}$ or L^2, can be extended to functions in L^p, $1 \leqslant p < \infty$.

Theorem 4.16. *For $f \in \mathscr{S}(\mathbb{R})$, the following assertions are true:*

(i) (Kolmogorov) *H is of weak type* $(1,1)$*:*

$$\mathfrak{m}(\{x \in \mathbb{R} : |Hf(x)| > \alpha\}) \leqslant \frac{C}{\alpha}\|f\|_1.$$

(ii) (M. Riesz) *H is of type (p,p), $1 < p < \infty$:*

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

Proof. (i) Fix $\alpha > 0$. From the Calderón-Zygmund decomposition of f at height α (Theorem 3.20), there exist two functions g and b such that $f = g + b$ and

(1) $\|g\|_1 \leqslant \|f\|_1$ and $\|g\|_\infty \leqslant 2\alpha$.

(2) $b = \sum_j b_j$, where each b_j is supported in a dyadic interval I_j satisfying $\int_{I_j} b_j(x)dx = 0$ and $\|b_j\|_1 \leqslant 4\alpha\mathfrak{m}(I_j)$. Furthermore, the intervals I_j and I_k have disjoint interiors when $j \neq k$.

(3) $\sum_j \mathfrak{m}(I_j) \leqslant \alpha^{-1}\|f\|_1$.

Let $2I_j$ be the interval with the same center as I_j and twice the length, and let $\Omega = \cup_j I_j$ and $\Omega^* = \cup_j 2I_j$. Then $\mathfrak{m}(\Omega^*) \leqslant 2\mathfrak{m}(\Omega) \leqslant 2\alpha^{-1}\|f\|_1$.

Since $Hf = Hg + Hb$, from parts (iv) and (vi) of Proposition 2.15, (4.23) and (1), we have

$$\begin{aligned}(Hf)_*(\alpha) \leqslant& (Hg)_*(\alpha/2) + (Hb)_*(\alpha/2)\\ \leqslant& (\alpha/2)^{-2}\int_{\mathbb{R}} |Hg(x)|^2 dx + \mathfrak{m}(\Omega^*) + \mathfrak{m}(\{x \notin \Omega^* : |Hb(x)| > \alpha/2\})\\ \leqslant& \frac{4}{\alpha^2}\int_{\mathbb{R}} |g(x)|^2 dx + 2\alpha^{-1}\|f\|_1 + 2\alpha^{-1}\int_{\mathbb{R}\setminus\Omega^*} |Hb(x)|dx\\ \leqslant& \frac{8}{\alpha}\int_{\mathbb{R}} |g(x)|dx + \frac{2}{\alpha}\|f\|_1 + \frac{2}{\alpha}\int_{\mathbb{R}\setminus\Omega^*}\sum_j |Hb_j(x)|dx\\ \leqslant& \frac{8}{\alpha}\|f\|_1 + \frac{2}{\alpha}\|f\|_1 + \frac{2}{\alpha}\sum_j\int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx.\end{aligned}$$

For $x \notin 2I_j$, we have

$$Hb_j(x) = \frac{1}{\pi}\,\text{p.v.}\int_{I_j}\frac{b_j(y)}{x-y}dy = \frac{1}{\pi}\int_{I_j}\frac{b_j(y)}{x-y}dy,$$

since $\operatorname{supp} b_j \subset I_j$ and $|x-y| \geqslant \mathfrak{m}(I_j)/2$ for $y \in I_j$. Denote the center of I_j by c_j, then, since b_j is mean zero, we have

$$\int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx = \int_{\mathbb{R}\setminus 2I_j}\left|\frac{1}{\pi}\int_{I_j}\frac{b_j(y)}{x-y}dy\right|dx$$

$$=\frac{1}{\pi}\int_{\mathbb{R}\setminus 2I_j}\left|\int_{I_j} b_j(y)\left(\frac{1}{x-y}-\frac{1}{x-c_j}\right)dy\right|dx$$
$$\leqslant\frac{1}{\pi}\int_{I_j}|b_j(y)|\left(\int_{\mathbb{R}\setminus 2I_j}\frac{|y-c_j|}{|x-y||x-c_j|}dx\right)dy$$
$$\leqslant\frac{1}{\pi}\int_{I_j}|b_j(y)|\left(\int_{\mathbb{R}\setminus 2I_j}\frac{\mathrm{m}(I_j)}{|x-c_j|^2}dx\right)dy.$$

The last inequality follows from the fact that $|y-c_j|<\mathrm{m}(I_j)/2$ and $|x-y|>|x-c_j|/2$. Since $|x-c_j|>\mathrm{m}(I_j)$, the inner integral equals

$$2\mathrm{m}(I_j)\int_{\mathrm{m}(I_j)}^{\infty}\frac{1}{r^2}dr=2\mathrm{m}(I_j)\frac{1}{\mathrm{m}(I_j)}=2.$$

Thus, by (2) and (3),

$$(Hf)_*(\alpha)\leqslant\frac{10}{\alpha}\|f\|_1+\frac{4}{\alpha\pi}\sum_j\int_{I_j}|b_j(y)|dy\leqslant\frac{10}{\alpha}\|f\|_1+\frac{4}{\alpha\pi}\sum_j 4\alpha\mathrm{m}(I_j)$$
$$\leqslant\frac{10}{\alpha}\|f\|_1+\frac{16}{\pi}\frac{1}{\alpha}\|f\|_1=\frac{10+16/\pi}{\alpha}\|f\|_1.$$

(ii) Since H is of weak type $(1,1)$ and of type $(2,2)$, by the Marcinkiewicz interpolation theorem, we have the strong (p,p) inequality for $1<p<2$. If $p>2$, we apply the dual estimate with the help of (4.25) and the result for $p<2$:

$$\|Hf\|_p=\sup_{\|g\|_{p'}\leqslant 1}|\langle Hf,g\rangle|=\sup_{\|g\|_{p'}\leqslant 1}|\langle f,Hg\rangle|\leqslant\|f\|_p\sup_{\|g\|_{p'}\leqslant 1}\|Hg\|_{p'}\leqslant C_{p'}\|f\|_p.$$

This completes the proof. $\square$

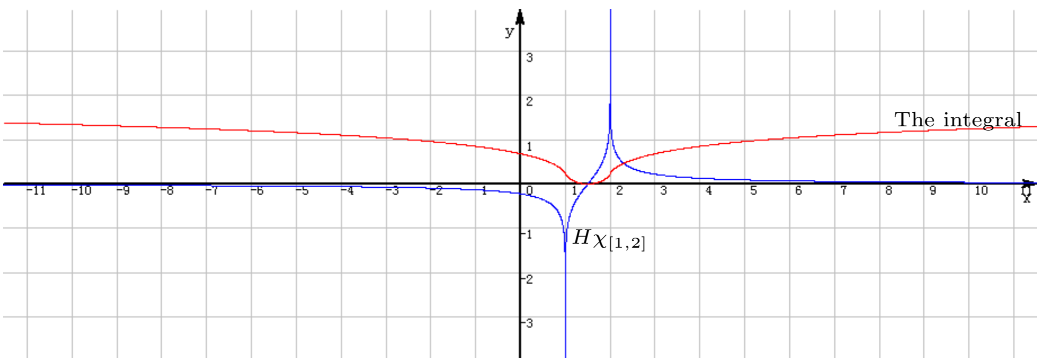
Remark 4.17. i) Recall from the proof of the Marcinkiewicz interpolation theorem that the coefficient

$$C_p=\begin{cases}\dfrac{10+16/\pi}{1-1/p}+\dfrac{(1/2)^{1/2}}{1/p-1/2}+2^{1/2}, & 1<p<2,\\[2ex] (10+16/\pi)p+\dfrac{(1/2)^{1/2}}{1/2-1/p}+2^{1/2}, & p>2.\end{cases}$$

So the constant C_p tends to infinity as p tends to 1 or ∞. More precisely,

$$C_p=O(p)\text{ as }p\to\infty,\text{ and }C_p=O((p-1)^{-1})\text{ as }p\to 1.$$

ii) The strong (p,p) inequality is false if $p=1$ or $p=\infty$, this can easily be seen from the previous example $H\chi_{[a,b]}=\frac{1}{\pi}\ln\frac{|x-a|}{|x-b|}$ which is neither integrable nor bounded. See the following figure.



iii) By using the inequalities in Theorem 4.16, we can extend the Hilbert transform to functions in L^p, $1 \leqslant p < \infty$. If $f \in L^1$ and $\{f_n\}$ is a sequence of functions in $\mathscr{S}$ that converges to f in L^1, then by the weak $(1,1)$ inequality the sequence $\{Hf_n\}$ is a Cauchy sequence in measure: for any $\varepsilon > 0$,

$$\lim_{m,n\to\infty} \mathfrak{m}(\{x \in \mathbb{R} : |(Hf_n - Hf_m)(x)| > \varepsilon\}) = 0.$$

Therefore, it converges in measure to a measurable function which we define to be the Hilbert transform of f.

If $f \in L^p$, $1 < p < \infty$, and $\{f_n\}$ is a sequence of functions in $\mathscr{S}$ that converges to f in L^p, by the strong (p,p) inequality, $\{Hf_n\}$ is a Cauchy sequence in L^p, so it converges to a function in L^p which we call the Hilbert transform of f.

In either case, a subsequence of $\{Hf_n\}$, depending on f, converges pointwise almost everywhere to Hf as defined.

4.3 The Calderón-Zygmund theorem

From this section on, we are going to consider singular integrals whose kernels have the same essential properties as the kernel of the Hilbert transform. We can generalize Theorem 4.16 to get the following result.

Theorem 4.18 (Calderón-Zygmund Theorem)**.** *Let K be a tempered distribution in $\mathbb{R}^n$ which coincides with a locally integrable function on $\mathbb{R}^n \setminus \{0\}$ and satisfies*

$$|\widehat{K}(\xi)| \leqslant B, \tag{4.27}$$

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \quad y \in \mathbb{R}^n. \tag{4.28}$$

Then we have the strong (p,p) estimate for $1 < p < \infty$

$$\|K * f\|_p \leqslant C_p \|f\|_p, \tag{4.29}$$

and the weak $(1,1)$ estimate

$$(K * f)_*(\alpha) \leqslant \frac{C}{\alpha}\|f\|_1. \tag{4.30}$$

We will show that these inequalities are true for $f \in \mathscr{S}$, but they can be extended to arbitrary $f \in L^p$ as we did for the Hilbert transform. Condition (4.28) is usually referred to as the *Hörmander condition*; in practice it is often deduced from another stronger condition called the *gradient condition* (i.e., (4.31) as below).

Proposition 4.19. *The Hörmander condition (4.28) holds if for every $x \neq 0$*

$$|\nabla K(x)| \leqslant \frac{C}{|x|^{n+1}}. \tag{4.31}$$

Proof. By the integral mean value theorem and (4.31), we have

$$\begin{aligned}
&\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant \int_{|x|\geqslant 2|y|} \int_0^1 |\nabla K(x-\theta y)||y|d\theta dx \\
\leqslant &\int_0^1 \int_{|x|\geqslant 2|y|} \frac{C|y|}{|x-\theta y|^{n+1}} dx d\theta \leqslant \int_0^1 \int_{|x|\geqslant 2|y|} \frac{C|y|}{(|x|/2)^{n+1}} dx d\theta \\
\leqslant &2^{n+1} C|y|\omega_{n-1} \int_{2|y|}^\infty \frac{1}{r^2} dr = 2^{n+1}C|y|\omega_{n-1}\frac{1}{2|y|} = 2^n C\omega_{n-1}.
\end{aligned}$$

This completes the proof. □

Proof of Theorem 4.18. Since the proof is (essentially) a repetition of the proof of Theorem 4.16, we will omit the details.

Let $f \in \mathscr{S}$ and $Tf = K * f$. From (4.27), it follows that

$$\begin{aligned}\|Tf\|_2 =& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\widehat{Tf}\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\widehat{K}\hat{f}\|_2 \\ \leqslant& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\widehat{K}\|_\infty \|\hat{f}\|_2 \leqslant B \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\hat{f}\|_2 \\ =& B\|f\|_2,\end{aligned} \tag{4.32}$$

by the Plancherel theorem (Theorem 1.26) and part (vi) in Proposition 1.3.

It will suffice to prove that T is of weak type $(1,1)$ since the strong (p,p) inequality, $1 < p < 2$, follows from the interpolation, and for $p > 2$ it follows from the duality since the conjugate operator T' has kernel $K'(x) = K(-x)$ which also satisfies (4.27) and (4.28). In fact,

$$\begin{aligned}\langle Tf, \varphi\rangle =& \int_{\mathbb{R}^n} Tf(x)\varphi(x)dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x-y)f(y)dy\varphi(x)dx \\ =& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(-(y-x))\varphi(x)dxf(y)dy = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (K' * \varphi)(y)f(y)dy \\ =& \langle f, T'\varphi\rangle.\end{aligned}$$

To show that f is of weak type $(1,1)$, fix $\alpha > 0$ and form the Calderón-Zygmund decomposition of f at height α. Then as in Theorem 4.16, we can write $f = g + b$, where

(i) $\|g\|_1 \leqslant \|f\|_1$ and $\|g\|_\infty \leqslant 2^n\alpha$.

(ii) $b = \sum_j b_j$, where each b_j is supported in a dyadic cube Q_j satisfying $\int_{Q_j} b_j(x)dx = 0$ and $\|b_j\|_1 \leqslant 2^{n+1}\alpha \mathfrak{m}(Q_j)$. Furthermore, the cubes Q_j and Q_k have disjoint interiors when $j \neq k$.

(iii) $\sum_j \mathfrak{m}(Q_j) \leqslant \alpha^{-1}\|f\|_1$.

The argument now proceeds as before, and the proof reduces to showing that

$$\int_{\mathbb{R}^n \setminus Q_j^*} |Tb_j(x)|dx \leqslant C \int_{Q_j} |b_j(x)|dx, \tag{4.33}$$

where Q_j^* is the cube with the same center as Q_j and whose sides are $2\sqrt{n}$ times longer. Denote their common center by c_j. Inequality (4.33) follows from the Hörmander condition (4.28): since each b_j has zero average, if $x \notin Q_j^*$

$$Tb_j(x) = \int_{Q_j} K(x-y)b_j(y)dy = \int_{Q_j} [K(x-y) - K(x-c_j)]b_j(y)dy;$$

hence,

$$\int_{\mathbb{R}^n \setminus Q_j^*} |Tb_j(x)|dx \leqslant \int_{Q_j}\left(\int_{\mathbb{R}^n \setminus Q_j^*} |K(x-y) - K(x-c_j)|dx\right)|b_j(y)|dy.$$

However, by changing variables $x - c_j = x'$ and $y - c_j = y'$, and the fact that $|x - c_j| \geqslant 2|y - c_j|$ for all $x \notin Q_j^*$ and $y \in Q_j$ as an obvious geometric consideration shows, and (4.28), we get

$$\int_{\mathbb{R}^n \setminus Q_j^*} |K(x-y) - K(x-c_j)|dx \leqslant \int_{|x'| \geqslant 2|y'|} |K(x'-y') - K(x')|dx' \leqslant B.$$

This completes the proof. □

4.4 Truncated integrals

There is still an element which may be considered unsatisfactory in our formulation, and this is because of the following related points:

1) The L^2 boundedness of the operator has been assumed via the hypothesis that $\widehat{K} \in L^\infty$ and not obtained as a consequence of some condition on the kernel K;

2) An extraneous condition such as $K \in L^2$ subsists in the hypothesis; and for this reason our results do not directly treat the "principal-value" singular integrals, those which exist because of the cancelation of positive and negative values. However, from what we have done, it is now a relatively simple matter to obtain a theorem which covers the cases of interest.

Definition 4.20. Suppose that $K(x) \in L^1_{\text{loc}}(\mathbb{R}^n \setminus \{0\})$ and satisfies the following conditions:

$$\begin{gathered} |K(x)| \leqslant B|x|^{-n}, \quad \forall x \neq 0, \\ \int_{|x| \geqslant 2|y|} |K(x-y) - K(x)| dx \leqslant B, \quad \forall y \neq 0, \end{gathered} \tag{4.34}$$

and

$$\int_{R_1 < |x| < R_2} K(x) dx = 0, \quad \forall 0 < R_1 < R_2 < \infty. \tag{4.35}$$

Then K is called the *Calderón-Zygmund kernel*, where B is a constant independent of x and y.

Theorem 4.21. *Suppose that K is a Calderón-Zygmund kernel. For $\varepsilon > 0$ and $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, let*

$$T_\varepsilon f(x) = \int_{|y| \geqslant \varepsilon} f(x-y) K(y) dy. \tag{4.36}$$

Then the following conclusions hold.

(i) *We have*

$$\|T_\varepsilon f\|_p \leqslant A_p \|f\|_p \tag{4.37}$$

where A_p is independent of f and ε.

(ii) *For any $f \in L^p(\mathbb{R}^n)$, $\lim_{\varepsilon \to 0} T_\varepsilon(f)$ exists in the sense of L^p norm. That is, there exists an operator T such that*

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

(iii) $\|Tf\|_p \leqslant A_p \|f\|_p$ *for $f \in L^p(\mathbb{R}^n)$.*

Remark 4.22. 1) The linear operator T defined by (ii) of Theorem 4.21 is called the *Calderón-Zygmund singular integral operator*. T_ε is also called the *truncated operator* of T.

2) The cancelation property alluded to is contained in condition (4.35). This hypothesis, together with (4.34), allows us to prove the L^2 boundedness and from this the L^p convergence of the truncated integrals (4.37).

3) We should point out that the kernel $K(x) = \frac{1}{\pi x}$, $x \in \mathbb{R}^1$, clearly satisfies the hypotheses of Theorem 4.21. Therefore, we have the existence of the Hilbert transform in the sense that if $f \in L^p(\mathbb{R})$, $1 < p < \infty$, then

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

exists in the L^p norm and the resulting operator is bounded in L^p, as has shown in Theorem 4.16.

For L^2 boundedness, we have the following lemma.

Lemma 4.23. *Suppose K satisfies the conditions (4.34) and (4.35) of the above theorem with bound B. Let*

$$K_\varepsilon(x) = \begin{cases} K(x), & |x| \geqslant \varepsilon, \\ 0, & |x| < \varepsilon. \end{cases}$$

Then, we have the estimate

$$\sup_\xi |\widehat{K}_\varepsilon(\xi)| \leqslant CB, \quad \varepsilon > 0, \tag{4.38}$$

where C depends only on the dimension n.

Proof. First, we prove the inequality (4.38) for the special case $\varepsilon = 1$. Since $\hat{K}_1(0) = 0$, thus we can assume $\xi \neq 0$ and have

$$\begin{aligned}
\widehat{K}_1(\xi) =& \lim_{R\to\infty} \int_{|x|\leqslant R} e^{-\omega i x\cdot\xi} K_1(x)dx \\
=& \int_{|x|<2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi} K_1(x)dx + \lim_{R\to\infty} \int_{2\pi/(|\omega||\xi|)<|x|\leqslant R} e^{-\omega i x\cdot\xi} K_1(x)dx \\
=:& I_1 + I_2.
\end{aligned}$$

By the condition (4.35), $\int_{1<|x|<2\pi/(|\omega||\xi|)} K(x)dx = 0$ which implies

$$\int_{|x|<2\pi/(|\omega||\xi|)} K_1(x)dx = 0.$$

Thus, $\int_{|x|<2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi} K_1(x)dx = \int_{|x|<2\pi/(|\omega||\xi|)} [e^{-\omega i x\cdot\xi} - 1] K_1(x)dx$. Hence, from the fact $|e^{i\theta} - 1| \leqslant |\theta|$ (see Section 1.1) and the first condition in (4.34), we get

$$\begin{aligned}
|I_1| \leqslant& \int_{|x|<2\pi/(|\omega||\xi|)} |\omega||x||\xi||K_1(x)|dx \leqslant |\omega|B|\xi| \int_{|x|<2\pi/(|\omega||\xi|)} |x|^{-n+1}dx \\
=& \omega_{n-1} B|\omega||\xi| \int_0^{2\pi/(|\omega||\xi|)} dr = 2\pi\omega_{n-1}B.
\end{aligned}$$

To estimate I_2, choose $z = z(\xi)$ such that $e^{-\omega i\xi\cdot z} = -1$. This choice can be realized if $z = \pi\xi/(\omega|\xi|^2)$, with $|z| = \pi/(|\omega||\xi|)$. Since, by changing variables $x + z = y$, we get

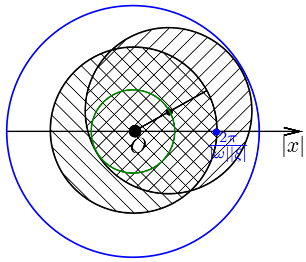
$$\begin{aligned}
\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} K_1(x)dx =& -\int_{\mathbb{R}^n} e^{-\omega i(x+z)\cdot\xi} K_1(x)dx = -\int_{\mathbb{R}^n} e^{-\omega i y\cdot\xi} K_1(y-z)dy \\
=& -\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} K_1(x-z)dx,
\end{aligned}$$

which implies $\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} K_1(x)dx = \frac{1}{2}\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi}[K_1(x) - K_1(x-z)]dx$, then we have

$$\begin{aligned}
I_2 =& \left(\lim_{R\to\infty} \int_{|x|\leqslant R} - \int_{|x|\leqslant 2\pi/(|\omega||\xi|)}\right) e^{-\omega i x\cdot\xi} K_1(x)dx \\
=& \frac{1}{2} \lim_{R\to\infty} \int_{|x|\leqslant R} e^{-\omega i x\cdot\xi}[K_1(x) - K_1(x-z)]dx - \int_{|x|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi} K_1(x)dx \\
=& \frac{1}{2} \lim_{R\to\infty} \int_{2\pi/(|\omega||\xi|)\leqslant|x|\leqslant R} e^{-\omega i x\cdot\xi}[K_1(x) - K_1(x-z)]dx \\
&- \frac{1}{2}\int_{|x|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi} K_1(x)dx - \frac{1}{2}\int_{|x|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi} K_1(x-z)dx.
\end{aligned}$$

The last two integrals are equal to, in view of the integration by parts,

$$
\begin{aligned}
&-\frac{1}{2}\int_{|x|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi}K_1(x)dx - \frac{1}{2}\int_{|y+z|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i(y+z)\cdot\xi}K_1(y)dy\\
=&-\frac{1}{2}\int_{|x|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi}K_1(x)dx + \frac{1}{2}\int_{|x+z|\leqslant 2\pi/(|\omega||\xi|)} e^{-\omega i x\cdot\xi}K_1(x)dx\\
=&-\frac{1}{2}\int_{\substack{|x|\leqslant 2\pi/(|\omega||\xi|)\\ |x+z|>2\pi/(|\omega||\xi|)}} e^{-\omega i x\cdot\xi}K_1(x)dx + \frac{1}{2}\int_{\substack{|x+z|\leqslant 2\pi/(|\omega||\xi|)\\ |x|>2\pi/(|\omega||\xi|)}} e^{-\omega i x\cdot\xi}K_1(x)dx.
\end{aligned}
$$



For the first integral, we have $2\pi/(|\omega||\xi|)\geqslant|x|\geqslant|x+z|-|z|>2\pi/(|\omega||\xi|)-\pi/(|\omega||\xi|)=\pi/(|\omega||\xi|)$, and for the second one, $2\pi/(|\omega||\xi|)<|x|\leqslant|x+z|+|z|\leqslant 3\pi/(|\omega||\xi|)$. These two integrals are taken over a region contained in the spherical shell, $\pi/(|\omega||\xi|)<|x|\leqslant 3\pi/(|\omega||\xi|)$ (see the figure), and is bounded by $\frac{1}{2}B\omega_{n-1}\ln 3$ since $|K_1(x)|\leqslant B|x|^{-n}$. By $|z|=\pi/(|\omega||\xi|)$ and the condition (4.34), the first integral of I_2 is majorized by

$$
\frac{1}{2}\int_{|x|\geqslant 2\pi/(|\omega||\xi|)}|K_1(x-z)-K_1(x)|dx=\frac{1}{2}\int_{|x|\geqslant 2|z|}|K_1(x-z)-K_1(x)|dx\leqslant\frac{1}{2}B.
$$

Thus, we have obtained

$$
|\widehat{K}_1(\xi)|\leqslant 2\pi\omega_{n-1}B+\frac{1}{2}B+\frac{1}{2}B\omega_{n-1}\ln 3\leqslant CB,
$$

where C depends only on n. We finish the proof for K_1.

To pass to the case of general K_ε, we use a simple observation (*dilation argument*) whose significance carries over to the whole theory presented in this chapter.

Let δ_ε be the dilation by the factor $\varepsilon>0$, i.e., $(\delta_\varepsilon f)(x)=f(\varepsilon x)$. Thus if T is a convolution operator

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

then

$$
\delta_{\varepsilon^{-1}}T\delta_\varepsilon f(x)=\int_{\mathbb{R}^n}\varphi(\varepsilon^{-1}x-y)f(\varepsilon y)dy=\varepsilon^{-n}\int_{\mathbb{R}^n}\varphi(\varepsilon^{-1}(x-z))f(z)dz=\varphi_\varepsilon*f,
$$

where $\varphi_\varepsilon(x)=\varepsilon^{-n}\varphi(\varepsilon^{-1}x)$. In our case, if T corresponds to the kernel $K(x)$, then $\delta_{\varepsilon^{-1}}T\delta_\varepsilon$ corresponds to the kernel $\varepsilon^{-n}K(\varepsilon^{-1}x)$. *Notice that if K satisfies the assumptions of our theorem, then $\varepsilon^{-n}K(\varepsilon^{-1}x)$ also satisfies these assumptions with the same bounds.* (A similar remark holds for the assumptions of all the theorems in this chapter.) Now, with our K given, let $K'=\varepsilon^nK(\varepsilon x)$. Then K' satisfies the conditions of our lemma with the same bound B, and so if we denote

$$
K_1'(x)=\begin{cases}K'(x), & |x|\geqslant 1,\\ 0, & |x|<1,\end{cases}
$$

then we know that $|\widehat{K}_1'(\xi)|\leqslant CB$. The Fourier transform of $\varepsilon^{-n}K_1'(\varepsilon^{-1}x)$ is $\widehat{K}_1'(\varepsilon\xi)$ which is again bounded by CB; however $\varepsilon^{-n}K_1'(\varepsilon^{-1}x)=K_\varepsilon(x)$, therefore the lemma is completely proved. □

We can now prove Theorem 4.21.

Proof of Theorem 4.21. Since K satisfies the conditions (4.34) and (4.35), then $K_\varepsilon(x)$ satisfies the same conditions with bounds not greater than CB. By Lemma 4.23 and Theorem 4.18, we have that the L^p boundedness of the operators $\{K_\varepsilon\}_{\varepsilon>0}$, are uniformly bounded.

Next, we prove that $\{T_\varepsilon f_1\}_{\varepsilon>0}$ is a Cauchy sequence in L^p provided $f_1\in C_c^1(\mathbb{R}^n)$. In fact, we have

$$
\begin{aligned}
T_\varepsilon f_1(x) - T_\eta f_1(x) = & \int_{|y|\geqslant\varepsilon} K(y) f_1(x-y)dy - \int_{|y|\geqslant\eta} K(y) f_1(x-y)dy \\
= & \operatorname{sgn}(\eta-\varepsilon) \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} K(y)[f_1(x-y)-f_1(x)]dy,
\end{aligned}
$$

because of the cancelation condition (4.35). For $p \in (1,\infty)$, we get, by the mean value theorem with some $\theta \in [0,1]$, Minkowski's inequality and (4.34), that

$$
\begin{aligned}
\|T_\varepsilon f_1 - T_\eta f_1\|_p \leqslant & \left\| \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} |K(y)||\nabla f_1(x-\theta y)||y|dy \right\|_p \\
\leqslant & \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} |K(y)|\|\nabla f_1(x-\theta y)\|_p |y|dy \\
\leqslant & C \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} |K(y)||y|dy \\
\leqslant & CB \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} |y|^{-n+1}dy \\
= & CB\omega_{n-1} \int_{\min(\varepsilon,\eta)}^{\max(\varepsilon,\eta)} dr \\
= & CB\omega_{n-1}|\eta-\varepsilon|
\end{aligned}
$$

which tends to 0 as $\varepsilon,\eta \to 0$. Thus, we obtain $T_\varepsilon f_1$ converges in L^p as $\varepsilon \to 0$ by the completeness of L^p.

Finally, an arbitrary $f \in L^p$ can be written as $f = f_1 + f_2$ where f_1 is of the type described above and $\|f_2\|_p$ is small. We apply the basic inequality (4.37) for f_2 to get $\|T_\varepsilon f_2\|_p \leqslant C\|f_2\|_p$, then we see that $\lim_{\varepsilon\to 0} T_\varepsilon f$ exists in L^p norm; that the limiting operator T also satisfies the inequality (4.37) is then obvious. Thus, we complete the proof of the theorem. □

4.5 Singular integral operators commuted with dilations

In this section, we shall consider those operators which not only commute with translations but also with dilations. Among these we shall study the class of singular integral operators, falling under the scope of Theorem 4.21.

If T corresponds to the kernel $K(x)$, then as we have already pointed out, $\delta_{\varepsilon^{-1}} T \delta_\varepsilon$ corresponds to the kernel $\varepsilon^{-n} K(\varepsilon^{-1}x)$. So if $\delta_{\varepsilon^{-1}} T \delta_\varepsilon = T$ we are back to the requirement $K(x) = \varepsilon^{-n} K(\varepsilon^{-1}x)$, i.e., $K(\varepsilon x) = \varepsilon^{-n} K(x)$, $\varepsilon > 0$; that is K is homogeneous of degree $-n$. Put another way

$$
K(x) = \frac{\Omega(x)}{|x|^n}, \tag{4.39}
$$

with Ω homogeneous of degree 0, i.e., $\Omega(\varepsilon x) = \Omega(x)$, $\varepsilon > 0$. This condition on Ω is equivalent with the fact that it is constant on rays emanating from the origin; in particular, Ω is completely determined by its restriction to the unit sphere S^{n-1}.

Let us try to reinterpret the conditions of Theorem 4.21 in terms of Ω.

1) By (4.34), $\Omega(x)$ must be bounded and consequently integrable on S^{n-1}; and another condition $\int_{|x|\geqslant 2|y|} \left| \frac{\Omega(x-y)}{|x-y|^n} - \frac{\Omega(x)}{|x|^n} \right| dx \leqslant C$ which is not easily restated precisely in terms of Ω. However, what is evident is that it requires a certain continuity of Ω. Here we shall

content ourselves in treating the case where Ω satisfies the following "*Dini-type*" condition suggested by (4.34):

$$\text{if } w(\eta) := \sup_{\substack{|x-x'|\leqslant\eta \\ |x|=|x'|=1}} |\Omega(x) - \Omega(x')|, \quad \text{then } \int_0^1 \frac{w(\eta)}{\eta} d\eta < \infty. \tag{4.40}$$

Of course, any Ω which is of class C^1, or even merely Lipschitz continuous, satisfies the condition (4.40).

2) The cancelation condition (4.35) is then the same as the condition

$$\int_{S^{n-1}} \Omega(x) d\sigma(x) = 0 \tag{4.41}$$

where $d\sigma(x)$ is the induced Euclidean measure on S^{n-1}. In fact, this equation implies that

$$\begin{aligned}\int_{R_1<|x|<R_2} K(x)dx &= \int_{R_1}^{R_2} \int_{S^{n-1}} \frac{\Omega(rx')}{r^n} d\sigma(x') r^{n-1} dr \\ &= \ln\left(\frac{R_2}{R_1}\right) \int_{S^{n-1}} \Omega(x')d\sigma(x').\end{aligned}$$

Theorem 4.24. *Let $\Omega \in L^\infty(S^{n-1})$ be homogeneous of degree 0, and suppose that Ω satisfies the smoothness property (4.40), and the cancelation property (4.41) above. For $1 < p < \infty$, and $f \in L^p(\mathbb{R}^n)$, let*

$$T_\varepsilon f(x) = \int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy.$$

(a) Then there exists a bound A_p (independent of f and ε) such that

$$\|T_\varepsilon f\|_p \leqslant A_p \|f\|_p.$$

(b) $\lim_{\varepsilon\to 0} T_\varepsilon f = Tf$ exists in L^p norm, and

$$\|Tf\|_p \leqslant A_p \|f\|_p.$$

(c) If $f \in L^2(\mathbb{R}^n)$, then the Fourier transforms of f and Tf are related by $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$, where m is a homogeneous function of degree 0. Explicitly,

$$m(\xi) = \int_{S^{n-1}} \left[-\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(\xi\cdot x) + \ln(1/|\xi\cdot x|)\right] \Omega(x)d\sigma(x), \quad |\xi| = 1. \tag{4.42}$$

Proof. The conclusions (a) and (b) are immediately consequences of Theorem 4.21, once we have shown that any $K(x)$ of the form $\frac{\Omega(x)}{|x|^n}$ satisfies

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \tag{4.43}$$

if Ω is as in condition (4.40). Indeed,

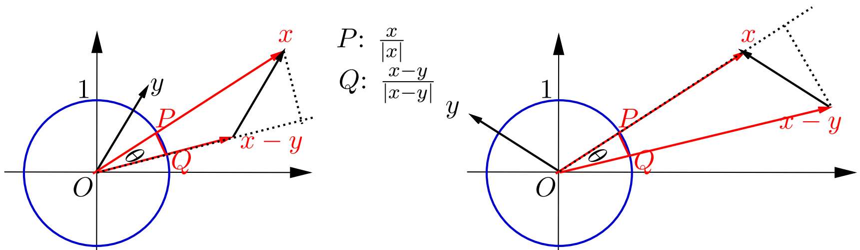
$$K(x-y) - K(x) = \frac{\Omega(x-y) - \Omega(x)}{|x-y|^n} + \Omega(x)\left[\frac{1}{|x-y|^n} - \frac{1}{|x|^n}\right].$$

The second group of terms is bounded since Ω is bounded and

$$\begin{aligned}&\int_{|x|\geqslant 2|y|} \left|\frac{1}{|x-y|^n} - \frac{1}{|x|^n}\right| dx = \int_{|x|\geqslant 2|y|} \left|\frac{|x|^n - |x-y|^n}{|x-y|^n|x|^n}\right| dx \\ =& \int_{|x|\geqslant 2|y|} \frac{||x| - |x-y|| \sum_{j=0}^{n-1} |x|^{n-1-j}|x-y|^j}{|x-y|^n|x|^n} dx \\ \leqslant& \int_{|x|\geqslant 2|y|} |y| \sum_{j=0}^{n-1} |x|^{-j-1}|x-y|^{j-n} dx\end{aligned}$$

$$\leqslant \int_{|x|\geqslant 2|y|} |y| \sum_{j=0}^{n-1} |x|^{-j-1}(|x|/2)^{j-n} dx \quad (\text{since } |x-y| \geqslant |x|-|y| \geqslant |x|/2)$$
$$= \int_{|x|\geqslant 2|y|} |y| \sum_{j=0}^{n-1} 2^{n-j}|x|^{-n-1} dx = 2(2^n-1)|y| \int_{|x|\geqslant 2|y|} |x|^{-n-1} dx$$
$$=2(2^n-1)|y|\omega_{n-1}\frac{1}{2|y|} = (2^n-1)\omega_{n-1}.$$

To estimate the first group of terms, we notice that if $|x| \geqslant 2|y|$, the distance $|PQ|$ between the projections of $x-y$ and x on the unit sphere as in the following picture.



Case 1: $|x| \geq |x-y|$, $\sin\theta \leq \frac{|y|}{|x|}$ Case 2: $|x| \leq |x-y|$, $\sin\theta \leq \frac{|y|}{|x-y|} \leq \frac{|y|}{|x|}$

By the sine theorem, we have $\frac{\sin\theta}{|PQ|} = \frac{\sin\frac{\pi-\theta}{2}}{|OP|}$ where $|OP| = 1$. Since $|y| \leqslant |x|/2$, we have $\theta \leqslant \frac{\pi}{2}$ and so $\cos\theta \geqslant 0$. Thus, $\cos\frac{\theta}{2} = \sqrt{\frac{1+\cos\theta}{2}} \geqslant 1/\sqrt{2}$. Then, we have

$$\left|\frac{x-y}{|x-y|} - \frac{x}{|x|}\right| = |PQ| = \frac{\sin\theta}{\sin(\frac{\pi}{2}-\frac{\theta}{2})} = \frac{\sin\theta}{\cos\frac{\theta}{2}} \leqslant \sqrt{2}\frac{|y|}{|x|} \leqslant 2\frac{|y|}{|x|}$$

since $\sin\theta \leqslant \frac{|y|}{|x|}$ for both cases. So the integral corresponding to the first group of terms is dominated by

$$2^n \int_{|x|\geqslant 2|y|} w\left(2\frac{|y|}{|x|}\right)\frac{dx}{|x|^n} = 2^n \int_{|z|\geqslant 2} w(2/|z|)\frac{dz}{|z|^n} = 2^n\omega_{n-1}\int_2^\infty w(2/r)\frac{dr}{r}$$
$$=2^n\omega_{n-1}\int_0^1 \frac{w(\eta)d\eta}{\eta} < \infty$$

in view of changes of variables $x = |y|z$ and the Dini-type condition (4.40).

Now, we prove (c). Since T is a bounded linear operator on L^2 which commutes with translations, we know, by Theorem 1.62 and Proposition 1.3, that T can be realized in terms of a multiplier m such that $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$. For such operators, the fact that they commute with dilations is equivalent with the property that the multiplier is homogeneous of degree 0.

For our particular operators we have not only the existence of m but also an explicit expression of the multiplier in terms of the kernel. This formula is deduced as follows.

Since $K(x)$ is not integrable, we first consider its truncated function. Let $0 < \varepsilon < \eta < \infty$, and

$$K_{\varepsilon,\eta}(x) = \begin{cases} \dfrac{\Omega(x)}{|x|^n}, & \varepsilon \leqslant |x| \leqslant \eta, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $K_{\varepsilon,\eta} \in L^1(\mathbb{R}^n)$. If $f \in L^2(\mathbb{R}^n)$ then $\widehat{K_{\varepsilon,\eta} * f}(\xi) = \widehat{K_{\varepsilon,\eta}}(\xi)\hat{f}(\xi)$.

We shall prove two facts about $\widehat{K_{\varepsilon,\eta}}(\xi)$.

(i) $\sup_\xi |\widehat{K_{\varepsilon,\eta}}(\xi)| \leqslant A$, with A independent of ε and η;

(ii) if $\xi \neq 0$, $\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}} \widehat{K_{\varepsilon,\eta}}(\xi) = m(\xi)$, see (4.42).

For this purpose, it is convenient to introduce polar coordinates. Let $x = rx'$, $r = |x|$, $x' = x/|x| \in S^{n-1}$, and $\xi = R\xi'$, $R = |\xi|$, $\xi' = \xi/|\xi| \in S^{n-1}$. Then we have

$$\begin{aligned}\widehat{K_{\varepsilon,\eta}}(\xi) =& \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} K_{\varepsilon,\eta}(x)dx = \int_{\varepsilon\leqslant|x|\leqslant\eta} e^{-\omega i x\cdot\xi}\frac{\Omega(x)}{|x|^n}dx\\ =& \int_{S^{n-1}} \Omega(x')\left(\int_\varepsilon^\eta e^{-\omega i Rrx'\cdot\xi'} r^{-n}r^{n-1}dr\right)d\sigma(x')\\ =& \int_{S^{n-1}} \Omega(x')\left(\int_\varepsilon^\eta e^{-\omega i Rrx'\cdot\xi'} \frac{dr}{r}\right)d\sigma(x').\end{aligned}$$

Since

$$\int_{S^{n-1}} \Omega(x')d\sigma(x') = 0,$$

we can introduce the factor $\cos(|\omega|Rr)$ (which does not depend on x') in the integral defining $\widehat{K_{\varepsilon,\eta}}(\xi)$. We shall also need the auxiliary integral

$$I_{\varepsilon,\eta}(\xi,x') = \int_\varepsilon^\eta [e^{-\omega i Rrx'\cdot\xi'} - \cos(|\omega|Rr)]\frac{dr}{r}, \qquad R>0.$$

Thus, it follows

$$\widehat{K_{\varepsilon,\eta}}(\xi) = \int_{S^{n-1}} I_{\varepsilon,\eta}(\xi,x')\Omega(x')d\sigma(x').$$

Now, we first consider $I_{\varepsilon,\eta}(\xi,x')$. For its imaginary part, we have, by changing variable $\omega Rr(x'\cdot\xi') = t$, that

$$\Im I_{\varepsilon,\eta}(\xi,x') = -\int_\varepsilon^\eta \frac{\sin\omega Rr(x'\cdot\xi')}{r}dr = -\operatorname{sgn}(\omega)\operatorname{sgn}(x'\cdot\xi')\int_{|\omega|R\varepsilon(x'\cdot\xi')}^{|\omega|R\eta(x'\cdot\xi')}\frac{\sin t}{t}dt,$$

converges to

$$-\operatorname{sgn}(\omega)\operatorname{sgn}(x'\cdot\xi')\int_0^\infty \frac{\sin t}{t}dt = -\frac{\pi}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(x'\cdot\xi'),$$

as $\varepsilon\to 0$ and $\eta\to\infty$.

For its real part, since $\cos r$ is an even function, we have

$$\Re I_{\varepsilon,\eta}(\xi,x') = \int_\varepsilon^\eta [\cos(|\omega|Rr|x'\cdot\xi'|) - \cos(|\omega|Rr)]\frac{dr}{r}.$$

If $x'\cdot\xi' = \pm 1$, then $\Re I_{\varepsilon,\eta}(\xi,x') = 0$. Now we assume $0<\varepsilon<1<\eta$. For the case $x'\cdot\xi' \neq \pm 1$, we get the absolute value of its real part

$$\begin{aligned}|\Re I_{\varepsilon,\eta}(\xi,x')| \leqslant& \left|\int_\varepsilon^1 -2\sin\frac{|\omega|}{2}Rr(|x'\cdot\xi'|+1)\sin\frac{|\omega|}{2}Rr(|x'\cdot\xi'|-1)\frac{dr}{r}\right|\\ &+\left|\int_1^\eta \cos|\omega|Rr|x'\cdot\xi'|\frac{dr}{r} - \int_1^\eta \cos|\omega|Rr\frac{dr}{r}\right|\\ \leqslant&\frac{|\omega|^2}{2}R^2(1-|x'\cdot\xi'|^2)\int_\varepsilon^1 rdr + \left|\int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R\eta|\xi'\cdot x'|}\frac{\cos t}{t}dt - \int_{|\omega|R}^{|\omega|R\eta}\frac{\cos t}{t}dt\right|\\ \leqslant&\frac{|\omega|^2}{4}R^2 + I_1.\end{aligned}$$

If $\eta|\xi'\cdot x'|>1$, then we have

$$\begin{aligned}I_1 =& \left|\int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R}\frac{\cos t}{t}dt - \int_{|\omega|R\eta|\xi'\cdot x'|}^{|\omega|R\eta}\frac{\cos t}{t}dt\right| \leqslant \int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R}\frac{dt}{t} + \int_{|\omega|R\eta|\xi'\cdot x'|}^{|\omega|R\eta}\frac{dt}{t}\\ \leqslant& 2\ln(1/|\xi'\cdot x'|).\end{aligned}$$

If $0<\eta|\xi'\cdot x'|\leqslant 1$, then

$$I_1 \leqslant \int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R/|\xi'\cdot x'|} \frac{dt}{t} \leqslant 2\ln(1/|\xi'\cdot x'|).$$

Thus,

$$|\Re I_{\varepsilon,\eta}(\xi,x')| \leqslant \frac{|\omega|^2}{4}R^2 + 2\ln(1/|\xi'\cdot x'|),$$

and so the real part converges as $\varepsilon \to 0$ and $\eta \to \infty$. By the fundamental theorem of calculus, we can write

$$\begin{aligned}
&\int_\varepsilon^\eta \frac{\cos(\lambda r)-\cos(\mu r)}{r}dr = -\int_\varepsilon^\eta\int_\mu^\lambda \sin(tr)dtdr = -\int_\mu^\lambda\int_\varepsilon^\eta \sin(tr)drdt\\
=&\int_\mu^\lambda\int_\varepsilon^\eta \frac{\partial_r\cos(tr)}{t}drdt = \int_\mu^\lambda \frac{\cos(t\eta)-\cos(t\varepsilon)}{t}dt\\
=&\int_{\mu\eta}^{\lambda\eta}\frac{\cos(s)}{s}ds - \int_\mu^\lambda \frac{\cos(t\varepsilon)}{t}dt = \frac{\sin s}{s}\Big|_{\mu\eta}^{\lambda\eta} + \int_{\mu\eta}^{\lambda\eta}\frac{\sin s}{s^2}ds - \int_\mu^\lambda\frac{\cos(t\varepsilon)}{t}dt\\
\to&0-\int_\mu^\lambda \frac{1}{t}dt = -\ln(\lambda/\mu) = \ln(\mu/\lambda), \text{ as } \eta\to\infty,\ \varepsilon\to 0.
\end{aligned}$$

Take $\lambda = |\omega|R|x'\cdot\xi'|$, and $\mu = |\omega|R$. So

$$\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}} \Re(I_{\varepsilon,\eta}(\xi,x')) = \int_0^\infty [\cos|\omega|Rr(x'\cdot\xi') - \cos|\omega|Rr]\frac{dr}{r} = \ln(1/|x'\cdot\xi'|).$$

By the properties of $I_{\varepsilon,\eta}$ just proved, we have

$$\begin{aligned}
|\widehat{K_{\varepsilon,\eta}}(\xi)| \leqslant& \int_{S^{n-1}}\left[\frac{\pi}{2} + \frac{|\omega|^2}{4}R^2 + 2\ln(1/|\xi'\cdot x'|)\right]|\varOmega(x')|d\sigma(x')\\
\leqslant& C(\frac{\pi}{2}+\frac{|\omega|^2}{4}R^2)\omega_{n-1} + 2C\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)d\sigma(x').
\end{aligned}$$

For $n=1$, we have $S^0=\{-1,1\}$ and then $\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)d\sigma(x') = 2\ln 1 = 0$. For $n\geqslant 2$, we can pick an orthogonal matrix A such that $Ae_1=\xi'$, and so by changes of variables and using the notation $\bar{y}=(y_2,y_3,...,y_n)$,

$$\begin{aligned}
&\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)d\sigma(x') = \int_{S^{n-1}}\ln(1/|Ae_1\cdot x'|)d\sigma(x')\\
=&\int_{S^{n-1}}\ln(1/|e_1\cdot A^{-1}x'|)d\sigma(x') \overset{A^{-1}x'=y}{=\!=\!=\!=} \int_{S^{n-1}}\ln(1/|e_1\cdot y|)d\sigma(y)\\
=&\int_{S^{n-1}}\ln(1/|y_1|)d\sigma(y) = \int_{-1}^1\ln(1/|y_1|)\int_{\sqrt{1-y_1^2}S^{n-2}}d\sigma(\bar{y})\frac{dy_1}{\sqrt{1-y_1^2}}\\
\overset{\bar{z}=\bar{y}/\sqrt{1-y_1^2}}{=\!=\!=\!=\!=\!=}&\int_{-1}^1\ln(1/|y_1|)\int_{S^{n-2}}(1-y_1^2)^{(n-3)/2}d\sigma(\bar{z})dy_1\\
=&\omega_{n-2}\int_{-1}^1\ln(1/|y_1|)(1-y_1^2)^{(n-3)/2}dy_1\\
=&2\omega_{n-2}\int_0^1\ln(1/|y_1|)(1-y_1^2)^{(n-3)/2}dy_1\\
\overset{y_1=\cos\theta}{=\!=\!=\!=}&2\omega_{n-2}\int_0^{\pi/2}\ln(1/\cos\theta)(\sin\theta)^{n-2}d\theta = 2\omega_{n-2}I_2.
\end{aligned}$$

For $n\geqslant 3$, we have, by integration by parts,

$$I_2 \leqslant \int_0^{\pi/2}\ln(1/\cos\theta)\sin\theta d\theta = \int_0^{\pi/2}\sin\theta d\theta = 1.$$

For $n=2$, we have, by the formula $\int_0^{\pi/2}\ln(\cos\theta)d\theta = -\frac{\pi}{2}\ln 2$ (see [GR, 4.225.3, p.531]),

$$I_2 = \int_0^{\pi/2} \ln(1/\cos\theta)d\theta = -\int_0^{\pi/2} \ln(\cos\theta)d\theta = \frac{\pi}{2}\ln 2.$$

Hence, $\int_{S^{n-1}} \ln(1/|\xi' \cdot x'|)d\sigma(x') \leqslant C$ for any $\xi' \in S^{n-1}$.

Thus, we have proved the uniform boundedness of $\widehat{K_{\varepsilon,\eta}}(\xi)$, i.e., (i). In view of the limit of $I_{\varepsilon,\eta}(\xi, x')$ as $\varepsilon \to 0$, $\eta \to \infty$ just proved, and the dominated convergence theorem, we get

$$\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}} \widehat{K_{\varepsilon,\eta}}(\xi) = m(\xi),$$

if $\xi \neq 0$, that is (ii).

By the Plancherel theorem, if $f \in L^2(\mathbb{R}^n)$, $K_{\varepsilon,\eta} * f$ converges in L^2 norm as $\varepsilon \to 0$ and $\eta \to \infty$, and the Fourier transform of this limit is $m(\xi)\hat{f}(\xi)$.

However, if we keep ε fixed and let $\eta \to \infty$, then clearly $\int K_{\varepsilon,\eta}(y)f(x-y)dy$ converges everywhere to $\int_{|y|\geqslant\varepsilon} K(y)f(x-y)dy$, which is $T_\varepsilon f$.

Letting now $\varepsilon \to 0$, we obtain the conclusion (c) and our theorem is completely proved. □

Remark 4.25. 1) In the theorem, the condition that Ω is mean zero on S^{n-1} is necessary and cannot be neglected. Since in the estimate

$$\int_{\mathbb{R}^n} \frac{\Omega(y)}{|y|^n} f(x-y)dy = \left[\int_{|y|\leqslant 1} + \int_{|y|>1}\right] \frac{\Omega(y)}{|y|^n} f(x-y)dy,$$

the main difficulty lies in the first integral. For instance, if we assume $\Omega(x) \equiv 1$, f is a nonzero constant, then this integral is divergent.

2) From the formula of the symbol $m(\xi)$, it is homogeneous of degree 0 in view of the mean zero property of Ω.

3) The proof of part (c) holds under very general conditions on Ω. Write $\Omega = \Omega_e + \Omega_o$ where Ω_e is the even part of Ω, $\Omega_e(x) = \Omega_e(-x)$, and $\Omega_o(x)$ is the odd part, $\Omega_o(-x) = -\Omega_o(x)$. Then, because of the uniform boundedness of the sine integral, i.e., $\Im I_{\varepsilon,\eta}(\xi, x')$, we required only $\int_{S^{n-1}} |\Omega_o(x')|d\sigma(x') < \infty$, i.e., the integrability of the odd part. For the even part, the proof requires the uniform boundedness of

$$\int_{S^{n-1}} |\Omega_e(x')| \ln(1/|\xi' \cdot x'|)d\sigma(x').$$

This observation is suggestive of certain generalizations of Theorem 4.21, see [Ste70, §6.5, p.49–50].

4.6 The maximal singular integral operator

Theorem 4.24 guaranteed the existence of the singular integral transformation

$$\lim_{\varepsilon\to 0} \int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy \tag{4.44}$$

in the sense of convergence in the L^p norm. The natural counterpart of this result is that of convergence almost everywhere. For the questions involving almost everywhere convergence, it is best to consider also the corresponding maximal function.

Theorem 4.26. *Suppose that Ω satisfies the conditions of the previous theorem. For $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, consider*

$$T_\varepsilon f(x) = \int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy, \quad \varepsilon > 0.$$

(The integral converges absolutely for every x.)

(a) $\lim_{\varepsilon\to 0} T_\varepsilon f(x)$ *exists for almost every* x.

(b) Let $T^* f(x) = \sup_{\varepsilon>0} |T_\varepsilon f(x)|$. *If* $f \in L^1(\mathbb{R}^n)$, *then the mapping* $f \to T^* f$ *is of weak type* $(1,1)$.

(c) If $1 < p < \infty$, *then* $\|T^* f\|_p \leqslant A_p \|f\|_p$.

Proof. The argument for the theorem presents itself in three stages.

The first one is the proof of inequality (c) which can be obtained as a relatively easy consequence of the L^p norm existence of $\lim_{\varepsilon\to 0} T_\varepsilon$, already proved, and certain general properties of "approximations to the identity".

Let $Tf(x) = \lim_{\varepsilon\to 0} T_\varepsilon f(x)$, where the limit is taken in the L^p norm. Its existence is guaranteed by Theorem 4.24. We shall prove this part by showing the following *Cotlar inequality*

$$T^* f(x) \leqslant M(Tf)(x) + CMf(x).$$

Let φ be a smooth non-negative function on $\mathbb{R}^n$, which is supported in the unit ball, has integral equal to one, and which is also radial and decreasing in $|x|$. Consider

$$K_\varepsilon(x) = \begin{cases} \frac{\Omega(x)}{|x|^n}, & |x| \geqslant \varepsilon, \\ 0, & |x| < \varepsilon. \end{cases}$$

This leads us to another function Φ defined by

$$\Phi = \varphi * K - K_1, \tag{4.45}$$

where $\varphi * K = \lim_{\varepsilon\to 0} \varphi * K_\varepsilon = \lim_{\varepsilon\to 0} \int_{|x-y|\geqslant\varepsilon} K(x-y)\varphi(y)dy$.

We shall need to prove that the smallest decreasing radial majorant of Φ is integrable (so as to apply Theorem 4.10). In fact, if $|x| < 1$, then

$$\begin{aligned} |\Phi| =& |\varphi * K| = \left| \int_{\mathbb{R}^n} K(y)\varphi(x-y)dy \right| = \left| \int_{\mathbb{R}^n} K(y)(\varphi(x-y) - \varphi(x))dy \right| \\ \leqslant & \int_{\mathbb{R}^n} |K(y)||\varphi(x-y) - \varphi(x)|dy \leqslant C \int_{\mathbb{R}^n} \frac{|\varphi(x-y) - \varphi(x)|}{|y|^n} dy \leqslant C, \end{aligned}$$

since (4.41) implies $\int_{\mathbb{R}^n} K(y)dy = 0$ and by the smoothness of φ.

If $1 \leqslant |x| \leqslant 2$, then $\Phi = \varphi * K - K$ is again bounded by the same reason and K is bounded in this case.

Finally if $|x| \geqslant 2$,

$$\Phi(x) = \int_{\mathbb{R}^n} K(x-y)\varphi(y)dy - K(x) = \int_{|y|\leqslant 1} [K(x-y) - K(x)]\varphi(y)dy.$$

Similar to (4.43), we can get the bound for $|y| \leqslant 1$

$$\int_{|x|\geqslant 2} |K(x-y) - K(x)|dx \leqslant \int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant C.$$

Thus we obtain

$$\int_{|x|\geqslant 2} |\Phi(x)|dx \leqslant C \int_{|y|\leqslant 1} \varphi(y)dy \leqslant C.$$

Therefore, we have proved that $\Phi \in L^1(\mathbb{R}^n)$ from three cases discussed above.

From (4.45), it follows, because the singular integral operator $\varphi \to \varphi * K$ commutes with dilations, that

$$\varphi_\varepsilon * K - K_\varepsilon = \Phi_\varepsilon, \quad \text{with } \Phi_\varepsilon(x) = \varepsilon^{-n}\Phi(x/\varepsilon). \tag{4.46}$$

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

$$(\varphi_\varepsilon * K) * f(x) = Tf * \varphi_\varepsilon(x), \tag{4.47}$$

where the identity holds for every x. In fact, we notice first that

$$(\varphi_\varepsilon * K_\delta) * f(x) = T_\delta f * \varphi_\varepsilon(x), \quad \text{for every } \delta > 0 \tag{4.48}$$

because both sides of (4.48) are equal for each x to the absolutely convergent double integral $\int_{z\in\mathbb{R}^n}\int_{|y|\geqslant\delta} K(y)f(z-y)\varphi_\varepsilon(x-z)dydz$. Moreover, $\varphi_\varepsilon \in L^q(\mathbb{R}^n)$, with $1<q<\infty$ and $1/p+1/q=1$, so $\varphi_\varepsilon * K_\delta \to \varphi_\varepsilon * K$ in L^q norm, and $T_\delta f \to Tf$ in L^p norm, as $\delta\to 0$, by Theorem 4.24. This proves (4.47), and so by (4.46)

$$T_\varepsilon f = K_\varepsilon * f = \varphi_\varepsilon * K * f - \Phi_\varepsilon * f = Tf * \varphi_\varepsilon - f * \Phi_\varepsilon.$$

Passing to the supremum over ε and applying Theorem 4.10, part (a), Theorem 3.9 for maximal funtions and Theorem 4.24, we get

$$\begin{aligned}\|T^*f\|_p \leqslant& \|\sup_{\varepsilon>0}|Tf*\varphi_\varepsilon|\|_p + \|\sup_{\varepsilon>0}|f*\Phi_\varepsilon|\|_p\\ \leqslant& C\|M(Tf)\|_p + C\|Mf\|_p \leqslant C\|Tf\|_p + C\|f\|_p \leqslant C\|f\|_p.\end{aligned}$$

Thus, we have proved (c).

The second and most difficult stage of the proof is the conclusion (b). Here the argument proceeds in the main as in the proof of the weak type $(1,1)$ result for singular integrals in Theorem 4.18. We review it with deliberate brevity so as to avoid a repetition of details already examined.

For a given $\alpha>0$, we split $f=g+b$ as in the proof of Theorem 4.18. We also consider for each cube Q_j its mate Q_j^*, which has the same center c_j but whose side length is expanded $2\sqrt{n}$ times. The following geometric remarks concerning these cubes are nearly obvious (The first one has given in the proof of Theorem 4.18).

(i) If $x\notin Q_j^*$, then $|x-c_j|\geqslant 2|y-c_j|$ for all $y\in Q_j$, as an obvious geometric consideration shows.

(ii) Suppose $x\in\mathbb{R}^n\setminus Q_j^*$ and assume that for some $y\in Q_j$, $|x-y|=\varepsilon$. Then the closed ball centered at x, of radius $\gamma_n\varepsilon$, contains Q_j, i.e., $B(x,r)\supset Q_j$, if $r=\gamma_n\varepsilon$.

(iii) Under the same hypotheses as (ii), we have that $|x-y|\geqslant\gamma_n'\varepsilon$, for every $y\in Q_j$.

Here γ_n and γ_n' depend only on the dimension n, and not the particular cube Q_j.

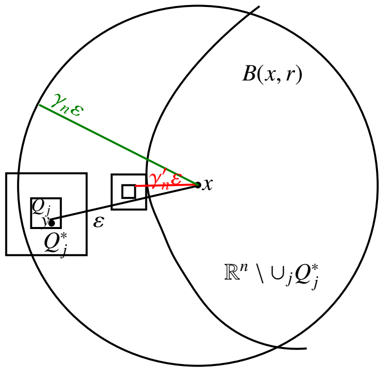


Fig. 4.1 Observation for (ii) and (iii)

With these observations, and following the development in the proof of Theorem 4.18, we shall prove that if $x\in\mathbb{R}^n\setminus\cup_j Q_j^*$,

$$\begin{aligned}\sup_{\varepsilon>0}|T_\varepsilon b(x)| \leqslant& \sum_j\int_{Q_j}|K(x-y)-K(x-c_j)||b(y)|dy\\ &+ C\sup_{r>0}\frac{1}{m(B(x,r))}\int_{B(x,r)}|b(y)|dy,\end{aligned} \tag{4.49}$$

with $K(x)=\frac{\Omega(x)}{|x|^n}$.

The addition of the maximal function to the r.h.s of (4.49) is the main new element of the proof.

To prove (4.49), fix $x\in\mathbb{R}^n\setminus\cup_j Q_j^*$, and $\varepsilon>0$. Now the cubes Q_j fall into three classes:

1) for all $y\in Q_j$, $|x-y|<\varepsilon$;

2) for all $y\in Q_j$, $|x-y|>\varepsilon$;

3) there is a $y\in Q_j$, such that $|x-y|=\varepsilon$.

We now examine

$$T_\varepsilon b(x)=\sum_j\int_{Q_j}K_\varepsilon(x-y)b(y)dy. \tag{4.50}$$

Case 1). $K_\varepsilon(x-y)=0$ if $|x-y|<\varepsilon$, and so the integral over the cube Q_j in (4.50) is zero.

Case 2). $K_\varepsilon(x-y)=K(x-y)$, if $|x-y|>\varepsilon$, and therefore this integral over Q_j equals

$$\int_{Q_j}K(x-y)b(y)dy=\int_{Q_j}[K(x-y)-K(x-c_j)]b(y)dy.$$

This term is majorized in absolute value by

$$\int_{Q_j}|K(x-y)-K(x-c_j)||b(y)|dy,$$

which expression appears in the r.h.s. of (4.49).

Case 3). We write simply

$$\left|\int_{Q_j}K_\varepsilon(x-y)b(y)dy\right|\leqslant\int_{Q_j}|K_\varepsilon(x-y)||b(y)|dy=\int_{Q_j\cap B(x,r)}|K_\varepsilon(x-y)||b(y)|dy,$$

by (ii), with $r=\gamma_n\varepsilon$. However, by (iii) and the fact that Ω is bounded, we have

$$|K_\varepsilon(x-y)|=\left|\frac{\Omega(x-y)}{|x-y|^n}\right|\leqslant\frac{C}{(\gamma_n'\varepsilon)^n}.$$

Thus, in this case,

$$\left|\int_{Q_j}K_\varepsilon(x-y)b(y)dy\right|\leqslant\frac{C}{m(B(x,r))}\int_{Q_j\cap B(x,r)}|b(y)|dy.$$

If we add over all cubes Q_j, we finally obtain, for $r=\gamma_n\varepsilon$,

$$|T_\varepsilon b(x)|\leqslant\sum_j\int_{Q_j}|K(x-y)-K(x-c_j)||b(y)|dy+C\frac{1}{m(B(x,r))}\int_{B(x,r)}|b(y)|dy.$$

Taking the supremum over ε gives (4.49).

This inequality can be written in the form

$$|T^*b(x)|\leqslant\Sigma+CMb(x),\quad x\in F^*,$$

and so

$$\begin{aligned}&\mathfrak{m}(\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:|T^*b(x)|>\alpha/2\})\\ \leqslant&\mathfrak{m}(\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:\Sigma>\alpha/4\})+\mathfrak{m}(\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:CMb(x)>\alpha/4\}).\end{aligned}$$

The first term in the r.h.s. is similar to (4.33), and we can get $\int_{\mathbb{R}^n\setminus\cup_jQ_j^*}\Sigma(x)dx\leqslant C\|b\|_1$ which implies $\mathfrak{m}(\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:\Sigma>\alpha/4\})\leqslant\frac{4C}{\alpha}\|b\|_1$.

For the second one, by Theorem 3.9, i.e., the weak type estimate for the maximal function M, we get $m(\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:CMb(x)>\alpha/4\})\leqslant\frac{C}{\alpha}\|b\|_1$.

The weak type $(1,1)$ property of T^* then follows as in the proof of the same property for T, in Theorem 4.18 for more details.

The final stage of the proof, the passage from the inequalities of T^* to the existence of the limits almost everywhere, follows the familiar pattern described in the proof of the Lebesgue differential theorem (i.e., Theorem3.13).

More precisely, for any $f\in L^p(\mathbb{R}^n)$, $1\leqslant p<\infty$, let

$$\Lambda f(x)=\left|\limsup_{\varepsilon\to0}T_\varepsilon f(x)-\liminf_{\varepsilon\to0}T_\varepsilon f(x)\right|.$$

Clearly, $\Lambda f(x)\leqslant 2T^*f(x)$. Now write $f=f_1+f_2$ where $f_1\in C_c^1$, and $\|f_2\|_p\leqslant\delta$.

We have already proved in the proof of Theorem 4.21 that $T_\varepsilon f_1$ converges uniformly as $\varepsilon \to 0$, so $\Lambda f_1(x) \equiv 0$. By (4.37), we have $\|\Lambda f_2\|_p \leqslant 2A_p\|f_2\|_p \leqslant 2A_p\delta$ if $1 < p < \infty$. This shows $\Lambda f_2 = 0$, almost everywhere, thus by $\Lambda f(x) \leqslant \Lambda f_1(x) + \Lambda f_2(x)$, we have $\Lambda f = 0$ almost everywhere. So $\lim_{\varepsilon\to 0} T_\varepsilon f$ exists almost everywhere if $1 < p < \infty$.

In the case $p = 1$, we get similarly

$$\mathfrak{m}(\{x : \Lambda f(x) > \alpha\}) \leqslant \frac{A}{\alpha}\|f_2\|_1 \leqslant \frac{A\delta}{\alpha},$$

and so again $\Lambda f(x) = 0$ almost everywhere, which implies that $\lim_{\varepsilon\to 0} T_\varepsilon f(x)$ exists almost everywhere. □

4.7 Vector-valued analogues

It is interesting to point out that the results of this chapter, where our functions were assumes to take real or complex values, can be extended to the case of functions taking their values in a Hilbert space. We present this generalization because it can be put to good use in several problems. An indication of this usefulness is given in the Littlewood-Paley theory.

We begin by reviewing quickly certain aspects of integration theory in this context.

Let $\mathscr{H}$ be a separable Hilbert space. Then a function $f(x)$, from $\mathbb{R}^n$ to $\mathscr{H}$ is *measurable* if the scalar valued functions $(f(x), \varphi)$ are measurable, where $(\cdot,\cdot)$ denotes the inner product of $\mathscr{H}$, and φ denotes an arbitrary vector of $\mathscr{H}$.

If $f(x)$ is such a measurable function, then $|f(x)|$ is also measurable (as a function with non-negative values), where $|\cdot|$ denotes the norm of $\mathscr{H}$.

Thus, $L^p(\mathbb{R}^n, \mathscr{H})$ is defined as the equivalent classes of measurable functions $f(x)$ from $\mathbb{R}^n$ to $\mathscr{H}$, with the property that the norm $\|f\|_p = (\int_{\mathbb{R}^n} |f(x)|^p dx)^{1/p}$ is finite, when $p < \infty$; when $p = \infty$ there is a similar definition, except $\|f\|_\infty = \operatorname{ess\,sup}|f(x)|$.

Next, let $\mathscr{H}_1$ and $\mathscr{H}_2$ be two separable Hilbert spaces, and let $L(\mathscr{H}_1, \mathscr{H}_2)$ denote the Banach space of bounded linear operators from $\mathscr{H}_1$ to $\mathscr{H}_2$, with the usual operator norm.

We say that a function $f(x)$, from $\mathbb{R}^n$ to $L(\mathscr{H}_1, \mathscr{H}_2)$ is measurable if $f(x)\varphi$ is an $\mathscr{H}_2$-valued measurable function for every $\varphi \in \mathscr{H}_1$. In this case $|f(x)|$ is also measurable and we can define the space $L^p(\mathbb{R}^n, L(\mathscr{H}_1, \mathscr{H}_2))$, as before; here again $|\cdot|$ denotes the norm, this time in $L(\mathscr{H}_1, \mathscr{H}_2)$.

The usual facts about convolution hold in this setting. For example, suppose $K(x) \in L^q(\mathbb{R}^n, L(\mathscr{H}_1, \mathscr{H}_2))$ and $f(x) \in L^p(\mathbb{R}^n, \mathscr{H}_1)$, then $g(x) = \int_{\mathbb{R}^n} K(x-y)f(y)dy$ converges in the norm of $\mathscr{H}_2$ for almost every x, and

$$|g(x)| \leqslant \int_{\mathbb{R}^n} |K(x-y)f(y)|dy \leqslant \int_{\mathbb{R}^n} |K(x-y)||f(y)|dy.$$

Also $\|g\|_r \leqslant \|K\|_q\|f\|_p$, if $1/r = 1/p + 1/q - 1$, with $1 \leqslant r \leqslant \infty$.

Suppose that $f(x) \in L^1(\mathbb{R}^n, \mathscr{H})$. Then we can define its Fourier transform $\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-\omega ix\cdot\xi} f(x)dx$ which is an element of $L^\infty(\mathbb{R}^n, \mathscr{H})$. If $f \in L^1(\mathbb{R}^n, \mathscr{H}) \cap L^2(\mathbb{R}^n, \mathscr{H})$, then $\hat{f}(\xi) \in L^2(\mathbb{R}^n, \mathscr{H})$ with $\|\hat{f}\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|f\|_2$. The Fourier transform can then be extended by continuity to a unitary mapping of the Hilbert space $L^2(\mathbb{R}^n, \mathscr{H})$ to itself, up to a constant multiplication.

These facts can be obtained easily from the scalar-valued case by introducing an arbitrary orthonormal basis in $\mathscr{H}$.

Now suppose that $\mathscr{H}_1$ and $\mathscr{H}_2$ are two given Hilbert spaces. Assume that $f(x)$ takes values in $\mathscr{H}_1$, and $K(x)$ takes values in $L(\mathscr{H}_1, \mathscr{H}_2)$. Then

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

whenever defined, takes values in $\mathscr{H}_2$.

Theorem 4.27. *The results in this chapter, in particular Theorem 4.18, Proposition 4.19, Theorems 4.21, 4.24 and 4.26 are valid in the more general context where f takes its value in $\mathscr{H}_1$, K takes its values in $L(\mathscr{H}_1, \mathscr{H}_2)$ and Tf and $T_\varepsilon f$ take their value in $\mathscr{H}_2$, and where throughout the absolute value $|\cdot|$ is replaced by the appropriate norm in $\mathscr{H}_1$, $L(\mathscr{H}_1, \mathscr{H}_2)$ or $\mathscr{H}_2$ respectively.*

This theorem is not a corollary of the scalar-valued case treated in any obvious way. However, its proof consists of nothing but a identical repetition of the arguments given for the scalar-valued case, if we take into account the remarks made in the above paragraphs. So, we leave the proof to the interested reader.

Remark 4.28. 1) The final bounds obtained do not depend on the Hilbert spaces $\mathscr{H}_1$ or $\mathscr{H}_2$, but only on B, p, and n, as in the scalar-valued case.

2) Most of the argument goes through in the even greater generality of Banach space-valued functions, appropriately defined. The Hilbert space structure is used only in the L^2 theory when applying the variant of Plancherel's formula.

The Hilbert space structure also enters in the following corollary.

Corollary 4.29. *With the same assumptions as in Theorem 4.27, if in addition*

$$\|Tf\|_2 = c\|f\|_2, \quad c>0, \quad f \in L^2(\mathbb{R}^n, \mathscr{H}_1),$$

then $\|f\|_p \leqslant A_p'\|Tf\|_p$, if $f \in L^p(\mathbb{R}^n, \mathscr{H}_1)$, if $1<p<\infty$.

Proof. We remark that the $L^2(\mathbb{R}^n, \mathscr{H}_j)$ are Hilbert spaces. In fact, let $(\cdot,\cdot)_j$ denote the inner product of $\mathscr{H}_j$, $j=1,2$, and let $\langle\cdot,\cdot\rangle_j$ denote the corresponding inner product in $L^2(\mathbb{R}^n, \mathscr{H}_j)$; that is

$$\langle f, g\rangle_j = \int_{\mathbb{R}^n} (f(x), g(x))_j dx.$$

Now T is a bounded linear transformation from the Hilbert space $L^2(\mathbb{R}^n, \mathscr{H}_1)$ to the Hilbert space $L^2(\mathbb{R}^n, \mathscr{H}_2)$, and so by the general theory of inner products there exists a unique adjoint transformation $\tilde{T}$, from $L^2(\mathbb{R}^n, \mathscr{H}_2)$ to $L^2(\mathbb{R}^n, \mathscr{H}_1)$, which satisfies the characterizing property

$$\langle Tf_1, f_2\rangle_2 = \langle f_1, \tilde{T}f_2\rangle_1, \quad \text{with } f_j \in L^2(\mathbb{R}^n, \mathscr{H}_j).$$

But our assumption is equivalent with the identity (see the theory of Hilbert spaces, e.g. [Din07, Chapter 6])

$$\langle Tf, Tg\rangle_2 = c^2\langle f, g\rangle_1, \quad \text{for all } f, g \in L^2(\mathbb{R}^n, \mathscr{H}_1).$$

Thus using the definition of the adjoint, $\langle \tilde{T}Tf, g\rangle_1 = c^2\langle f, g\rangle_1$, and so the assumption can be restated as

$$\tilde{T}Tf = c^2 f, \quad f \in L^2(\mathbb{R}^n, \mathscr{H}_1). \tag{4.51}$$

$\tilde{T}$ is again an operator of the same kind as T but it takes function with values in $\mathscr{H}_2$ to functions with values in $\mathscr{H}_1$, and its kernel $\tilde{K}(x) = K^*(-x)$, where $*$ denotes the adjoint of an element in $L(\mathscr{H}_1, \mathscr{H}_2)$.

This is obvious on the formal level since

$$\begin{aligned}\langle Tf_1, f_2\rangle_2 &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (K(x-y)f_1(y), f_2(x))_2 dydx \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (f_1(y), K^*(-(y-x))f_2(x))_1 dxdy = \langle f_1, \tilde{T}f_2\rangle_1.\end{aligned}$$

The rigorous justification of this identity is achieved by a simple limiting argument. We will not tire the reader with the routine details.

This being said we have only to add the remark that $K^*(-x)$ satisfies the same conditions as $K(x)$, and so we have, for it, similar conclusions as for K (with the same bounds). Thus by (4.51),

$$c^2\|f\|_p = \|\tilde{T}Tf\|_p \leqslant A_p\|Tf\|_p.$$

This proves the corollary with $A_p' = A_p/c^2$. □

Remark 4.30. This corollary applies in particular to the singular integrals commuted with dilations, then the condition required is that the multiplier $m(\xi)$ have constant absolute value. This is the case, for example, when T is the Hilbert transform, $K(x) = \frac{1}{\pi x}$, and $m(\xi) = -i\,\mathrm{sgn}\,(\omega)\,\mathrm{sgn}\,(\xi)$.

Chapter 5
Riesz Transforms and Spherical Harmonics

5.1 The Riesz transforms

We look for the operators in $\mathbb{R}^n$ which have the analogous structural characterization as the Hilbert transform. We begin by making a few remarks about the interaction of rotations with the n-dimensional Fourier transform. We shall need the following elementary observation.

Let ρ denote any rotation about the origin in $\mathbb{R}^n$. Denote also by ρ its induced action on functions, $\rho(f)(x) = f(\rho x)$. Then

$$\begin{aligned}(\mathscr{F}\rho)f(\xi) =&\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} f(\rho x)dx = \int_{\mathbb{R}^n} e^{-\omega i \rho^{-1}y\cdot\xi} f(y)dy = \int_{\mathbb{R}^n} e^{-\omega i y\cdot\rho\xi} f(y)dy\\ =&\mathscr{F}f(\rho\xi) = \rho\mathscr{F}f(\xi),\end{aligned}$$

that is,

$$\mathscr{F}\rho = \rho\mathscr{F}.$$

Let $\ell(x) = (\ell_1(x), \ell_2(x), ..., \ell_n(x))$ be an n-tuple of functions defined on $\mathbb{R}^n$. For any rotation ρ about the origin, write $\rho = (\rho_{jk})$ for its matrix realization. Suppose that ℓ transforms like a vector. Symbolically this can be written as

$$\ell(\rho x) = \rho(\ell(x)),$$

or more explicitly

$$\ell_j(\rho x) = \sum_k \rho_{jk}\ell_k(x), \quad \text{for every rotation } \rho. \tag{5.1}$$

Lemma 5.1. *Suppose ℓ is homogeneous of degree 0, i.e., $\ell(\varepsilon x) = \ell(x)$, for $\varepsilon > 0$. If ℓ transforms according to (5.1) then $\ell(x) = c\frac{x}{|x|}$ for some constant c; that is*

$$\ell_j(x) = c\frac{x_j}{|x|}. \tag{5.2}$$

Proof. It suffices to consider $x \in S^{n-1}$ due to the homogeneousness of degree 0 for ℓ. Now, let e_1, e_2, ..., e_n denote the usual unit vectors along the axes. Set $c = \ell_1(e_1)$. We can see that $\ell_j(e_1) = 0$, if $j \neq 1$.

In fact, we take a rotation arbitrarily such that e_1 fixed under the acting of ρ, i.e., $\rho e_1 = e_1$. Thus, we also have $e_1 = \rho^{-1}\rho e_1 = \rho^{-1}e_1 = \rho^\top e_1$. From $\rho e_1 = \rho^\top e_1 = e_1$, we get $\rho_{11} = 1$ and $\rho_{1k} = \rho_{j1} = 0$ for $k \neq 1$ and $j \neq 1$. So $\rho = \begin{pmatrix}1 & 0\\ 0 & A\end{pmatrix}$. Because $\begin{pmatrix}1 & 0\\ 0 & A\end{pmatrix}^{-1} = \begin{pmatrix}1 & 0\\ 0 & A^{-1}\end{pmatrix}$ and $\rho^{-1} = \rho^\top$, we obtain $A^{-1} = A^\top$ and $\det A = 1$, i.e., A

is a rotation in $\mathbb{R}^{n-1}$. On the other hand, by (5.1), we get $\ell_j(e_1) = \sum_{k=2}^n \rho_{jk}\ell_k(e_1)$ for $j = 2, ..., n$. That is, the $n-1$ dimensional vector $(\ell_2(e_1), \ell_3(e_1), \cdots, \ell_n(e_1))$ is left fixed by all the rotations on this $n-1$ dimensional vector space. Thus, we have to take $\ell_2(e_1) = \ell_3(e_1) = \cdots = \ell_n(e_1) = 0$.

Inserting again in (5.1) gives $\ell_j(\rho e_1) = \rho_{j1}\ell_1(e_1) = c\rho_{j1}$. If we take a rotation such that $\rho e_1 = x$, then we have $\rho_{j1} = x_j$, so $\ell_j(x) = cx_j$, $(|x| = 1)$, which proves the lemma. □

We now define the n Riesz transforms. For $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, we set

$$R_j f(x) = \lim_{\varepsilon\to 0} c_n \int_{|y|\geqslant\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)dy, \quad j = 1, ..., n, \tag{5.3}$$

with $c_n = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}$ where $1/c_n = \frac{\pi^{(n+1)/2}}{\Gamma((n+1)/2)}$ is half the surface area of the unit sphere S^n of $\mathbb{R}^{n+1}$. Thus, R_j is defined by the kernel $K_j(x) = \frac{\Omega_j(x)}{|x|^n}$, and $\Omega_j(x) = c_n \frac{x_j}{|x|}$.

Next, we derive the multipliers which correspond to the Riesz transforms, and which in fact justify their definition. Denote

$$\Omega(x) = (\Omega_1(x), \Omega_2(x), ..., \Omega_n(x)), \text{ and } m(\xi) = (m_1(\xi), m_2(\xi), ..., m_n(\xi)).$$

Let us recall the formula (4.42), i.e.,

$$m(\xi) = \int_{S^{n-1}} \Phi(\xi \cdot x)\Omega(x)d\sigma(x), \quad |\xi| = 1, \tag{5.4}$$

with $\Phi(t) = -\frac{\pi i}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(t) + \ln|1/t|$. For any rotation ρ, since Ω commutes with any rotations, i.e., $\Omega(\rho x) = \rho(\Omega(x))$, we have, by changes of variables,

$$\begin{aligned}
\rho(m(\xi)) &= \int_{S^{n-1}} \Phi(\xi\cdot x)\rho(\Omega(x))d\sigma(x) = \int_{S^{n-1}} \Phi(\xi\cdot x)\Omega(\rho x)d\sigma(x)\\
&= \int_{S^{n-1}} \Phi(\xi\cdot\rho^{-1}y)\Omega(y)d\sigma(y) = \int_{S^{n-1}} \Phi(\rho\xi\cdot y)\Omega(y)d\sigma(y)\\
&= m(\rho\xi).
\end{aligned}$$

Thus, m commutes with rotations and so m satisfies (5.1). However, the m_j are each homogeneous of degree 0, so Lemma 5.1 shows that $m_j(\xi) = c\frac{\xi_j}{|\xi|}$, with

$$\begin{aligned}
c = m_1(e_1) &= \int_{S^{n-1}} \Phi(e_1\cdot x)\Omega_1(x)d\sigma(x)\\
&= \int_{S^{n-1}} [-\frac{\pi i}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(x_1) + \ln|1/x_1|]c_n x_1 d\sigma(x)\\
&= -\operatorname{sgn}(\omega)\frac{\pi i}{2}c_n \int_{S^{n-1}} |x_1|d\sigma(x) \quad \text{(the 2nd is 0 since it is odd w.r.t. } x_1\text{)}\\
&= -\operatorname{sgn}(\omega)\frac{\pi i}{2}\frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}\frac{2\pi^{(n-1)/2}}{\Gamma((n+1)/2)} = -\operatorname{sgn}(\omega)i.
\end{aligned}$$

Here we have used the fact $\int_{S^{n-1}} |x_1|d\sigma(x) = 2\pi^{(n-1)/2}/\Gamma((n+1)/2)$. Therefore, we obtain

$$\widehat{R_j f}(\xi) = -\operatorname{sgn}(\omega)i\frac{\xi_j}{|\xi|}\hat{f}(\xi), \quad j = 1, ..., n. \tag{5.5}$$

This identity and Plancherel's theorem also imply the following "unitary" character of the Riesz transforms

$$\sum_{j=1}^n \|R_j f\|_2^2 = \|f\|_2^2.$$

By $m(\rho\xi) = \rho(m(\xi))$ proved above, we have $m_j(\rho\xi) = \sum_k \rho_{jk}m_k(\xi)$ for any rotation ρ and then $m_j(\rho\xi)\hat{f}(\xi) = \sum_k \rho_{jk}m_k(\xi)\hat{f}(\xi)$. Taking the inverse Fourier transform, it follows

$$\mathscr{F}^{-1}m_j(\rho\xi)\hat{f}(\xi) = \mathscr{F}^{-1}\sum_k \rho_{jk}m_k(\xi)\hat{f}(\xi) = \sum_k \rho_{jk}\mathscr{F}^{-1}m_k(\xi)\hat{f}(\xi) = \sum_k \rho_{jk}R_k f.$$

But by changes of variables, we have

$$\begin{aligned}\mathscr{F}^{-1}m_j(\rho\xi)\hat{f}(\xi) &= \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} m_j(\rho\xi)\hat{f}(\xi)d\xi \\ &= \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i\rho x\cdot\eta} m_j(\eta)\hat{f}(\rho^{-1}\eta)d\eta \\ &=(\mathscr{F}^{-1}(m_j(\xi)\hat{f}(\rho^{-1}\xi)))(\rho x) = \rho\mathscr{F}^{-1}(m_j(\xi)\hat{f}(\rho^{-1}\xi))(x) \\ &=\rho R_j\rho^{-1}f,\end{aligned}$$

since the Fourier transform commutes with rotations. Therefore, it reaches

$$\rho R_j\rho^{-1}f = \sum_k \rho_{jk}R_k f, \tag{5.6}$$

which is the statement that under rotations in $\mathbb{R}^n$, the Riesz operators transform in the same manner as the components of a vector.

We have the following characterization of Riesz transforms.

Proposition 5.2. *Let $T = (T_1, T_2, ..., T_n)$ be an n-tuple of bounded linear transforms on $L^2(\mathbb{R}^n)$. Suppose*

(a) *Each T_j commutes with translations of $\mathbb{R}^n$;*
(b) *Each T_j commutes with dilations of $\mathbb{R}^n$;*
(c) *For every rotation $\rho = (\rho_{jk})$ of $\mathbb{R}^n$, $\rho T_j\rho^{-1}f = \sum_k \rho_{jk}T_k f$.*

Then the T_j is a constant multiple of the Riesz transforms, i.e., there exists a constant c such that $T_j = cR_j$, $j = 1, ..., n$.

Proof. All the elements of the proof have already been discussed. We bring them together.

(i) Since the T_j is bounded linear on $L^2(\mathbb{R}^n)$ and commutes with translations, by Theorem 1.62 they can be each realized by bounded multipliers m_j, i.e., $\mathscr{F}(T_j f) = m_j\hat{f}$.

(ii) Since the T_j commutes with dilations, i.e., $T_j\delta_\varepsilon f = \delta_\varepsilon T_j f$, in view of Proposition 1.3, we see that $\mathscr{F}T_j\delta_\varepsilon f = m_j(\xi)\mathscr{F}\delta_\varepsilon f = m_j(\xi)\varepsilon^{-n}\delta_{\varepsilon^{-1}}\hat{f}(\xi) = m_j(\xi)\varepsilon^{-n}\hat{f}(\xi/\varepsilon)$ and $\mathscr{F}\delta_\varepsilon T_j f = \varepsilon^{-n}\delta_{\varepsilon^{-1}}\mathscr{F}T_j f = \varepsilon^{-n}\delta_{\varepsilon^{-1}}(m_j\hat{f}) = \varepsilon^{-n}m_j(\xi/\varepsilon)\hat{f}(\xi/\varepsilon)$, which imply $m_j(\xi) = m_j(\xi/\varepsilon)$ or equivalently $m_j(\varepsilon\xi) = m_j(\xi)$, $\varepsilon > 0$; that is, each m_j is homogeneous of degree 0.

(iii) Finally, assumption (c) has a consequence by taking the Fourier transform, i.e., the relation (5.1), and so by Lemma 5.1, we can obtain the desired conclusion. □

One of the important applications of the Riesz transforms is that they can be used to mediate between various combinations of partial derivatives of a function.

Proposition 5.3. *Suppose $f \in C_c^2(\mathbb{R}^n)$. Let $\Delta f = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}$. Then we have the a priori bound*

$$\left\|\frac{\partial^2 f}{\partial x_j\partial x_k}\right\|_p \leqslant A_p\|\Delta f\|_p, \quad 1 < p < \infty. \tag{5.7}$$

Proof. Since $\mathscr{F}(\partial_{x_j}f)(\xi) = \omega i\xi_j\mathscr{F}f(\xi)$, we have

$$\begin{aligned}\mathscr{F}\left(\frac{\partial^2 f}{\partial x_j\partial x_k}\right)(\xi) =& -\omega^2\xi_j\xi_k\mathscr{F}f(\xi) \\ =& -\left(-\operatorname{sgn}(\omega)\frac{i\xi_j}{|\xi|}\right)\left(-\operatorname{sgn}(\omega)\frac{i\xi_k}{|\xi|}\right)(-\omega^2|\xi|^2)\mathscr{F}f(\xi) \\ =& -\mathscr{F}R_jR_k\Delta f.\end{aligned}$$

Thus, $\frac{\partial^2 f}{\partial x_j\partial x_k} = -R_jR_k\Delta f$. By the L^p boundedness of the Riesz transforms, we have the desired result. □

Proposition 5.4. *Suppose $f \in C_c^1(\mathbb{R}^2)$. Then we have the a priori bound*
$$\left\|\frac{\partial f}{\partial x_1}\right\|_p + \left\|\frac{\partial f}{\partial x_2}\right\|_p \leqslant A_p \left\|\frac{\partial f}{\partial x_1} + i\frac{\partial f}{\partial x_2}\right\|_p, \quad 1 < p < \infty.$$

Proof. The proof is similar to the previous one. Indeed, we have
$$\begin{aligned}\mathscr{F}\partial_{x_j} f =&\omega i\xi_j \mathscr{F} f(\xi) = \omega\frac{i\xi_j}{|\xi|}|\xi|\mathscr{F} f(\xi) = \omega\frac{i\xi_j}{|\xi|}\frac{\xi_1^2+\xi_2^2}{|\xi|}\mathscr{F} f(\xi)\\ =&\omega\frac{i\xi_j}{|\xi|}\frac{(\xi_1 - i\xi_2)(\xi_1+i\xi_2)}{|\xi|}\mathscr{F} f(\xi)\\ =&-\frac{-\operatorname{sgn}(\omega)i\xi_j}{|\xi|}\frac{-\operatorname{sgn}(\omega)i(\xi_1-i\xi_2)}{|\xi|}\mathscr{F}(\partial_{x_1} f + i\partial_{x_2} f)\\ =&-\mathscr{F} R_j(R_1 - iR_2)(\partial_{x_1} f + i\partial_{x_2} f).\end{aligned}$$
That is, $\partial_{x_j} f = -R_j(R_1 - iR_2)(\partial_{x_1} f + i\partial_{x_2} f)$. Also by the L^p boundedness of the Riesz transforms, we can obtain the result. □

We shall now tie together the Riesz transforms and the theory of harmonic functions, more particularly Poisson integrals. Since we are interested here mainly in the formal aspects we shall restrict ourselves to the L^2 case. For L^p case, one can see the further results in [Ste70, §4.3 and §4.4, p.78].

Theorem 5.5. *Let f and f_1, ..., f_n all belong to $L^2(\mathbb{R}^n)$, and let their respective Poisson integrals be $u_0(x,y) = P_y * f$, $u_1(x,y) = P_y * f_1$, ..., $u_n(x,y) = P_y * f_n$. Then a necessary and sufficient condition of*
$$f_j = R_j(f), \quad j = 1, ..., n, \tag{5.8}$$
is that the following generalized Cauchy-Riemann equations hold:
$$\begin{cases}\displaystyle\sum_{j=0}^n \frac{\partial u_j}{\partial x_j} = 0,\\ \displaystyle\frac{\partial u_j}{\partial x_k} = \frac{\partial u_k}{\partial x_j}, \quad j \neq k, \quad \text{with } x_0 = y.\end{cases} \tag{5.9}$$

Remark 5.6. At least locally, the system (5.9) is equivalent with the existence of a harmonic function g of the $n+1$ variables, such that $u_j = \frac{\partial g}{\partial x_j}$, $j = 0, 1, 2, ..., n$.

Proof. Suppose $f_j = R_j f$, then $\widehat{f_j}(\xi) = -\operatorname{sgn}(\omega)\frac{i\xi_j}{|\xi|}\hat{f}(\xi)$, and so by (4.15)
$$u_j(x,y) = -\operatorname{sgn}(\omega)\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \hat{f}(\xi)\frac{i\xi_j}{|\xi|}e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi, \quad j = 1, ..., n,$$
and
$$u_0(x,y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \hat{f}(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi.$$

The equation (5.9) can then be immediately verified by differentiation under the integral sign, which is justified by the rapid convergence of the integrals in question.

Conversely, let $u_j(x,y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \widehat{f_j}(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi$, $j = 0, 1, ..., n$ with $f_0 = f$. Then the fact that $\frac{\partial u_0}{\partial x_j} = \frac{\partial u_j}{\partial x_0} = \frac{\partial u_j}{\partial y}$, $j = 1, ..., n$, and Fourier inversion theorem, show that
$$\omega i\xi_j \widehat{f_0}(\xi)e^{-|\omega\xi|y} = -|\omega\xi|\widehat{f_j}(\xi)e^{-|\omega\xi|y},$$
therefore $\widehat{f_j}(\xi) = -\operatorname{sgn}(\omega)\frac{i\xi_j}{|\xi|}\widehat{f_0}(\xi)$, and so
$$f_j = R_j f_0 = R_j f, \quad j = 1, ..., n.$$
□

5.2 Spherical harmonics and higher Riesz transforms

We return to the consideration of special transforms of the form

$$Tf(x) = \lim_{\varepsilon \to 0} \int_{|y| \geqslant \varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy, \tag{5.10}$$

where Ω is homogeneous of degree 0 and its integral over S^{n-1} vanishes.

We have already considered the example, i.e., the case of Riesz transforms, $\Omega_j(y) = c\frac{y_j}{|y|}$, $j = 1, ..., n$. For $n = 1$, $\Omega(y) = c \operatorname{sgn} y$, and this is the only possible case, i.e., the Hilbert transform. To study the matter further for $n > 1$, we recall the expression

$$m(\xi) = \int_{S^{n-1}} \Lambda(y \cdot \xi)\Omega(y)d\sigma(y), \quad |\xi| = 1$$

where m is the multiplier arising from the transform (5.10).

We have already remarked that the mapping $\Omega \to m$ commutes with rotations. We shall therefore consider the functions on the sphere S^{n-1} (more particularly the space $L^2(S^{n-1})$) from the point of view of its decomposition under the action of rotations. As is well known, this decomposition is in terms of the spherical harmonics, and it is with a brief review of their properties that we begin.

We fix our attention, as always, on $\mathbb{R}^n$, and we shall consider polynomials in $\mathbb{R}^n$ which are also harmonic.

Definition 5.7. Denote $\alpha = (\alpha_1, ..., \alpha_n)$, $|\alpha| = \sum_{j=1}^n \alpha_j$ and $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Let $\mathscr{P}_k$ denote the linear space of all homogeneous polynomials of degree k, i.e.,

$$\mathscr{P}_k := \left\{ P(x) = \sum a_\alpha x^\alpha : |\alpha| = k \right\}.$$

Each such polynomial corresponds its dual object, the differential operator $P(\partial_x) = \sum a_\alpha \partial_x^\alpha$, where $\partial_x^\alpha = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}$. On $\mathscr{P}_k$, we define a positive inner product $\langle P, Q \rangle = P(\partial_x)\bar{Q}$. Note that two distinct monomials x^α and $x^{\alpha'}$ in $\mathscr{P}_k$ are orthogonal w.r.t. it, since there exists at least one i such that $\alpha_i \geqslant \alpha_i'$, then $\partial_{x_i}^{\alpha_i} x_i^{\alpha_i'} = 0$. $\langle P, P \rangle = \sum |a_\alpha|^2 \alpha!$ where $\alpha! = (\alpha_1!) \cdots (\alpha_n!)$.

Definition 5.8. We define $\mathscr{H}_k$ to be the linear space of homogeneous polynomials of degree k which are harmonic: the *solid spherical harmonics of degree k*. That is,

$$\mathscr{H}_k := \{P(x) \in \mathscr{P}_k : \Delta P(x) = 0\}.$$

It will be convenient to restrict these polynomials to S^{n-1}, and there to define the standard inner product,

$$(P, Q) = \int_{S^{n-1}} P(x)\overline{Q(x)}d\sigma(x).$$

For a function f on S^{n-1}, we define the spherical LaplaceanΔ_S by

$$\Delta_S f(x) = \Delta f(x/|x|),$$

where $f(x/|x|)$ is the degree zero homogeneous extension of the function f to $\mathbb{R}^n \setminus \{0\}$, and Δ is the Laplacian of the Euclidean space.[1]

---

[1] This is implied by the well-known formula for the Euclidean Laplacian in spherical polar coordinates:

$$\Delta f = r^{1-n} \frac{\partial}{\partial r} \left(r^{n-1} \frac{\partial f}{\partial r} \right) + r^{-2} \Delta_S f.$$

Proposition 5.9. *We have the following properties.*

(1) *The finite dimensional spaces* $\{\mathscr{H}_k\}_{k=0}^{\infty}$ *are mutually orthogonal.*

(2) *Every homogeneous polynomial* $P \in \mathscr{P}_k$ *can be written in the form* $P = P_1 + |x|^2 P_2$, *where* $P_1 \in \mathscr{H}_k$ *and* $P_2 \in \mathscr{P}_{k-2}$.

(3) *Let* H_k *denote the linear space of restrictions of* $\mathscr{H}_k$ *to the unit sphere.*[2] *The elements of* H_k *are the surface spherical harmonics of degree* k, *i.e.,*

$$H_k = \{P(x) \in \mathscr{H}_k : |x| = 1\}.$$

Then $L^2(S^{n-1}) = \sum_{k=0}^{\infty} H_k$. *Here the* L^2 *space is taken w.r.t. usual measure, and the infinite direct sum is taken in the sense of Hilbert space theory. That is, if* $f \in L^2(S^{n-1})$, *then* f *has the development*

$$f(x) = \sum_{k=0}^{\infty} Y_k(x), \quad Y_k \in H_k, \tag{5.11}$$

where the convergence is in the $L^2(S^{n-1})$ *norm, and*

$$\int_{S^{n-1}} |f(x)|^2 d\sigma(x) = \sum_k \int_{S^{n-1}} |Y_k(x)|^2 d\sigma(x).$$

(4) *If* $Y_k(x) \in H_k$, *then* $\Delta_S Y_k(x) = -k(k+n-2)Y_k(x)$.

(5) *Suppose* f *has the development* (5.11). *Then* f *(after correction on a set of measure zero, if necessary) is indefinitely differentiable on* S^{n-1} *(i.e.,* $f \in C^\infty(S^{n-1})$*) if and only if*

$$\int_{S^{n-1}} |Y_k(x)|^2 d\sigma(x) = O(k^{-N}), \quad \text{as } k \to \infty, \text{ for each fixed } N. \tag{5.12}$$

Proof. (1) If $P \in \mathscr{P}_k$, i.e., $P(x) = \sum a_\alpha x^\alpha$ with $|\alpha| = k$, then

$$\sum_{j=1}^{n} x_j \partial_{x_j} P = \sum_{j=1}^{n} x_j \sum a_\alpha \alpha_j x_1^{\alpha_1} \cdots x_j^{\alpha_j - 1} \cdots x_n^{\alpha_n} = \sum_{j=1}^{n} \alpha_j \sum a_\alpha x^\alpha = kP.$$

On S^{n-1}, it follows $kP = \frac{\partial P}{\partial \nu}$ where $\frac{\partial}{\partial \nu}$ denotes differentiation w.r.t. the outward normal vector. Thus, for $P \in \mathscr{H}_k$, and $Q \in \mathscr{H}_j$, then by Green's theorem

$$\begin{aligned}(k-j)\int_{S^{n-1}} P\bar{Q} d\sigma(x) &= \int_{S^{n-1}} \left(\bar{Q}\frac{\partial P}{\partial \nu} - P\frac{\partial \bar{Q}}{\partial \nu}\right) d\sigma(x) \\ &= \int_{|x|\leqslant 1} [\bar{Q}\Delta P - P\Delta\bar{Q}] dx = 0,\end{aligned}$$

where Δ is the Laplacean on $\mathbb{R}^n$.

(2) Indeed, let $|x|^2\mathscr{P}_{k-2}$ be the subspace of $\mathscr{P}_k$ of all polynomials of the form $|x|^2 P_2$ where $P_2 \in \mathscr{P}_{k-2}$. Then its orthogonal complement w.r.t. $\langle \cdot, \cdot \rangle$ is exactly $\mathscr{H}_k$. In fact, P_1 is in this orthogonal complement if and only if $\langle |x|^2 P_2, P_1 \rangle = 0$ for all P_2. But $\langle |x|^2 P_2, P_1 \rangle = (P_2(\partial_x)\Delta)\overline{P_1} = \langle P_2, \Delta P_1 \rangle$, so $\Delta P_1 = 0$ and thus $\mathscr{P}_k = \mathscr{H}_k \oplus |x|^2 \mathscr{P}_{k-2}$, which proves the conclusion. In addition, we have for $P \in \mathscr{P}_k$

$$P(x) = P_k(x) + |x|^2 P_{k-2}(x) + \cdots + \begin{cases} |x|^k P_0(x), & k \text{ even}, \\ |x|^{k-1} P_1(x), & k \text{ odd}, \end{cases}$$

where $P_j \in \mathscr{H}_j$ by noticing that $\mathscr{P}_j = \mathscr{H}_j$ for $j = 0, 1$.

(3) In fact, by the further result in (2), if $|x| = 1$, then we have

$$P(x) = P_k(x) + P_{k-2}(x) + \cdots . + \begin{cases} P_0(x), & k \text{ even}, \\ P_1(x), & k \text{ odd}, \end{cases}$$

[2] Sometimes, in order to emphasize the distribution between $\mathscr{H}_k$ and H_k, the members of H_k are referred to as the *surface spherical harmonics*.

with $P_j \in \mathscr{H}_j$. That is, the restriction of any polynomial on the unit sphere is a finite linear combination of spherical harmonics. Since the restriction of polynomials is dense in $L^2(S^{n-1})$ in the norm (see [SW71, Corollary 2.3, p.141]) by the Weierstrass approximation theorem,[3] the conclusion is then established.

(4) In fact, for $|x|=1$, we have

$$\begin{aligned}\Delta_S Y_k(x) =&\Delta(|x|^{-k}Y_k(x)) = |x|^{-k}\Delta Y_k + \Delta(|x|^{-k})Y_k + 2\nabla(|x|^{-k})\cdot\nabla Y_k\\ =&(k^2+(2-n)k)|x|^{-k-2}Y_k - 2k^2|x|^{-k-2}Y_k\\ =&-k(k+n-2)|x|^{k-2}Y_k = -k(k+n-2)Y_k,\end{aligned}$$

since $\sum_{j=1}^n x_j\partial_{x_j}Y_k = kY_k$ for $Y_k \in \mathscr{P}_k$.

(5) To prove this, we write (5.11) as $f(x) = \sum_{k=0}^\infty a_k Y_k^0(x)$, where the Y_k^0 are normalized such that $\int_{S^{n-1}} |Y_k^0(x)|^2 d\sigma(x) = 1$. Our assertion is then equivalent with $a_k = O(k^{-N/2})$, as $k\to\infty$. If f is of class C^2, then an application of Green's theorem shows that

$$\int_{S^{n-1}} \Delta_S f\overline{Y_k^0}d\sigma = \int_{S^{n-1}} f\Delta_S\overline{Y_k^0}d\sigma.$$

Thus, if $f\in C^\infty$, then by (4)

$$\begin{aligned}\int_{S^{n-1}} \Delta_S^r f\overline{Y_k^0}d\sigma = \int_{S^{n-1}} f\Delta_S^r\overline{Y_k^0}d\sigma = [-k(k+n-2)]^r \int_{S^{n-1}}\sum_{j=0}^\infty a_j Y_j^0\overline{Y_k^0}d\sigma\\ =[-k(k+n-2)]^r a_k\int_{S^{n-1}}|Y_k^0|^2 d\sigma = a_k[-k(k+n-2)]^r.\end{aligned}$$

So $a_k = O(k^{-2r})$ for every r and therefore (5.12) holds.

To prove the converse, from (5.12), we have for any $r\in\mathbb{N}$

$$\begin{aligned}\|\Delta_S^r f\|_2^2 =&(\Delta_S^r f, \Delta_S^r f) = (\sum_{j=0}^\infty \Delta_S^r Y_j(x), \sum_{k=0}^\infty \Delta_S^r Y_k(x))\\ =&(\sum_{j=0}^\infty[-j(j+n-2)]^r Y_j(x), \sum_{k=0}^\infty[-k(k+n-2)]^r Y_k(x))\\ =&\sum_{k=0}^\infty[-k(k+n-2)]^{2r}(Y_k(x),Y_k(x))\\ =&\sum_{k=0}^\infty[-k(k+n-2)]^{2r}O(k^{-N})\leqslant C,\end{aligned}$$

if we take N large enough. Thus, $f\in C^\infty(S^{n-1})$. □

Theorem 5.10 (Hecke's identity)**.** *It holds*

$$\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}|x|^2}) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}(-i\,\mathrm{sgn}\,(\omega))^k P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2}, \quad \forall P_k\in\mathscr{H}_k(\mathbb{R}^n). \tag{5.13}$$

Proof. That is to prove

$$\int_{\mathbb{R}^n} P_k(x)e^{-\omega ix\cdot\xi-\frac{|\omega|}{2}|x|^2}dx = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}(-i\,\mathrm{sgn}\,(\omega))^k P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2}. \tag{5.14}$$

Applying the differential operator $P_k(\partial_\xi)$ to both sides of the identity (cf. Theorem 1.10)

$$\int_{\mathbb{R}^n} e^{-\omega ix\cdot\xi-\frac{|\omega|}{2}|x|^2}dx = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} e^{-\frac{|\omega|}{2}|\xi|^2},$$

[3] If g is continuous on S^{n-1}, we can approximate it uniformly by polynomials restricted to S^{n-1}.

we obtain

$$(-\omega i)^k \int_{\mathbb{R}^n} P_k(x) e^{-\omega i x\cdot\xi - \frac{|\omega|}{2}|x|^2} dx = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} Q(\xi) e^{-\frac{|\omega|}{2}|\xi|^2}.$$

Since $P_k(x)$ is polynomial, it is obvious analytic continuation $P_k(z)$ to all of $\mathbb{C}^n$. Thus, by a change of variable

$$\begin{aligned} Q(\xi) =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(x) e^{-\omega i x\cdot\xi - \frac{|\omega|}{2}|x|^2 + \frac{|\omega|}{2}|\xi|^2} dx \\ =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(x) e^{-\frac{|\omega|}{2}(x+i\,\mathrm{sgn}\,(\omega)\xi)^2} dx \\ =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(y - i\,\mathrm{sgn}\,(\omega)\xi) e^{-\frac{|\omega|}{2}|y|^2} dy. \end{aligned}$$

So,

$$\begin{aligned} Q(i\,\mathrm{sgn}\,(\omega)\xi) =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(y+\xi) e^{-\frac{|\omega|}{2}|y|^2} dy \\ =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_0^\infty r^{n-1} e^{-\frac{|\omega|}{2}r^2} \int_{S^{n-1}} P_k(\xi + r y') d\sigma(y') dr. \end{aligned}$$

Since P_k is harmonic, it satisfies the mean value property, i.e., Theorem 4.5, thus

$$\int_{S^{n-1}} P_k(\xi + r y') d\sigma(y') = \omega_{n-1} P_k(\xi) = P_k(\xi) \int_{S^{n-1}} d\sigma(y').$$

Hence

$$\begin{aligned} Q(i\,\mathrm{sgn}\,(\omega)\xi) =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} P_k(\xi) \int_0^\infty r^{n-1} e^{-\frac{|\omega|}{2}r^2} \int_{S^{n-1}} d\sigma(y') dr \\ =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} P_k(\xi) \int_{\mathbb{R}^n} e^{-\frac{|\omega|}{2}|x|^2} dx = (-\omega i)^k P_k(\xi). \end{aligned}$$

Thus, $Q(\xi) = (-\omega i)^k P_k(-i\,\mathrm{sgn}\,(\omega)\xi) = (-\omega i)^k (-i\,\mathrm{sgn}\,(\omega))^k P_k(\xi)$, which proves the theorem. □

The theorem implies the following generalization of itself, whose interest is that it links the various components of the decomposition of $L^2(\mathbb{R}^n)$, for different n.

If f is a radial function, we write $f = f(r)$, where $r = |x|$.

Corollary 5.11. *Let $P_k(x) \in \mathscr{H}_k(\mathbb{R}^n)$. Suppose that f is radial and $P_k(x)f(r) \in L^2(\mathbb{R}^n)$. Then the Fourier transform of $P_k(x)f(r)$ is also of the form $P_k(x)g(r)$, with g a radial function. Moreover, the induced transform $f \to g$, $T_{n,k}f = g$, depends essentially only on $n+2k$. More precisely, we have Bochner's relation*

$$T_{n,k} = \left(\frac{|\omega|}{2\pi}\right)^k (-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}. \tag{5.15}$$

Proof. Consider the Hilbert space of radial functions

$$\mathscr{R} = \left\{ f(r) : \|f\|^2 = \int_0^\infty |f(r)|^2 r^{2k+n-1} dr < \infty \right\},$$

with the indicated norm. Fix now $P_k(x)$, and assume that P_k is normalized, i.e.,

$$\int_{S^{n-1}} |P_k(x)|^2 d\sigma(x) = 1.$$

Our goal is to show that

$$(T_{n,k}f)(r) = \left(\frac{|\omega|}{2\pi}\right)^k (-i\,\mathrm{sgn}\,(\omega))^k (T_{n+2k,0}f)(r), \tag{5.16}$$

for each $f \in \mathscr{R}$.

First, if $f(r) = e^{-\frac{|\omega|}{2}r^2}$, then (5.16) is an immediate consequence of Theorem 5.10, i.e.,

$$\begin{aligned}(T_{n,k}e^{-\frac{|\omega|}{2}r^2})(R) =& \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k e^{-\frac{|\omega|}{2}R^2}\\ =& \left(\frac{|\omega|}{2\pi}\right)^{k} (-i\,\mathrm{sgn}\,(\omega))^k (T_{n+2k,0}e^{-\frac{|\omega|}{2}r^2})(R),\end{aligned}$$

which implies $T_{n,k}f = \left(\frac{|\omega|}{2\pi}\right)^{k} (-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}f$ for $f = e^{-\frac{|\omega|}{2}r^2}$.

Next, we consider $e^{-\frac{|\omega|}{2}\varepsilon r^2}$ for a fixed $\varepsilon > 0$. By the homogeneity of P_k and the interplay of dilations with the Fourier transform (cf. Proposition 1.3), i.e., $\mathscr{F}\delta_\varepsilon = \varepsilon^{-n}\delta_{\varepsilon^{-1}}\mathscr{F}$, and Hecke's identity, we get

$$\begin{aligned}&\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}) = \varepsilon^{-k/2}\mathscr{F}(P_k(\varepsilon^{1/2}x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2})\\ =&\varepsilon^{-k/2-n/2}\delta_{\varepsilon^{-1/2}}\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}|x|^2})\\ =&\varepsilon^{-k/2-n/2}\left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \delta_{\varepsilon^{-1/2}}(P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2})\\ =&\left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \varepsilon^{-k/2-n/2}P_k(\varepsilon^{-1/2}\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}\\ =&\left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \varepsilon^{-k-n/2}P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}.\end{aligned}$$

This shows that $T_{n,k}e^{-\frac{|\omega|}{2}\varepsilon r^2} = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \varepsilon^{-k-n/2}e^{-\frac{|\omega|}{2}r^2/\varepsilon}$, and so

$$\begin{aligned}T_{n+2k,0}e^{-\frac{|\omega|}{2}\varepsilon r^2} =& \left(\frac{|\omega|}{2\pi}\right)^{-k-n/2} (-i\,\mathrm{sgn}\,(\omega))^0 \varepsilon^{-0-(n+2k)/2}e^{-\frac{|\omega|}{2}r^2/\varepsilon}\\ =& \left(\frac{|\omega|}{2\pi}\right)^{-k-n/2} \varepsilon^{-k-n/2}e^{-\frac{|\omega|}{2}r^2/\varepsilon}.\end{aligned}$$

Thus, $T_{n,k}e^{-\frac{|\omega|}{2}\varepsilon r^2} = \left(\frac{|\omega|}{2\pi}\right)^{k} (-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}e^{-\frac{|\omega|}{2}\varepsilon r^2}$ for $\varepsilon > 0$.

To finish the proof, it suffices to see that the linear combination of $\{e^{-\frac{|\omega|}{2}\varepsilon r^2}\}_{0<\varepsilon<\infty}$ is dense in $\mathscr{R}$. Suppose the contrary, then there exists a (almost everywhere) non-zero $g \in \mathscr{R}$, such that g is orthogonal to every $e^{-\frac{|\omega|}{2}\varepsilon r^2}$ in the sense of $\mathscr{R}$, i.e.,

$$\int_0^\infty e^{-\frac{|\omega|}{2}\varepsilon r^2} g(r)r^{2k+n-1}dr = 0, \tag{5.17}$$

for all $\varepsilon > 0$. Let $\psi(s) = \int_0^s e^{-r^2}g(r)r^{n+2k-1}dr$ for $s \geqslant 0$. Then, putting $\varepsilon = 2(m+1)/|\omega|$, where m is a positive integer, and by integration by parts, we have

$$0 = \int_0^\infty e^{-mr^2}\psi'(r)dr = 2m\int_0^\infty e^{-mr^2}\psi(r)rdr.$$

By the change of variable $z = e^{-r^2}$, this equality is equivalent to

$$0 = \int_0^1 z^{m-1}\psi(\sqrt{\ln 1/z})dz, \quad m = 1, 2, ...$$

Since the polynomials are uniformly dense in the space of continuous functions on the closed interval $[0, 1]$, this can only be the case when $\psi(\sqrt{\ln 1/z}) = 0$ for all z in $[0, 1]$. Thus, $\psi'(r) = e^{-r^2}g(r)r^{n+2k-1} = 0$ for almost every $r \in (0, \infty)$, contradicting the hypothesis that $g(r)$ is not equal to 0 almost everywhere.

Since the operators $T_{n,k}$ and $\left(\frac{|\omega|}{2\pi}\right)^k(-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}$ are bounded and agree on the dense subspace, they must be equal. Thus, we have shown the desired result. □

We come now to what has been our main goal in our discussion of spherical harmonics.

Theorem 5.12. *Let $P_k(x) \in \mathscr{H}_k$, $k \geqslant 1$. Then the multiplier corresponding to the transform (5.10) with the kernel $\frac{P_k(x)}{|x|^{k+n}}$ is*

$$\gamma_k \frac{P_k(\xi)}{|\xi|^k}, \quad \text{with } \gamma_k = \pi^{n/2}(-i\,\mathrm{sgn}\,(\omega))^k \frac{\Gamma(k/2)}{\Gamma(k/2+n/2)}.$$

Remark 5.13. 1) If $k \geqslant 1$, then $P_k(x)$ is orthogonal to the constants on the sphere, and so its mean value over any sphere centered at the origin is zero.

2) The statement of the theorem can be interpreted as

$$\mathscr{F}\left(\frac{P_k(x)}{|x|^{k+n}}\right) = \gamma_k \frac{P_k(\xi)}{|\xi|^k}. \tag{5.18}$$

3) As such it will be derived from the following closely related fact,

$$\mathscr{F}\left(\frac{P_k(x)}{|x|^{k+n-\alpha}}\right) = \gamma_{k,\alpha} \frac{P_k(\xi)}{|\xi|^{k+\alpha}}, \tag{5.19}$$

where $\gamma_{k,\alpha} = \pi^{n/2}\left(\frac{|\omega|}{2}\right)^{-\alpha}(-i\,\mathrm{sgn}\,(\omega))^k \frac{\Gamma(k/2+\alpha/2)}{\Gamma(k/2+n/2-\alpha/2)}$.

Lemma 5.14. *The identity (5.19) holds in the sense that*

$$\int_{\mathbb{R}^n} \frac{P_k(x)}{|x|^{k+n-\alpha}} \hat{\varphi}(x)dx = \gamma_{k,\alpha} \int_{\mathbb{R}^n} \frac{P_k(\xi)}{|\xi|^{k+\alpha}} \varphi(\xi)d\xi, \quad \forall \varphi \in \mathscr{S}. \tag{5.20}$$

It is valid for all non-negative integer k and for $0 < \alpha < n$.

Remark 5.15. For the complex number α with $\Re\alpha \in (0, n)$, the lemma and (5.19) are also valid, see [SW71, Theorem 4.1, p.160-163].

Proof. From the proof of Corollary 5.11, we have already known that

$$\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}(-i\,\mathrm{sgn}\,(\omega))^k \varepsilon^{-k-n/2} P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon},$$

so we have by the multiplication formula,

$$\begin{aligned}\int_{\mathbb{R}^n} P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}\hat{\varphi}(x)dx &= \int_{\mathbb{R}^n} \mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2})(\xi)\varphi(\xi)d\xi \\ &= \left(\frac{|\omega|}{2\pi}\right)^{-n/2}(-i\,\mathrm{sgn}\,(\omega))^k \varepsilon^{-k-n/2} \int_{\mathbb{R}^n} P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}\varphi(\xi)d\xi,\end{aligned}$$

for $\varepsilon > 0$.

We now integrate both sides of the above w.r.t. ε, after having multiplied the equation by a suitable power of ε, ($\varepsilon^{\beta-1}$, $\beta = (k+n-\alpha)/2$, to be precise). That is

$$\begin{aligned}&\int_0^\infty \varepsilon^{\beta-1} \int_{\mathbb{R}^n} P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}\hat{\varphi}(x)dxd\varepsilon \\ =& \left(\frac{|\omega|}{2\pi}\right)^{-n/2}(-i\,\mathrm{sgn}\,(\omega))^k \int_0^\infty \varepsilon^{\beta-1}\varepsilon^{-k-n/2} \int_{\mathbb{R}^n} P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}\varphi(\xi)d\xi d\varepsilon.\end{aligned} \tag{5.21}$$

By changing the order of the double integral and a change of variable, we get

$$\begin{aligned}\text{l.h.s. of (5.21)} &= \int_{\mathbb{R}^n} P_k(x)\hat{\varphi}(x) \int_0^\infty \varepsilon^{\beta-1}e^{-\frac{|\omega|}{2}\varepsilon|x|^2}d\varepsilon dx \\ &\overset{t=|\omega|\varepsilon|x|^2/2}{=\!=\!=\!=} \int_{\mathbb{R}^n} P_k(x)\hat{\varphi}(x)\left(\frac{|\omega|}{2}|x|^2\right)^{-\beta} \int_0^\infty t^{\beta-1}e^{-t}dtdx\end{aligned}$$

$$= \left(\frac{|\omega|}{2}\right)^{-\beta} \Gamma(\beta) \int_{\mathbb{R}^n} P_k(x)\hat{\varphi}(x)|x|^{-2\beta} dx.$$

Similarly,

$$\begin{aligned}
\text{r.h.s. of (5.21)} &= \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi) \\
&\qquad \int_0^\infty \varepsilon^{-(k/2+\alpha/2+1)} e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon} d\varepsilon d\xi \\
&\overset{t=\frac{|\omega|}{2}|\xi|^2/\varepsilon}{=\!=\!=\!=} \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi) \left(\frac{|\omega|}{2}|\xi|^2\right)^{-(k+\alpha)/2} \\
&\qquad \int_0^\infty t^{k/2+\alpha/2-1} e^{-t} dt d\xi \\
&= \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \left(\frac{|\omega|}{2}\right)^{-(k+\alpha)/2} \Gamma(k/2+\alpha/2) \\
&\qquad \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi)|\xi|^{-(k+\alpha)} d\xi.
\end{aligned}$$

Thus, we get

$$\begin{aligned}
&\left(\frac{|\omega|}{2}\right)^{-(k+n-\alpha)/2} \Gamma((k+n-\alpha)/2) \int_{\mathbb{R}^n} P_k(x)\hat{\varphi}(x)|x|^{-(k+n-\alpha)} dx \\
=& \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (-i\,\mathrm{sgn}\,(\omega))^k \left(\frac{|\omega|}{2}\right)^{-(k+\alpha)/2} \Gamma(k/2+\alpha/2) \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi)|\xi|^{-(k+\alpha)} d\xi
\end{aligned}$$

which leads to (5.20).

Observe that when $0 < \alpha < n$ and $\varphi \in \mathscr{S}$, then double integrals in the above converge absolutely. Thus the formal argument just given establishes the lemma. □

Proof of Theorem 5.12. By the assumption that $k \geqslant 1$, we have that the integral of P_k over any sphere centered at the origin is zero. Thus for $\varphi \in \mathscr{S}$, we get

$$\int_{\mathbb{R}^n} \frac{P_k(x)}{|x|^{k+n-\alpha}} \hat{\varphi}(x) dx = \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n-\alpha}} [\hat{\varphi}(x) - \hat{\varphi}(0)] dx + \int_{|x|>1} \frac{P_k(x)}{|x|^{k+n-\alpha}} \hat{\varphi}(x) dx.$$

Obviously, the second term tends to $\int_{|x|>1} \frac{P_k(x)}{|x|^{k+n}} \hat{\varphi}(x) dx$ as $\alpha \to 0$ by the dominated convergence theorem. As in the proof of part (c) of Theorem 4.26, $\frac{P_k(x)}{|x|^{k+n}} [\hat{\varphi}(x) - \hat{\varphi}(0)]$ is locally integrable, thus we have, by the dominated convergence theorem, the limit of the first term in the r.h.s. of the above

$$\begin{aligned}
&\lim_{\alpha\to 0} \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n-\alpha}} [\hat{\varphi}(x) - \hat{\varphi}(0)] dx = \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n}} [\hat{\varphi}(x) - \hat{\varphi}(0)] dx \\
=& \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n}} \hat{\varphi}(x) dx = \lim_{\varepsilon\to 0} \int_{\varepsilon\leqslant|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n}} \hat{\varphi}(x) dx.
\end{aligned}$$

Thus, we obtain

$$\lim_{\alpha\to 0+} \int_{\mathbb{R}^n} \frac{P_k(x)}{|x|^{k+n-\alpha}} \hat{\varphi}(x) dx = \lim_{\varepsilon\to 0} \int_{|x|\geqslant\varepsilon} \frac{P_k(x)}{|x|^{k+n}} \hat{\varphi}(x) dx. \tag{5.22}$$

Similarly,

$$\lim_{\alpha\to 0+} \int_{\mathbb{R}^n} \frac{P_k(\xi)}{|\xi|^{k+\alpha}} \varphi(\xi) d\xi = \lim_{\varepsilon\to 0} \int_{|\xi|\geqslant\varepsilon} \frac{P_k(\xi)}{|\xi|^k} \varphi(\xi) d\xi.$$

Thus, by Lemma 5.11, we complete the proof with $\gamma_k = \lim_{\alpha\to 0} \gamma_{k,\alpha}$. □

For fixed $k \geqslant 1$, the linear space of operators in (5.10), where $\Omega(y) = \frac{P_k(y)}{|y|^k}$ and $P_k \in \mathscr{H}_k$, form a natural generalization of the Riesz transforms; the latter arise in the special case $k = 1$. Those for $k > 1$, we call the *higher Riesz transforms*, with k as the degree of the higher Riesz transforms, they can also be characterized by their invariance properties (see [Ste70, §4.8, p.79]).

5.3 Equivalence between two classes of transforms

We now consider two classes of transforms, defined on $L^2(\mathbb{R}^n)$. The first class consists of all transforms of the form

$$Tf = c \cdot f + \lim_{\varepsilon \to 0} \int_{|y| \geqslant \varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy, \tag{5.23}$$

where c is a constant, $\Omega \in C^\infty(S^{n-1})$ is a homogeneous function of degree 0, and the integral $\int_{S^{n-1}} \Omega(x)d\sigma(x) = 0$. The second class is given by those transforms T for which

$$\mathscr{F}(Tf)(\xi) = m(\xi)\hat{f}(\xi) \tag{5.24}$$

where the multiplier $m \in C^\infty(S^{n-1})$ is homogeneous of degree 0.

Theorem 5.16. *The two classes of transforms, defined by* (5.23) *and* (5.24) *respectively, are identical.*

Proof. First, support that T is of the form (5.23). Then by Theorem 4.24, T is of the form (5.24) with m homogeneous of degree 0 and

$$m(\xi) = c + \int_{S^{n-1}} \left[-\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(\xi \cdot x) + \ln(1/|\xi \cdot x|) \right] \Omega(x) d\sigma(x), \quad |\xi| = 1. \tag{5.25}$$

Now, we need to show $m \in C^\infty(S^{n-1})$. Write the spherical harmonic developments

$$\Omega(x) = \sum_{k=1}^\infty Y_k(x),\ m(x) = \sum_{k=0}^\infty \tilde{Y}_k(x),\ \Omega_N(x) = \sum_{k=1}^N Y_k(x),\ m_N(x) = \sum_{k=0}^N \tilde{Y}_k(x), \tag{5.26}$$

where $Y_k, \tilde{Y}_k \in H_k$ in view of part (3) in Proposition 5.9. k starts from 1 in the development of Ω, since $\int_{S^{n-1}} \Omega(x)dx = 0$ implies that $\Omega(x)$ is orthogonal to constants, and H_0 contains only constants.

Then, by Theorem 5.12, if $\Omega = \Omega_N$, then $m(x) = m_N(x)$, with

$$\tilde{Y}_k(x) = \gamma_k Y_k(x), \quad k \geqslant 1.$$

But $m_M(x) - m_N(x) = \int_{S^{n-1}} \left[-\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(y \cdot x) + \ln \frac{1}{|y \cdot x|} \right] [\Omega_M(y) - \Omega_N(y)] d\sigma(y)$. Moreover by Hölder's inequality,

$$\begin{aligned} &\sup_{x \in S^{n-1}} |m_M(x) - m_N(x)| \\ &\leqslant \left(\sup_x \int_{S^{n-1}} \left| -\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(y \cdot x) + \ln(1/|y \cdot x|) \right|^2 d\sigma(y) \right)^{1/2} \\ &\quad \times \left(\int_{S^{n-1}} |\Omega_M(y) - \Omega_N(y)|^2 d\sigma(y) \right)^{1/2} \to 0, \end{aligned} \tag{5.27}$$

as M, $N \to \infty$, since[4] for $n = 1$, $S^0 = \{-1, 1\}$,

$$\int_{S^0} \left| -\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(y \cdot x) + \ln(1/|y \cdot x|) \right|^2 d\sigma(y) = \frac{\pi^2}{2},$$

[4] There the argument is similar with some part of the proof of Theorem 4.24.

and for $n \geqslant 2$, we can pick a orthogonal matrix A satisfying $Ae_1 = x$ and $\det A = 1$ for $|x| = 1$, and then by a change of variable,

$$\begin{aligned}
&\sup_x \int_{S^{n-1}} \left| -\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(y \cdot x) + \ln(1/|y \cdot x|) \right|^2 d\sigma(y) \\
=&\sup_x \int_{S^{n-1}} \left[\frac{\pi^2}{4} + (\ln(1/|y \cdot x|))^2 \right] d\sigma(y) \\
=&\frac{\pi^2}{4}\omega_{n-1} + \sup_x \int_{S^{n-1}} (\ln |y \cdot Ae_1|)^2 d\sigma(y) \\
=&\frac{\pi^2}{4}\omega_{n-1} + \sup_x \int_{S^{n-1}} (\ln |A^{-1}y \cdot e_1|)^2 d\sigma(y) \\
\overset{z=A^{-1}y}{=\!=\!=\!=}&\frac{\pi^2}{4}\omega_{n-1} + \int_{S^{n-1}} (\ln |z_1|)^2 d\sigma(z) < \infty.
\end{aligned}$$

Here, we have used the boundedness of the integral in the r.h.s., i.e., (with the notation $\bar{z} = (z_2, ..., z_n)$, cf. [Gra04, p.A-20,p.267])

$$\begin{aligned}
\int_{S^{n-1}} (\ln |z_1|)^2 d\sigma(z) =& \int_{-1}^{1} (\ln |z_1|)^2 \int_{\sqrt{1-z_1^2}S^{n-2}} d\sigma(\bar{z}) \frac{dz_1}{\sqrt{1-z_1^2}} \\
\overset{y=\bar{z}/\sqrt{1-z_1^2}}{=\!=\!=\!=\!=\!=}& \int_{-1}^{1} (\ln |z_1|)^2 \int_{S^{n-2}} (1-z_1^2)^{(n-3)/2} d\sigma(y) dz_1 \\
=&\omega_{n-2} \int_{-1}^{1} (\ln |z_1|)^2 (1-z_1^2)^{(n-3)/2} dz_1 \\
\overset{z_1=\cos\theta}{=\!=\!=\!=}&\omega_{n-2} \int_0^{\pi} (\ln |\cos\theta|)^2 (\sin\theta)^{n-2} d\theta = \omega_{n-2} I_1.
\end{aligned}$$

If $n \geqslant 3$, then, by integration by parts,

$$I_1 \leqslant \int_0^{\pi} (\ln|\cos\theta|)^2 \sin\theta d\theta = -2\int_0^{\pi} \ln|\cos\theta| \sin\theta d\theta = 2\int_0^{\pi} \sin\theta d\theta = 4.$$

If $n = 2$, then, by the formula $\int_0^{\pi/2} (\ln(\cos\theta))^2 d\theta = \frac{\pi}{2}[(\ln 2)^2 + \pi^2/12]$, cf. [GR, 4.225.8, p.531], we get

$$I_1 = \int_0^{\pi} (\ln|\cos\theta|)^2 d\theta = 2\int_0^{\pi/2} (\ln(\cos\theta))^2 d\theta = \pi[(\ln 2)^2 + \pi^2/12].$$

Thus, (5.27) shows that

$$m(x) = c + \sum_{k=1}^{\infty} \gamma_k Y_k(x).$$

Since $\Omega \in C^\infty$, we have, in view of part (5) of Proposition 5.9, that

$$\int_{S^{n-1}} |Y_k(x)|^2 d\sigma(x) = O(k^{-N})$$

as $k \to \infty$ for every fixed N. However, by the explicit form of γ_k, we see that $\gamma_k \sim k^{-n/2}$, so $m(x)$ is also indefinitely differentiable on the unit sphere, i.e., $m \in C^\infty(S^{n-1})$.

Conversely, suppose $m(x) \in C^\infty(S^{n-1})$ and let its spherical harmonic development be as in (5.26). Set $c = \tilde{Y}_0$, and $Y_k(x) = \frac{1}{\gamma_k}\tilde{Y}_k(x)$. Then $\Omega(x)$, given by (5.26), has mean value zero in the sphere, and is again indefinitely differentiable there. But as we have just seen the multiplier corresponding to this transform is m; so the theorem is proved. □

As an application of this theorem and a final illustration of the singular integral transforms we shall give the generalization of the estimates for partial derivatives given in 5.1.

Let $P(x) \in \mathscr{P}_k(\mathbb{R}^n)$. We shall say that P is *elliptic* if $P(x)$ vanishes only at the origin. For any polynomial P, we consider also its corresponding differential polynomial. Thus, if $P(x) = \sum a_\alpha x^\alpha$, we write $P(\frac{\partial}{\partial x}) = \sum a_\alpha (\frac{\partial}{\partial x})^\alpha$ as in the previous definition.

Corollary 5.17. *Suppose P is a homogeneous elliptic polynomial of degree k. Let $(\frac{\partial}{\partial x})^\alpha$ be any differential monomial of degree k. Assume $f \in C_c^k$, then we have the a priori estimate*

$$\left\| \left(\frac{\partial}{\partial x}\right)^\alpha f \right\|_p \leqslant A_p \left\| P\left(\frac{\partial}{\partial x}\right) f \right\|_p, \quad 1 < p < \infty. \tag{5.28}$$

Proof. From the Fourier transform of $\left(\frac{\partial}{\partial x}\right)^\alpha f$ and $P\left(\frac{\partial}{\partial x}\right) f$,

$$\mathscr{F}\left(P\left(\frac{\partial}{\partial x}\right) f\right)(\xi) = \int_{\mathbb{R}^n} e^{-\omega i x \cdot \xi} P\left(\frac{\partial}{\partial x}\right) f(x) dx = (\omega i)^k P(\xi) \hat{f}(\xi),$$

and

$$\mathscr{F}\left(\left(\frac{\partial}{\partial x}\right)^\alpha f\right)(\xi) = (\omega i)^k \xi^\alpha \hat{f}(\xi),$$

we have the following relation

$$P(\xi)\mathscr{F}\left(\left(\frac{\partial}{\partial x}\right)^\alpha f\right)(\xi) = \xi^\alpha \mathscr{F}\left(P\left(\frac{\partial}{\partial x}\right) f\right)(\xi).$$

Since $P(\xi)$ is non-vanishing except at the origin, $\frac{\xi^\alpha}{P(\xi)}$ is homogenous of degree 0 and is indefinitely differentiable on the unit sphere. Thus

$$\left(\frac{\partial}{\partial x}\right)^\alpha f = T\left(P\left(\frac{\partial}{\partial x}\right) f\right),$$

where T is one of the transforms of the type given by (5.24). By Theorem 5.16, T is also given by (5.23) and hence by the result of Theorem 4.24, we get the estimate (5.28). □

Chapter 6
The Littlewood-Paley g-function and Multipliers

In harmonic analysis, Littlewood-Paley theory is a term used to describe a theoretical framework used to extend certain results about L^2 functions to L^p functions for $1 < p < \infty$. It is typically used as a substitute for orthogonality arguments which only apply to L^p functions when $p = 2$. One implementation involves studying a function by decomposing it in terms of functions with localized frequencies, and using the Littlewood-Paley g-function to compare it with its Poisson integral. The 1-variable case was originated by J. E. Littlewood and R. Paley (1931, 1937, 1938) and developed further by Zygmund and Marcinkiewicz in the 1930s using complex function theory (Zygmund 2002 [1935], chapters XIV, XV). E. M. Stein later extended the theory to higher dimensions using real variable techniques.

6.1 The Littlewood-Paley g-function

The g-function is a nonlinear operator which allows one to give a useful characterization of the L^p norm of a function on $\mathbb{R}^n$ in terms of the behavior of its Poisson integral. This characterization will be used not only in this chapter, but also in the succeeding chapter dealing with function spaces.

Let $f \in L^p(\mathbb{R}^n)$ and write $u(x,y)$ for its Poisson integral

$$u(x,y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i\xi\cdot x} e^{-|\omega\xi|y} \hat{f}(\xi)d\xi = \int_{\mathbb{R}^n} P_y(t)f(x-t)dt$$

as defined in (4.15) and (4.17). Let Δ denote the Laplace operator in $\mathbb{R}^{n+1}_+$, that is $\Delta = \frac{\partial^2}{\partial y^2} + \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$; ∇ is the corresponding gradient, $|\nabla u(x,y)|^2 = |\frac{\partial u}{\partial y}|^2 + |\nabla_x u(x,y)|^2$, where $|\nabla_x u(x,y)|^2 = \sum_{j=1}^n |\frac{\partial u}{\partial x_j}|^2$.

Definition 6.1. With the above notations, we define the *Littlewood-Paley g-function* $g(f)(x)$, by

$$g(f)(x) = \left(\int_0^\infty |\nabla u(x,y)|^2 y dy\right)^{1/2}. \tag{6.1}$$

We can also define two *partial g-functions*, one dealing with the y differentiation and the other with the x differentiations,

$$g_1(f)(x) = \left(\int_0^\infty \left|\frac{\partial u}{\partial y}(x,y)\right|^2 y dy\right)^{1/2}, \qquad g_x(f)(x) = \left(\int_0^\infty |\nabla_x u(x,y)|^2 y dy\right)^{1/2}. \tag{6.2}$$

Obviously, $g^2 = g_1^2 + g_x^2$.

The basic result for g is the following.

Theorem 6.2. *Suppose $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Then $g(f)(x) \in L^p(\mathbb{R}^n)$, and*

$$A_p' \|f\|_p \leqslant \|g(f)\|_p \leqslant A_p \|f\|_p. \tag{6.3}$$

Proof. *Step 1: We first consider the simple case $p = 2$.* For $f \in L^2(\mathbb{R}^n)$, we have

$$\|g(f)\|_2^2 = \int_{\mathbb{R}^n} \int_0^\infty |\nabla u(x,y)|^2 y dy dx = \int_0^\infty y \int_{\mathbb{R}^n} |\nabla u(x,y)|^2 dx dy.$$

In view of the identity

$$u(x,y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i \xi \cdot x} e^{-|\omega \xi| y} \hat{f}(\xi) d\xi,$$

we have

$$\frac{\partial u}{\partial y} = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} -|\omega \xi| \hat{f}(\xi) e^{\omega i \xi \cdot x} e^{-|\omega \xi| y} d\xi,$$

and

$$\frac{\partial u}{\partial x_j} = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} \omega i \xi_j \hat{f}(\xi) e^{\omega i \xi \cdot x} e^{-|\omega \xi| y} d\xi.$$

Thus, by Plancherel's formula,

$$\begin{aligned}
&\int_{\mathbb{R}^n} |\nabla u(x,y)|^2 dx = \int_{\mathbb{R}^n} \left[\left|\frac{\partial u}{\partial y}\right|^2 + \sum_{j=1}^n \left|\frac{\partial u}{\partial x_j}\right|^2 \right] dx = \left\|\frac{\partial u}{\partial y}\right\|_{L_x^2}^2 + \sum_{j=1}^n \left\|\frac{\partial u}{\partial x_j}\right\|_{L_x^2}^2 \\
=& \left[\|\mathscr{F}^{-1}(-|\omega \xi| \hat{f}(\xi) e^{-|\omega \xi| y})\|_2^2 + \sum_{j=1}^n \|\mathscr{F}^{-1}(\omega i \xi_j \hat{f}(\xi) e^{-|\omega \xi| y})\|_2^2 \right] \\
=& \left(\frac{|\omega|}{2\pi}\right)^n \left[\| -|\omega \xi| \hat{f}(\xi) e^{-|\omega \xi| y}\|_2^2 + \sum_{j=1}^n \|\omega i \xi_j \hat{f}(\xi) e^{-|\omega \xi| y}\|_2^2 \right] \\
=& 2 \left(\frac{|\omega|}{2\pi}\right)^n \omega^2 \||\xi| \hat{f}(\xi) e^{-|\omega \xi| y}\|_2^2 = \int_{\mathbb{R}^n} 2 \left(\frac{|\omega|}{2\pi}\right)^n \omega^2 |\xi|^2 |\hat{f}(\xi)|^2 e^{-2|\omega \xi| y} d\xi,
\end{aligned}$$

and so

$$\begin{aligned}
\|g(f)\|_2^2 =& \int_0^\infty y \int_{\mathbb{R}^n} 2 \left(\frac{|\omega|}{2\pi}\right)^n \omega^2 |\xi|^2 |\hat{f}(\xi)|^2 e^{-2|\omega \xi| y} d\xi dy \\
=& \int_{\mathbb{R}^n} 2 \left(\frac{|\omega|}{2\pi}\right)^n \omega^2 |\xi|^2 |\hat{f}(\xi)|^2 \int_0^\infty y e^{-2|\omega \xi| y} dy d\xi \\
=& \int_{\mathbb{R}^n} 2 \left(\frac{|\omega|}{2\pi}\right)^n \omega^2 |\xi|^2 |\hat{f}(\xi)|^2 \frac{1}{4\omega^2 |\xi|^2} d\xi = \frac{1}{2} \left(\frac{|\omega|}{2\pi}\right)^n \|\hat{f}\|_2^2 \\
=& \frac{1}{2} \|f\|_2^2.
\end{aligned}$$

Hence,

$$\|g(f)\|_2 = 2^{-1/2} \|f\|_2. \tag{6.4}$$

We have also obtained $\|g_1(f)\|_2 = \|g_x(f)\|_2 = \frac{1}{2}\|f\|_2$.

Step 2: We consider the case $p \neq 2$ and prove $\|g(f)\|_p \leqslant A_p \|f\|_p$. We define the Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ which are to be consider now. $\mathscr{H}_1$ is the one-dimensional Hilbert space of complex numbers. To define $\mathscr{H}_2$, we define first $\mathscr{H}_2^0$ as the L^2 space on $(0,\infty)$ with measure $y dy$, i.e.,

$$\mathscr{H}_2^0 = \left\{ f : |f|^2 = \int_0^\infty |f(y)|^2 y dy < \infty \right\}.$$

Let $\mathscr{H}_2$ be the direct sum of $n+1$ copies of $\mathscr{H}_2^0$; so the elements of $\mathscr{H}_2$ can be represented as $(n+1)$ component vectors whose entries belong to $\mathscr{H}_2^0$. Since $\mathscr{H}_1$ is the same as the complex numbers, then $L(\mathscr{H}_1, \mathscr{H}_2)$ is of course identifiable with $\mathscr{H}_2$. Now let $\varepsilon > 0$, and keep it temporarily fixed.

Define

$$K_\varepsilon(x) = \left(\frac{\partial P_{y+\varepsilon}(x)}{\partial y}, \frac{\partial P_{y+\varepsilon}(x)}{\partial x_1}, \cdots, \frac{\partial P_{y+\varepsilon}(x)}{\partial x_n} \right).$$

Notice that for each fixed x, $K_\varepsilon(x) \in \mathscr{H}_2$. This is the same as saying that

$$\int_0^\infty \left| \frac{\partial P_{y+\varepsilon}(x)}{\partial y} \right|^2 y dy < \infty \text{ and } \int_0^\infty \left| \frac{\partial P_{y+\varepsilon}(x)}{\partial x_j} \right|^2 y dy < \infty, \text{ for } j = 1, ..., n.$$

In fact, since $P_y(x) = \frac{c_n y}{(|x|^2+y^2)^{(n+1)/2}}$, we have that both $\frac{\partial P_y}{\partial y}$ and $\frac{\partial P_y}{\partial x_j}$ are bounded by $\frac{A}{(|x|^2+y^2)^{(n+1)/2}}$. So the norm in $\mathscr{H}_2$ of $K_\varepsilon(x)$,

$$|K_\varepsilon(x)|^2 \leqslant A^2(n+1) \int_0^\infty \frac{y dy}{(|x|^2 + (y+\varepsilon)^2)^{n+1}} \leqslant A^2(n+1) \int_0^\infty \frac{dy}{(y+\varepsilon)^{2n+1}} \leqslant C_\varepsilon,$$

and in another way

$$|K_\varepsilon(x)|^2 \leqslant A^2(n+1) \int_\varepsilon^\infty \frac{y dy}{(|x|^2 + y^2)^{n+1}} = \frac{A^2(n+1)}{2n}(|x|^2 + \varepsilon^2)^{-n} \leqslant C|x|^{-2n}.$$

Thus,

$$|K_\varepsilon(x)| \in L^1_{\text{loc}}(\mathbb{R}^n). \tag{6.5}$$

Similarly,

$$\left| \frac{\partial K_\varepsilon(x)}{\partial x_j} \right|^2 \leqslant C \int_0^\infty \frac{y dy}{(|x|^2 + y^2)^{n+2}} \leqslant C \int_\varepsilon^\infty \frac{y dy}{(|x|^2 + y^2)^{n+2}} \leqslant C|x|^{-2n-2}.$$

Therefore, K_ε satisfies the gradient condition, i.e.,

$$\left| \frac{\partial K_\varepsilon(x)}{\partial x_j} \right| \leqslant C|x|^{-(n+1)}, \tag{6.6}$$

with C independent of ε.

Now we consider the operator T_ε defined by

$$T_\varepsilon f(x) = \int_{\mathbb{R}^n} K_\varepsilon(t) f(x-t) dt.$$

The function f is complex-valued (take their value in $\mathscr{H}_1$), but $T_\varepsilon f(x)$ takes its value in $\mathscr{H}_2$. Observe that

$$|T_\varepsilon f(x)| = \left(\int_0^\infty |\nabla u(x, y+\varepsilon)|^2 y dy \right)^{\frac{1}{2}} \leqslant \left(\int_\varepsilon^\infty |\nabla u(x,y)|^2 y dy \right)^{\frac{1}{2}} \leqslant g(f)(x). \tag{6.7}$$

Hence, $\|T_\varepsilon f(x)\|_2 \leqslant 2^{-1/2} \|f\|_2$, if $f \in L^2(\mathbb{R}^n)$, by (6.4). Therefore,

$$|\hat{K}_\varepsilon(x)| \leqslant 2^{-1/2}. \tag{6.8}$$

Because of (6.5), (6.6) and (6.8), by Theorem 4.27 (cf. Theorem 4.18), we get $\|T_\varepsilon f\|_p \leqslant A_p \|f\|_p$, $1 < p < \infty$ with A_p independent of ε. By (6.7), for each x, $|T_\varepsilon f(x)|$ increases to $g(f)(x)$, as $\varepsilon \to 0$, so we obtain finally

$$\|g(f)\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty. \tag{6.9}$$

Step 3: To derive the converse inequalities,

$$A'_p \|f\|_p \leqslant \|g(f)\|_p, \quad 1 < p < \infty. \tag{6.10}$$

In the first step, we have shown that $\|g_1(f)\|_2 = \frac{1}{2}\|f\|_2$ for $f \in L^2(\mathbb{R}^n)$. Let u_1, u_2 are the Poisson integrals of $f_1, f_2 \in L^2$, respectively. Then we have $\|g_1(f_1+f_2)\|_2^2 = \frac{1}{4}\|f_1+f_2\|_2^2$, i.e., $\int_{\mathbb{R}^n}\int_0^\infty |\frac{\partial(u_1+u_2)}{\partial y}|^2 y dy dx = \frac{1}{4}\int_{\mathbb{R}^n} |f_1+f_2|^2 dx$. It leads to the identity

$$4\int_{\mathbb{R}^n}\int_0^\infty \frac{\partial u_1}{\partial y}(x,y)\overline{\frac{\partial u_2}{\partial y}}(x,y) y dy dx = \int_{\mathbb{R}^n} f_1(x)\overline{f_2(x)}dx.$$

This identity, in turn, leads to the inequality, by Hölder's inequality and the definition of g_1,

$$\frac{1}{4}\left|\int_{\mathbb{R}^n} f_1(x)\overline{f_2(x)}dx\right| \leqslant \int_{\mathbb{R}^n} g_1(f_1)(x)g_1(f_2)(x)dx.$$

Suppose now in addition that $f_1 \in L^p(\mathbb{R}^n)$ and $f_2 \in L^{p'}(\mathbb{R}^n)$ with $\|f_2\|_{p'} \leqslant 1$ and $1/p + 1/p' = 1$. Then by Hölder inequality and the result (6.9).

$$\left|\int_{\mathbb{R}^n} f_1(x)\overline{f_2(x)}dx\right| \leqslant 4\|g_1(f_1)\|_p\|g_1(f_2)\|_{p'} \leqslant 4A_{p'}\|g_1(f_1)\|_p. \tag{6.11}$$

Now we take the supremum in (6.11) as f_2 ranges over all function in $L^2 \cap L^{p'}$, with $\|f_2\|_{p'} \leqslant 1$. Then, we obtain the desired result (6.10), with $A'_p = 1/4A_{p'}$, but where f is restricted to be in $L^2 \cap L^p$. The passage to the general case is provided by an easy limiting argument. Let f_m be a sequence of functions in $L^2 \cap L^p$, which converges in L^p norm to f. Notice that $|g(f_m)(x) - g(f_n)(x)| = \left|\|\nabla u_m\|_{L^2(0,\infty;ydy)} - \|\nabla u_n\|_{L^2(0,\infty;ydy)}\right| \leqslant \|\nabla u_m - \nabla u_n\|_{L^2(0,\infty;ydy)} = g(f_m - f_n)(x)$ by the triangle inequality. Thus, $\{g(f_m)\}$ is a Cauchy sequence in L^p and so converges to $g(f)$ in L^p, and we obtain the inequality (6.10) for f as a result of the corresponding inequalities for f_m. □

We have incidentally also proved the following, which we state as a corollary.

Corollary 6.3. *Suppose $f \in L^2(\mathbb{R}^n)$, and $g_1(f) \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Then $f \in L^p(\mathbb{R}^n)$, and $A'_p\|f\|_p \leqslant \|g_1(f)\|_p$.*

Remark 6.4. There are some very simple variants of the above that should be pointed out:

(i) The results hold also with $g_x(f)$ instead of $g(f)$. The direct inequality $\|g_x(f)\|_p \leqslant A_p\|f\|_p$ is of course a consequence of the one for g. The converse inequality is then proved in the same way as that for g_1.

(ii) For any integer $k > 1$, define

$$g_k(f)(x) = \left(\int_0^\infty \left|\frac{\partial^k u}{\partial y^k}(x,y)\right|^2 y^{2k-1}dy\right)^{1/2}.$$

Then the L^p inequalities hold for g_k as well. both (i) and (ii) are stated more systematically in [Ste70, Chapter IV, §7.2, p.112-113].

(iii) For later purpose, it will be useful to note that for each x, $g_k(f)(x) \geqslant A_k g_1(f)(x)$ where the bound A_k depends only on k.

It is easily verified from the Poisson integral formula that if $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, then

$$\frac{\partial^k u(x,y)}{\partial y^k} \to 0 \text{ for each } x, \quad \text{as } y \to \infty.$$

Thus,

$$\frac{\partial^k u(x,y)}{\partial y^k} = -\int_y^\infty \frac{\partial^{k+1}u(x,s)}{\partial s^{k+1}} s^k \frac{ds}{s^k}.$$

By Schwarz's inequality, therefore,

$$\left|\frac{\partial^k u(x,y)}{\partial y^k}\right|^2 \leqslant \left(\int_y^\infty \left|\frac{\partial^{k+1}u(x,s)}{\partial s^{k+1}}\right|^2 s^{2k}ds\right)\left(\int_y^\infty s^{-2k}ds\right).$$

Hence, by Hardy's inequality (2.18) (on p.41, with $q = r = 1$ there), we have

$$
\begin{aligned}
&(g_k(f)(x))^2 = \int_0^\infty \left|\frac{\partial^k u}{\partial y^k}(x,y)\right|^2 y^{2k-1}dy\\
\leqslant&\int_0^\infty \left(\int_y^\infty \left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2k}ds\right)\left(\int_y^\infty s^{-2k}ds\right) y^{2k-1}dy\\
=&\frac{1}{2k-1}\int_0^\infty \left(\int_y^\infty \left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2k}ds\right) dy \leqslant \frac{1}{2k-1}\int_0^\infty \left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2k+1}ds\\
=&\frac{1}{2k-1}\int_0^\infty \left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2(k+1)-1}ds\\
=&\frac{1}{2k-1}(g_{k+1}(f)(x))^2.
\end{aligned}
$$

Thus, the assertion is proved by the induction on k.

The proof that was given for the L^p inequalities for the g-function did not, in any essential way, depend on the theory of harmonic functions, despite the fact that this function was defined in terms of the Poisson integral. In effect, all that was really used is the fact that the Poisson kernels are suitable approximations to the identity.

There is, however, another approach, which can be carried out without recourse to the theory of singular integrals, but which leans heavily on characteristic properties of harmonic functions. We present it here (more precisely, we present that part which deals with $1 < p \leqslant 2$, for the inequality (6.9)), because its ideas can be adapted to other situations where the methods of Chapter 4 are not applicable. Everything will be based on the following three observations.

Lemma 6.5. *Suppose u is harmonic and strictly positive. Then*

$$\Delta u^p = p(p-1)u^{p-2}|\nabla u|^2. \tag{6.12}$$

Proof. The proof is straightforward. Indeed,

$$\partial_{x_j}u^p = pu^{p-1}\partial_{x_j}u, \quad \partial^2_{x_j}u^p = p(p-1)u^{p-2}(\partial_{x_j}u)^2 + pu^{p-1}\partial^2_{x_j}u,$$

which implies by summation

$$\Delta u^p = p(p-1)u^{p-2}|\nabla u|^2 + pu^{p-1}\Delta u = p(p-1)u^{p-2}|\nabla u|^2,$$

since $\Delta u = 0$. □

Lemma 6.6. *Suppose $F(x,y) \in C(\overline{\mathbb{R}^{n+1}_+}) \cap C^2(\mathbb{R}^{n+1}_+)$, and suitably small at infinity. Then*

$$\int_{\mathbb{R}^{n+1}_+} y\Delta F(x,y)dxdy = \int_{\mathbb{R}^n} F(x,0)dx. \tag{6.13}$$

Proof. We use Green's theorem

$$\int_D (u\Delta v - v\Delta u)dxdy = \int_{\partial D}\left(u\frac{\partial v}{\partial \mathcal{N}} - v\frac{\partial u}{\partial \mathcal{N}}\right)d\sigma$$

where $D = B_r \cap \mathbb{R}^{n+1}_+$, with B_r the ball of radius r in $\mathbb{R}^{n+1}$ centered at the origin, $\mathcal{N}$ is the outward normal vector. We take $v = F$, and $u = y$. Then, we will obtain our result (6.13) if

$$\int_D y\Delta F(x,y)dxdy \to \int_{\mathbb{R}^{n+1}_+} y\Delta F(x,y)dxdy, \text{ and } \int_{\partial D_0}\left(y\frac{\partial F}{\partial \mathcal{N}} - F\frac{\partial y}{\partial \mathcal{N}}\right)d\sigma \to 0,$$

as $r \to \infty$. Here ∂D_0 is the spherical part of the boundary of D. This will certainly be the case, if for example $\Delta F \geqslant 0$, and $|F| \leqslant O((|x|+y)^{-n-\varepsilon})$ and $|\nabla F| = O((|x|+y)^{-n-1-\varepsilon})$, as $|x| + y \to \infty$, for some $\varepsilon > 0$. □

Lemma 6.7. *If $u(x,y)$ is the Poisson integral of f, then*

$$\sup_{y>0}|u(x,y)| \leqslant Mf(x). \tag{6.14}$$

Proof. This is the same as the part (a) of Theorem 4.9. It can be proved with a similar argument as in the proof of part (a) for Theorem 4.10. □

Now we use these lemmas to give another proof for the inequality $\|g(f)\|_p \leqslant A_p\|f\|_p$, $1 < p \leqslant 2$.

Another proof of $\|g(f)\|_p \leqslant A_p\|f\|_p$, $1 < p \leqslant 2$. Suppose first $0 \leqslant f \in \mathscr{D}(\mathbb{R}^n)$ (and at least $f \neq 0$ on a nonzero measurable set). Then the Poisson integral u of f, $u(x,y) = \int_{\mathbb{R}^n} P_y(t)f(x-t)dt > 0$, since $P_y > 0$ for any $x \in \mathbb{R}^n$ and $y > 0$; and the majorizations $u(x,y) = O((|x|+y)^{-n})$ and $|\nabla u| = O((|x|+y)^{-n-1})$, as $|x|+y \to \infty$ are valid. We have, by Lemma 6.5, Lemma 6.7 and the hypothesis $1 < p \leqslant 2$,

$$\begin{aligned}(g(f)(x))^2 = &\int_0^\infty y|\nabla u(x,y)|^2 dy = \frac{1}{p(p-1)}\int_0^\infty yu^{2-p}\Delta u^p dy\\ \leqslant &\frac{[Mf(x)]^{2-p}}{p(p-1)}\int_0^\infty y\Delta u^p dy.\end{aligned}$$

We can write this as

$$g(f)(x) \leqslant C_p(Mf(x))^{(2-p)/2}(I(x))^{1/2}, \tag{6.15}$$

where $I(x) = \int_0^\infty y\Delta u^p dy$. However, by Lemma 6.6,

$$\int_{\mathbb{R}^n} I(x)dx = \int_{\mathbb{R}^{n+1}_+} y\Delta u^p dydx = \int_{\mathbb{R}^n} u^p(x,0)dx = \|f\|_p^p. \tag{6.16}$$

This immediately gives the desired result for $p = 2$.

Next, suppose $1 < p < 2$. By (6.15), Hölder's inequality, Theorem 3.9 and (6.16), we have, for $0 \leqslant f \in \mathscr{D}(\mathbb{R}^n)$,

$$\begin{aligned}&\int_{\mathbb{R}^n}(g(f)(x))^p dx \leqslant C_p^p\int_{\mathbb{R}^n}(Mf(x))^{p(2-p)/2}(I(x))^{p/2}dx\\ \leqslant &C_p^p\left(\int_{\mathbb{R}^n}(Mf(x))^p dx\right)^{1/r'}\left(\int_{\mathbb{R}^n} I(x)dx\right)^{1/r} \leqslant C_p'\|f\|_p^{p/r'}\|f\|_p^{p/r} = C_p'\|f\|_p^p,\end{aligned}$$

where $r = 2/p \in (1,2)$ and $1/r + 1/r' = 1$, then $r' = 2/(2-p)$.

Thus, $\|g(f)\|_p \leqslant A_p\|f\|_p$, $1 < p \leqslant 2$, whenever $0 \leqslant f \in \mathscr{D}(\mathbb{R}^n)$.

For general $f \in L^p(\mathbb{R}^n)$ (which we assume for simplicity to be real-valued), write $f = f^+ - f^-$ as its decomposition into positive and negative part; then we need only approximate in norm f^+ and f^-, each by a sequences of positive functions in $\mathscr{D}(\mathbb{R}^n)$. We omit the routine details that are needed to complete the proof. □

Unfortunately, the elegant argument just given is not valid for $p > 2$. There is, however, a more intricate variant of the same idea which does work for the case $p > 2$, but we do not intend to reproduce it here.

We shall, however, use the ideas above to obtain a significant generalization of the inequality for the g-functions.

Definition 6.8. Define the positive function

$$(g_\lambda^*(f)(x))^2 = \int_0^\infty\int_{\mathbb{R}^n}\left(\frac{y}{|t|+y}\right)^{\lambda n}|\nabla u(x-t,y)|^2 y^{1-n}dtdy. \tag{6.17}$$

Before going any further, we shall make a few comments that will help to clarify the meaning of the complicated expression (6.17).

First, $g_\lambda^*(f)(x)$ will turn out to be a pointwise majorant of $g(f)(x)$. To understand this situation better we have to introduce still another quantity, which is roughly midway between g and g_λ^*. It is defined as follows.

Definition 6.9. Let $\varGamma$ be a fixed proper cone in $\mathbb{R}_+^{n+1}$ with vertex at the origin and which contains $(0,1)$ in its interior. The exact form of $\varGamma$ will not really matter, but for the sake of definiteness let us choose for $\varGamma$ the up circular cone:

$$\varGamma = \left\{(t,y) \in \mathbb{R}_+^{n+1} : |t| < y, y > 0\right\}.$$

For any $x \in \mathbb{R}^n$, let $\varGamma(x)$ be the cone $\varGamma$ translated such that its vertex is at x. Now define the positive *Luzin's S-function* $S(f)(x)$ by

$$[S(f)(x)]^2 = \int_{\varGamma(x)} |\nabla u(t,y)|^2 y^{1-n} dydt = \int_{\varGamma} |\nabla u(x-t,y)|^2 y^{1-n} dydt. \tag{6.18}$$

We assert, as we shall momentarily prove, that

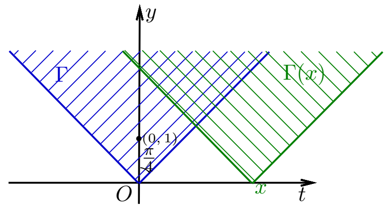


Fig. 6.1 $\varGamma$ and $\varGamma(x)$ for $n=1$

Proposition 6.10.

$$g(f)(x) \leqslant CS(f)(x) \leqslant C_\lambda g_\lambda^*(f)(x). \tag{6.19}$$

What interpretation can we put on the inequalities relating these three quantities? A hint is afforded by considering three corresponding approaches to the boundary for harmonic functions.

(a) With $u(x,y)$ the Poisson integral of $f(x)$, the simplest approach to the boundary point $x \in \mathbb{R}^n$ is obtained by letting $y \to 0$, (with x fixed). This is the perpendicular approach, and for it the appropriate limit exists almost everywhere, as we already know.

(b) Wider scope is obtained by allowing the variable point (t,y) to approach $(x,0)$ through any cone $\varGamma(x)$, (where vertex is x). This is the non-tangential approach which will be so important for us later. As the reader may have already realized, the relation of the S-function to the g-function is in some sense analogous to the relation between the non-tangential and the perpendicular approaches; we should add that the S-function is of decisive significance in its own right, but we shall not pursue that matter now.

(c) Finally, the widest scope is obtained by allowing the variable point (t,y) to approach $(x,0)$ in an arbitrary manner, i.e., the unrestricted approach. The function g_λ^* has the analogous role: it takes into account the unrestricted approach for Poisson integrals.

Notice that $g_\lambda^*(x)$ depends on λ. For each x, the smaller λ the greater $g_\lambda^*(x)$, and this behavior is such that that L^p boundedness of g_λ^* depends critically on the correct relation between p and λ. This last point is probably the main interest in g_λ^*, and is what makes its study more difficult than g or S.

After these various heuristic and imprecise indications, let us return to firm ground. The only thing for us to prove here is the assertion (6.19).

Proof of Proposition 6.10. The inequality $S(f)(x) \leqslant C_\lambda g_\lambda^*(f)(x)$ is obvious, since the integral (6.17) majorizes that part of the integral taken only over $\varGamma$, and

$$\left(\frac{y}{|t|+y}\right)^{\lambda n} \geqslant \frac{1}{2^{\lambda n}}$$

since $|t| < y$ there. The non-trivial part of the assertion is:

$$g(f)(x) \leqslant CS(f)(x).$$

It suffices to prove this inequality for $x = 0$. Let us denote by B_y the ball in $\mathbb{R}^{n+1}_+$ centered at $(0, y)$ and tangent to the boundary of the cone Γ; the radius of B_y is then proportional to y. Now the partial derivatives $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial x_k}$ are, like u, harmonic functions. Thus, by the mean value theorem of harmonic functions (i.e., Theorem 4.5 by noticing $(0, y)$ is the center of B_y),

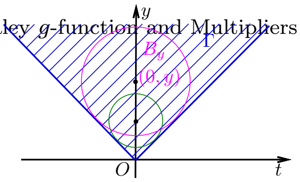


Fig. 6.2 Γ and B_y

$$\frac{\partial u(0,y)}{\partial y} = \frac{1}{m(B_y)} \int_{B_y} \frac{\partial u(x,s)}{\partial s} dxds$$

where $m(B_y)$ is the $n+1$ dimensional measure of B_y, i.e., $m(B_y) = cy^{n+1}$ for an appropriate constant c. By Schwarz's inequality

$$\left|\frac{\partial u(0,y)}{\partial y}\right|^2 \leqslant \frac{1}{(m(B_y))^2} \int_{B_y} \left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds \int_{B_y} dxds = \frac{1}{m(B_y)} \int_{B_y} \left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds.$$

If we integrate this inequality, we obtain

$$\int_0^\infty y \left|\frac{\partial u(0,y)}{\partial y}\right|^2 dy \leqslant \int_0^\infty c^{-1} y^{-n} \left(\int_{B_y} \left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds\right) dy.$$

However, $(x, s) \in B_y$ clearly implies that $c_1 s \leqslant y \leqslant c_2 s$, for two positive constants c_1 and c_2. Thus, apart from a multiplicative factor by changing the order of the double integrals, the last integral is majorized by

$$\int_\Gamma \left(\int_{c_1 s}^{c_2 s} y^{-n} dy\right) \left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds \leqslant c' \int_\Gamma \left|\frac{\partial u(x,s)}{\partial s}\right|^2 s^{1-n} dxds.$$

This is another way of saying that,

$$\int_0^\infty y \left|\frac{\partial u(0,y)}{\partial y}\right|^2 dy \leqslant c'' \int_\Gamma \left|\frac{\partial u(x,y)}{\partial y}\right|^2 y^{1-n} dxdy.$$

The same is true for the derivatives $\frac{\partial u}{\partial x_j}$, $j = 1, ..., n$, and adding the corresponding estimates proves our assertion. □

We are now in a position to state the L^p estimates concerning g^*_λ.

Theorem 6.11. *Let $\lambda > 1$ be a parameter. Suppose $f \in L^p(\mathbb{R}^n)$. Then*

(a) *For every $x \in \mathbb{R}^n$, $g(f)(x) \leqslant C_\lambda g^*_\lambda(f)(x)$.*

(b) *If $1 < p < \infty$, and $p > 2/\lambda$, then*

$$\|g^*_\lambda(f)\|_p \leqslant A_{p,\lambda} \|f\|_p. \tag{6.20}$$

Proof. The part (a) has already been proved in Proposition 6.10. Now, we prove (b).

For the case $p \geqslant 2$, only the assumption $\lambda > 1$ is relevant since $2/\lambda < 2 \leqslant p$.

Let ψ denote a positive function on $\mathbb{R}^n$, we claim that

$$\int_{\mathbb{R}^n} (g^*_\lambda(f)(x))^2 \psi(x) dx \leqslant A_\lambda \int_{\mathbb{R}^n} (g(f)(x))^2 (M\psi)(x) dx. \tag{6.21}$$

The l.h.s. of (6.21) equals

$$\int_0^\infty \int_{t \in \mathbb{R}^n} y |\nabla u(t,y)|^2 \left[\int_{x \in \mathbb{R}^n} \frac{\psi(x)}{(|t-x|+y)^{\lambda n}} y^{\lambda n} y^{-n} dx\right] dtdy,$$

so to prove (6.21), we must show that

$$\sup_{y>0} \int_{x \in \mathbb{R}^n} \frac{\psi(x)}{(|t-x|+y)^{\lambda n}} y^{\lambda n} y^{-n} dx \leqslant A_\lambda M\psi(t). \tag{6.22}$$

However, we know by Theorem 4.10, that

$$\sup_{\varepsilon>0} (\psi * \varphi_\varepsilon)(t) \leqslant AM\psi(t)$$

for appropriate φ, with $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$. Here, we have in fact $\varphi(x) = (1+|x|)^{-\lambda n}$, $\varepsilon = y$, and so with $\lambda > 1$ the hypotheses of that theorem are satisfied. This proves (6.22) and thus also (6.21).

The case $p = 2$ follows immediately from (6.21) by inserting in this inequality the function $\psi = 1$ (or by the definitions of $g^*_\lambda(f)$ and $g(f)$ directly), and using the L^2 result for g.

Suppose now $p > 2$; let us set $1/q + 2/p = 1$, and take the supremum of the l.h.s. of (6.21) over all $\psi \geqslant 0$, such that $\psi \in L^q(\mathbb{R}^n)$ and $\|\psi\|_q \leqslant 1$. Then, it gives $\|g^*_\lambda(f)\|_p^2$; Hölder's inequality yields an estimate for the right side:

$$A_\lambda \|g(f)\|_p^2 \|M\psi\|_q.$$

However, by the inequalities for the g-function, $\|g(f)\|_p \leqslant A'_p\|f\|_p$; and by the theorem of the maximal function $\|M\psi\|_q \leqslant A_q\|\psi\|_q \leqslant A''_q$, since $q > 1$, if $p < \infty$. If we substitute these in the above, we get the result:

$$\|g^*_\lambda(f)\|_p \leqslant A_{p,\lambda}\|f\|_p, \quad 2 \leqslant p < \infty, \quad \lambda > 1.$$

The inequalities for $p < 2$ will be proved by an adaptation of the reasoning used for g. Lemmas 6.5 and 6.6 will be equally applicable in the present situation, but we need more general version of Lemma 6.7, in order to majorize the unrestricted approach to the boundary of a Poisson integral.

It is at this stage where results which depend critically on the L^p class first make their appearance. Matters will depend on a variant of the maximal function which we define as follows. Let $\mu \geqslant 1$, and write $M_\mu f(x)$ for

$$M_\mu f(x) = \left(\sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)|^\mu dy\right)^{1/\mu}. \tag{6.23}$$

Then $M_1 f(x) = Mf(x)$, and $M_\mu f(x) = ((M|f|^\mu)(x))^{1/\mu}$. From the theorem of the maximal function, it immediately follows that, for $p > \mu$,

$$\|M_\mu f\|_p = \|((M|f|^\mu)(x))^{1/\mu}\|_p = \|((M|f|^\mu)(x))\|_{p/\mu}^{1/\mu} \leqslant \||f|^\mu\|_{p/\mu}^{1/\mu} = \|f\|_p. \tag{6.24}$$

This inequality fails for $p \leqslant \mu$, as in the special case $\mu = 1$.

The substitute for Lemma 6.7 is as follows.

Lemma 6.12. *Let $f \in L^p(\mathbb{R}^n)$, $p \geqslant \mu \geqslant 1$; if $u(x,y)$ is the Poisson integral of f, then*

$$|u(x-t,y)| \leqslant A\left(1 + \frac{|t|}{y}\right)^n Mf(x), \tag{6.25}$$

and more generally

$$|u(x-t,y)| \leqslant A_\mu\left(1 + \frac{|t|}{y}\right)^{n/\mu} M_\mu f(x). \tag{6.26}$$

We shall now complete the proof of the inequality (6.20) for the case $1 < p < 2$, with the restriction $p > 2/\lambda$.

Let us observe that we can always find a $\mu \in [1,p)$ such that if we set $\lambda' = \lambda - \frac{2-p}{\mu}$, then one still has $\lambda' > 1$. In fact, if $\mu = p$, then $\lambda - \frac{2-p}{\mu} > 1$ since $\lambda > 2/p$; this inequality can then be maintained by a small variation of μ. With this choice of μ, we have by Lemma 6.12

$$|u(x-t,y)| \left(\frac{y}{y+|t|}\right)^{n/\mu} \leqslant A_\mu M_\mu f(x). \tag{6.27}$$

We now proceed the argument with which we treated the function g.

$$\begin{aligned}(g_\lambda^*(f)(x))^2 =&\frac{1}{p(p-1)}\int_{\mathbb{R}_+^{n+1}} y^{1-n}\left(\frac{y}{y+|t|}\right)^{\lambda n} u^{2-p}(x-t,y)|\Delta u^p(x-t,y)|dtdy\\ \leqslant&\frac{1}{p(p-1)}A_\mu^{2-p}(M_\mu f(x))^{2-p}I^*(x),\end{aligned} \tag{6.28}$$

where

$$I^*(x)=\int_{\mathbb{R}_+^{n+1}} y^{1-n}\left(\frac{y}{y+|t|}\right)^{\lambda' n}\Delta u^p(x-t,y)dtdy.$$

It is clear that

$$\begin{aligned}\int_{\mathbb{R}^n} I^*(x)dx =&\int_{\mathbb{R}_+^{n+1}}\int_{\mathbb{R}_x^n} y^{1-n}\left(\frac{y}{y+|t-x|}\right)^{\lambda' n}\Delta u^p(t,y)dxdtdy\\ =&C_{\lambda'}\int_{\mathbb{R}_+^{n+1}} y\Delta u^p(t,y)dtdy.\end{aligned}$$

The last step follows from the fact that if $\lambda'>1$

$$\begin{aligned}y^{-n}\int_{\mathbb{R}^n}\left(\frac{y}{y+|t-x|}\right)^{\lambda' n}dx =&y^{-n}\int_{\mathbb{R}^n}\left(\frac{y}{y+|x|}\right)^{\lambda' n}dx \overset{x=yz}{=\!=}\int_{\mathbb{R}^n}\left(\frac{1}{1+|z|}\right)^{\lambda' n}dz\\ =&C_{\lambda'}<\infty.\end{aligned}$$

So, by Lemma 6.6

$$\int_{\mathbb{R}^n} I^*(x)dx=C_{\lambda'}\int_{\mathbb{R}^n} u^p(t,0)dt=C_{\lambda'}\|f\|_p^p. \tag{6.29}$$

Therefore, by (6.28), Hölder's inequality, (6.24) and (6.29),

$$\|g_\lambda^*(f)\|_p\leqslant C\|M_\mu f(x)^{1-p/2}(I^*(x))^{1/2}\|_p\leqslant C\|M_\mu f\|_p^{1-p/2}\|I^*\|_1^{1/2}\leqslant C\|f\|_p.$$

That is the desired result. □

Finally, we prove Lemma 6.12.

Proof of Lemma 6.12. One notices that (6.25) is unchanged by the dilation $(x,t,y)\to(\delta x,\delta t,\delta y)$, it is then clear that it suffices to prove (6.25) with $y=1$.

Setting $y=1$ in the Poisson kernel, we have $P_1(x)=c_n(1+|x|^2)^{-(n+1)/2}$, and $u(x-t,1)=f(x)*P_1(x-t)$, for each t. Theorem 4.10 shows that $|u(x-t,1)|\leqslant A_tMf(x)$, where $A_t=\int Q_t(x)dx$, and $Q_t(x)$ is the smallest decreasing radial majorant of $P_1(x-t)$, i.e.,

$$Q_t(x)=c_n\sup_{|x'|\geqslant|x|}\frac{1}{(1+|x'-t|^2)^{(n+1)/2}}.$$

For $Q_t(x)$, we have the easy estimates, $Q_t(x)\leqslant c_n$ for $|x|\leqslant 2t$ and $Q_t(x)\leqslant A'(1+|x|^2)^{-(n+1)/2}$, for $|x|\geqslant 2|t|$, from which it is obvious that $A_t\leqslant A(1+|t|)^n$ and hence (6.25) is proved.

Since $u(x-t,y)=\int_{\mathbb{R}^n}P_y(s)f(x-t-s)ds$, and $\int_{\mathbb{R}^n}P_y(s)ds=1$, by Hölder inequality, we have

$$u(x-t,y)\leqslant\|P_y^{1/\mu}f\|_\mu\|P_y^{1/\mu'}\|_{\mu'}\leqslant\left(\int_{\mathbb{R}^n}P_y(s)|f(x-t-s)|^\mu ds\right)^{1/\mu}=U^{1/\mu}(x-t,y),$$

where U is the Poisson integral of $|f|^\mu$. Apply (6.25) to U, it gives

$$|u(x-t,y)|\leqslant A^{1/\mu}(1+|t|/y)^{n/\mu}(M(|f|^\mu)(x))^{1/\mu}=A_\mu(1+|t|/y)^{n/\mu}M_\mu f(x),$$

and the Lemma is established. □

6.2 Fourier multipliers on L^p

In this section, we introduce briefly the Fourier multipliers on L^p, and we prove two (or three) main multiplier theorems.

In the study of PDEs, we often investigate the estimates of semigroups. For example, we consider the linear heat equation

$$u_t - \Delta u = 0, \quad u(0) = u_0.$$

It is clear that $u = \mathscr{F}^{-1} e^{-t|\omega\xi|^2} \mathscr{F} u_0 =: H(t)u_0$ is the solution of the above heat equation. The natural question is: Is $H(t)$ a bounded semigroup from L^p to L^p? In other word, is the following inequality true?

$$\|\mathscr{F}^{-1} e^{-t|\omega\xi|^2} \mathscr{F} u_0\|_p \lesssim \|u_0\|_p, \quad \text{for } 1 \leqslant p \leqslant \infty.$$

Of course, we have known that this estimate is true. From this example, we can give a general concept.

Definition 6.13. Let $\rho \in \mathscr{S}'$. ρ is called a Fourier multiplier on L^p if the convolution $(\mathscr{F}^{-1}\rho) * f \in L^p$ for all $f \in \mathscr{S}$, and if

$$\|\rho\|_{M_p} = \sup_{\|f\|_p = 1} \|(\mathscr{F}^{-1}\rho) * f\|_p$$

is finite. The linear space of all such ρ is denoted by M_p.

Since $\mathscr{S}$ is dense in L^p $(1 \leqslant p < \infty)$, the mapping from $\mathscr{S}$ to L^p: $f \to (\mathscr{F}^{-1}\rho) * f$ can be extended to a mapping from L^p to L^p with the same norm. We write $(\mathscr{F}^{-1}\rho) * f$ also for the values of the extended mapping.

For $p = \infty$ (as well as for $p = 2$) we can characterize M_p. Considering the map:

$$f \to (\mathscr{F}^{-1}\rho) * f \quad \text{ for } f \in \mathscr{S},$$

we have

$$\rho \in M_\infty \Leftrightarrow |\mathscr{F}^{-1}\rho * f(0)| \leqslant C\|f\|_\infty, \quad f \in \mathscr{S}. \tag{6.30}$$

Indeed, if $\rho \in M_\infty$, we have

$$|\mathscr{F}^{-1}\rho * f(0)| \leqslant \frac{\|\mathscr{F}^{-1}\rho * f\|_\infty}{\|f\|_\infty} \|f\|_\infty \leqslant C\|f\|_\infty.$$

On the other hand, if $|\mathscr{F}^{-1}\rho * f(0)| \leqslant C\|f\|_\infty$, we can get

$$\begin{aligned}\|\mathscr{F}^{-1}\rho * f\|_\infty = \sup_{x\in\mathbb{R}^n} |\mathscr{F}^{-1}\rho * f(x)| &= \sup_{x\in\mathbb{R}^n} |[(\mathscr{F}^{-1}\rho) * (f(x+\cdot))](0)| \\ &\leqslant C\|f(x+\cdot)\|_\infty = C\|f\|_\infty,\end{aligned}$$

which yields $\|\rho\|_{M_\infty} \leqslant C$, i.e., $\rho \in M_\infty$.

But (6.30) also means that $\mathscr{F}^{-1}\rho$ is a bounded measure on $\mathbb{R}^n$. Thus M_∞ is equal to the space of all Fourier transforms of bounded measures. Moreover, $\|\rho\|_{M_\infty}$ is equal to the total mass of $\mathscr{F}^{-1}\rho$. In view of the inequality above and the Hahn-Banach theorem, we may extend the mapping $f \to \mathscr{F}^{-1}\rho * f$ from $\mathscr{S}$ to L^∞ to a mapping from L^∞ to L^∞ without increasing its norm. We also write the extended mapping as $f \to \mathscr{F}^{-1}\rho * f$ for $f \in L^\infty$.

Theorem 6.14. *Let $1 \leqslant p \leqslant \infty$ and $1/p + 1/p' = 1$, then we have*

$$M_p = M_{p'} \quad \textit{(equal norms)}. \tag{6.31}$$

Moreover,

$$\begin{aligned} M_1 =& \left\{\rho \in \mathscr{S}' : \mathscr{F}^{-1}\rho \textit{ is a bounded measure}\right\} \\ \|\rho\|_{M_1} =& \textit{total mass of } \mathscr{F}^{-1}\rho = \int_{\mathbb{R}^n} |\mathscr{F}^{-1}\rho(x)|dx \end{aligned} \tag{6.32}$$

and

$$M_2 = L^\infty \quad \textit{(equal norm).} \tag{6.33}$$

For the norms $(1 \leqslant p_0, p_1 \leqslant \infty)$

$$\|\rho\|_{M_p} \leqslant \|\rho\|_{M_{p_0}}^{1-\theta} \|\rho\|_{M_{p_1}}^{\theta}, \quad \forall \rho \in M_{p_0} \cap M_{p_1} \tag{6.34}$$

if $1/p = (1-\theta)/p_0 + \theta/p_1$ $(0 \leqslant \theta \leqslant 1)$*. In particular, the norm* $\|\cdot\|_{M^p}$ *decreases with* p *in the interval* $1 \leqslant p \leqslant 2$*, and*

$$M_1 \subset M_p \subset M_q \subset M_2, \quad (1 \leqslant p \leqslant q \leqslant 2). \tag{6.35}$$

Proof. Let $f \in L^p$, $g \in L^{p'}$ and $\rho \in M_p$. Then, we have

$$\begin{aligned} \|\rho\|_{M_{p'}} =& \sup_{\|g\|_{p'}=1} \|(\mathscr{F}^{-1}\rho) * g\|_{p'} = \sup_{\|f\|_p=\|g\|_{p'}=1} |\langle (\mathscr{F}^{-1}\rho) * g(x), f(-x)\rangle| \\ =& \sup_{\|f\|_p=\|g\|_{p'}=1} |(\mathscr{F}^{-1}\rho) * g * f(0)| = \sup_{\|f\|_p=\|g\|_{p'}=1} |(\mathscr{F}^{-1}\rho) * f * g(0)| \\ =& \sup_{\|f\|_p=\|g\|_{p'}=1} |\int_{\mathbb{R}^n} ((\mathscr{F}^{-1}\rho) * f)(y) g(-y) dy| \\ =& \sup_{\|f\|_p=1} \|(\mathscr{F}^{-1}\rho) * f\|_p = \|\rho\|_{M_p}. \end{aligned}$$

The assertion (6.32) has already been established because of $M_1 = M_\infty$. The Plancherel theorem immediately gives (6.33). In fact,

$$\|\rho\|_{M_2} = \sup_{\|f\|_2=1} \|(\mathscr{F}^{-1}\rho) * f\|_2 = \sup_{\|f\|_2=1} \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\rho\hat{f}\|_2 \leqslant \|\rho\|_\infty.$$

On the other hand, for any $\varepsilon > 0$, we can choose a non-zero measurable set E such that $|\rho(\xi)| \geqslant \|\rho\|_\infty - \varepsilon$ for $\xi \in E$. Then choose a function $f \in L^2$ such that $\operatorname{supp} \mathscr{F} f \subset E$, we can obtain $\|\rho\|_{M_2} \geqslant \|\rho\|_\infty - \varepsilon$.

Invoking the Riesz-Thorin theorem, (6.34) follows, since the mapping $f \to (\mathscr{F}^{-1}\rho) * f$ maps $L^{p_0} \to L^{p_0}$ with norm $\|\rho\|_{M_{p_0}}$ and $L^{p_1} \to L^{p_1}$ with norm $\|\rho\|_{M_{p_1}}$.

Since $1/q = (1-\theta)/p + \theta/p'$ for some θ and $p \leqslant q \leqslant 2 \leqslant p'$, by using (6.34) with $p_0 = p$, $p_1 = p'$, we see that

$$\|\rho\|_{M_q} \leqslant \|\rho\|_{M_p},$$

from which (6.35) follows. □

Proposition 6.15. *Let* $1 \leqslant p \leqslant \infty$*. Then* M_p *is a Banach algebra under pointwise multiplication.*

Proof. It is clear that $\|\cdot\|_{M_p}$ is a norm. Note also that M_p is complete. Indeed, let $\{\rho_k\}$ is a Cauchy sequence in M_p. So does it in L^∞ because of $M_p \subset L^\infty$. Thus, it is convergent in L^∞ and we denote the limit by ρ. From $L^\infty \subset \mathscr{S}'$, we have $\mathscr{F}^{-1}\rho_k \mathscr{F} f \to \mathscr{F}^{-1}\rho \mathscr{F} f$ for any $f \in \mathscr{S}$ in sense of the strong topology on $\mathscr{S}'$. On the other hand, $\left\{\mathscr{F}^{-1}\rho_k \mathscr{F} f\right\}$ is also a Cauchy sequence in $L^p \subset \mathscr{S}'$, and converges to a function $g \in L^p$. By the uniqueness of limit in $\mathscr{S}'$, we know that $g = \mathscr{F}^{-1}\rho\mathscr{F} f$. Thus, $\|\rho_k - \rho\|_{M_p} \to 0$ as $k \to \infty$. Therefore, M_p is a Banach space.

Let $\rho_1 \in M_p$ and $\rho_2 \in M_p$. For any $f \in \mathscr{S}$, we have

$$\|(\mathscr{F}^{-1}\rho_1\rho_2) * f\|_p = \|(\mathscr{F}^{-1}\rho_1) * (\mathscr{F}^{-1}\rho_2) * f\|_p \leqslant \|\rho_1\|_{M_p} \|(\mathscr{F}^{-1}\rho_2) * f\|_p$$

$$\leqslant \|\rho_1\|_{M_p}\|\rho_2\|_{M_p}\|f\|_p,$$

which implies $\rho_1\rho_2 \in M_p$ and

$$\|\rho_1\rho_2\|_{M_p} \leqslant \|\rho_1\|_{M_p}\|\rho_2\|_{M_p}.$$

Thus, M_p is a Banach algebra. □

In order to clarify the next theorem we write $M_p = M_p(\mathbb{R}^n)$ for Fourier multipliers which are functions on $\mathbb{R}^n$. The next theorem says that $M_p(\mathbb{R}^n)$ is isometrically invariant under affine transforms of $\mathbb{R}^n$.

Theorem 6.16. *Let $a : \mathbb{R}^n \to \mathbb{R}^m$ be a surjective affine transform*[1] *with $n \geqslant m$, and $\rho \in M_p(\mathbb{R}^m)$. Then*

$$\|\rho(a(\cdot))\|_{M_p(\mathbb{R}^n)} = \|\rho\|_{M_p(\mathbb{R}^m)}.$$

If $m = n$, the mapping a^* is bijective. In particular, we have

$$\|\rho(c\cdot)\|_{M_p(\mathbb{R}^n)} = \|\rho(\cdot)\|_{M_p(\mathbb{R}^n)}, \quad \forall c \neq 0, \tag{6.36}$$

$$\|\rho(\langle x, \cdot\rangle)\|_{M_p(\mathbb{R}^n)} = \|\rho(\cdot)\|_{M_p(\mathbb{R})}, \quad \forall x \neq 0, \tag{6.37}$$

where $\langle x, \xi\rangle = \sum_{i=1}^n x_i\xi_i$.

Proof. It suffices to consider the case that $a : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transform. Make the coordinate transform

$$\eta_i = a_i(\xi),\ 1 \leqslant i \leqslant m; \quad \eta_j = \xi_j,\ m+1 \leqslant j \leqslant n, \tag{6.38}$$

which can be written as $\eta = A^{-1}\xi$ or $\xi = A\eta$ where $\det A \neq 0$. Let $A^\top$ be the transposed matrix of A. It is easy to see, for any $f \in \mathscr{S}(\mathbb{R}^n)$, that

$$\begin{aligned}
\mathscr{F}^{-1}\rho(a(\xi))\mathscr{F}f(x) &= \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\xi}\rho(a(\xi))\hat{f}(\xi)d\xi \\
&= |\det A| \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot A\eta}\rho(\eta_1, \cdots, \eta_m)\hat{f}(A\eta)d\eta \\
&= |\det A| \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i A^\top x\cdot\eta}\rho(\eta_1, \cdots, \eta_m)\hat{f}(A\eta)d\eta \\
&= |\det A|(\mathscr{F}^{-1}\rho(\eta_1, \cdots, \eta_m)\hat{f}(A\eta))(A^\top x) \\
&= \left[\mathscr{F}^{-1}\rho(\eta_1, \cdots, \eta_m)\left(\mathscr{F}f((A^\top)^{-1}\cdot)\right)(\eta)\right](A^\top x).
\end{aligned}$$

It follows from $\rho \in M_p(\mathbb{R}^m)$ that for any $f \in \mathscr{S}(\mathbb{R}^n)$

$$\begin{aligned}
&\|\mathscr{F}^{-1}\rho(a(\xi))\mathscr{F}f\|_p = |\det A|^{-1/p}\|\mathscr{F}^{-1}\rho(\eta_1, \cdots, \eta_m)\left(\mathscr{F}f((A^\top)^{-1}\cdot)\right)(\eta)\|_p \\
=&|\det A|^{-1/p}\left\|\left(\mathscr{F}^{-1}_{\eta_1,\cdots,\eta_m}\rho(\eta_1, \cdots, \eta_m)\right) * \|f((A^\top)^{-1}\cdot)\|_{L^p(\mathbb{R}^{n-m})}\right\|_{L^p(\mathbb{R}^m)} \\
\leqslant&\|\rho\|_{M_p(\mathbb{R}^m)}\|f\|_p.
\end{aligned}$$

Thus, we have

$$\|\rho(a(\cdot))\|_{M_p(\mathbb{R}^n)} \leqslant \|\rho\|_{M_p(\mathbb{R}^m)}. \tag{6.39}$$

Taking $f((A^\top)^{-1}\cdot) = f_1(x_1, \cdots, x_m)f_2(x_{m+1}, \cdots, x_n)$, one can conclude that the inverse inequality (6.39) also holds. □

Now we give a simple but very useful theorem for Fourier multipliers.

[1] An affine transform of $\mathbb{R}^n$ is a map $F : \mathbb{R}^n \to \mathbb{R}^n$ of the form $F(\mathbf{p}) = A\mathbf{p} + \mathbf{q}$ for all $\mathbf{p} \in \mathbb{R}^n$, where A is a linear transform of $\mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$.

Theorem 6.17 (Bernstein multiplier theorem)**.** *Assume that $k > n/2$ is an integer, and that $\partial_{x_j}^{\alpha}\rho \in L^2(\mathbb{R}^n)$, $j = 1, \cdots, n$ and $0 \leqslant \alpha \leqslant k$. Then we have $\rho \in M_p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, and*

$$\|\rho\|_{M_p} \lesssim \|\rho\|_2^{1-n/2k} \left(\sum_{j=1}^{n} \|\partial_{x_j}^k \rho\|_2 \right)^{n/2k}.$$

Proof. Let $t > 0$ and $J(x) = \sum_{j=1}^n |x_j|^k$. By the Cauchy-Schwartz inequality and the Plancherel theorem, we obtain

$$\int_{|x|>t} |\mathscr{F}^{-1}\rho(x)|dx = \int_{|x|>t} J(x)^{-1}J(x)|\mathscr{F}^{-1}\rho(x)|dx \lesssim t^{n/2-k} \sum_{j=1}^{n} \|\partial_{x_j}^k \rho\|_2.$$

Similarly, we have

$$\int_{|x|\leqslant t} |\mathscr{F}^{-1}\rho(x)|dx \lesssim t^{n/2}\|\rho\|_2.$$

Choosing t such that $\|\rho\|_2 = t^{-k} \sum_{j=1}^n \|\partial_{x_j}^k \rho\|_2$, we infer, with the help of Theorem 6.14, that

$$\|\rho\|_{M_p} \leqslant \|\rho\|_{M_1} = \int_{\mathbb{R}^n} |\mathscr{F}^{-1}\rho(x)|dx \lesssim \|\rho\|_2^{1-n/2k} \left(\sum_{j=1}^{n} \|\partial_{x_j}^k \rho\|_2 \right)^{n/2k}.$$

This completes the proof. □

The first application of the theory of the functions g and g_λ^* will be in the study of multipliers. Our main tool when proving theorems for the Sobolev and Besov spaces, defined in the following chapters, is the following theorem. Note that $1 < p < \infty$ here in contrast to the case in Theorem 6.17. We give the theorem as follows.

Theorem 6.18 (Mihlin multiplier theorem)**.** *Suppose that $\rho(\xi) \in C^k(\mathbb{R}^n \setminus \{0\})$ where $k > n/2$ is an integer. Assume also that for every differential monomial $\left(\frac{\partial}{\partial\xi}\right)^\alpha$, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, with $|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_n$, we have Mihlin's condition*

$$\left| \left(\frac{\partial}{\partial\xi}\right)^\alpha \rho(\xi) \right| \leqslant A|\xi|^{-|\alpha|}, \quad \textit{whenever } |\alpha| \leqslant k. \tag{6.40}$$

Then $\rho \in M_p$, $1 < p < \infty$, and

$$\|\rho\|_{M_p} \leqslant C_{p,n}A.$$

The proof of the theorem leads to a generalization of its statement which we formulate as a corollary.

Corollary 6.19 (Hörmander multiplier theorem)**.** *The assumption* (6.40) *can be replaced by the weaker assumptions, i.e., Hörmander's condition*

$$\begin{aligned} |\rho(\xi)| &\leqslant A, \\ \sup_{0<R<\infty} R^{2|\alpha|-n} \int_{R\leqslant|\xi|\leqslant 2R} \left| \left(\frac{\partial}{\partial\xi}\right)^\alpha \rho(\xi) \right|^2 d\xi &\leqslant A, \quad |\alpha| \leqslant k. \end{aligned} \tag{6.41}$$

The theorem and its corollary will be consequences of the following lemma. Its statement illuminates at the same time the nature of the multiplier transforms considered here, and the role played by the g-functions and their variants.

Lemma 6.20. *Under the assumptions of Theorem 6.18 or Corollary 6.19, let us set for $f \in L^2(\mathbb{R}^n)$*

$$F(x) = T_\rho f(x) = (\mathscr{F}^{-1}(\rho(\xi)) * f)(x).$$

Then

$$g_1(F)(x) \leqslant A_\lambda g_\lambda^*(f)(x), \quad \text{where } \lambda = 2k/n. \tag{6.42}$$

Thus in view of the lemma, the g-functions and their variants are the characterizing expressions which deal at once with all the multipliers considered. On the other hand, the fact that the relation (6.42) is pointwise shows that to a large extent the mapping T_ρ is "semi-local".

Proof of Theorem 6.18 and Corollary 6.19. The conclusion is deduced from the lemma as follows. Our assumption on k is such that $\lambda = 2k/n > 1$. Thus, Theorem 6.11 shows us that

$$\|g_\lambda^*(f)(x)\|_p \leqslant A_{\lambda,p}\|f\|_p, \quad 2 \leqslant p < \infty, \text{ if } f \in L^2 \cap L^p.$$

However, by Corollary 6.3, $A_p'\|F\|_p \leqslant \|g_1(F)(x)\|_p$, therefore by Lemma 6.20,

$$\|T_\rho f\|_p = \|F\|_p \leqslant A_\lambda \|g_\lambda^*(f)(x)\|_p \leqslant A_p\|f\|_p, \quad \text{if } 2 \leqslant p < \infty \text{ and } f \in L^2 \cap L^p.$$

That is, $\rho \in M_p$, $2 \leqslant p < \infty$. By duality, i.e., (6.31) of Theorem 6.14, we have also $\rho \in M_p$, $1 < p \leqslant 2$, which gives the assertion of the theorem. □

Now we shall prove Lemma 6.20.

Proof of Lemma 6.20. Let $u(x,y)$ denote the Poisson integral of f, and $U(x,y)$ the Poisson integral of F. Then with ˆ denoting the Fourier transform w.r.t. the x variable, we have

$$\hat{u}(\xi, y) = e^{-|\omega\xi|y}\hat{f}(\xi), \text{ and } \widehat{U}(\xi, y) = e^{-|\omega\xi|y}\widehat{F}(\xi) = e^{-|\omega\xi|y}\rho(\xi)\hat{f}(\xi).$$

Define $M(x,y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i x\cdot\xi} e^{-|\omega\xi|y}\rho(\xi)d\xi$. Then clearly $\widehat{M}(\xi,y) = e^{-|\omega\xi|y}\rho(\xi)$, and so

$$\widehat{U}(\xi, y_1 + y_2) = \widehat{M}(\xi, y_1)\hat{u}(\xi, y_2), \quad y = y_1 + y_2, \quad y_1, y_2 > 0.$$

This can be written as

$$U(x, y_1 + y_2) = \int_{\mathbb{R}^n} M(t, y_1)u(x - t, y_2)dt.$$

We differentiate this relation k times w.r.t. y_1 and once w.r.t. y_2, and set $y_1 = y_2 = y/2$. This gives us the identity

$$U^{(k+1)}(x,y) = \int_{\mathbb{R}^n} M^{(k)}(t, y/2)u^{(1)}(x - t, y/2)dt. \tag{6.43}$$

Here the superscripts denote the differentiation w.r.t. y.

Next, we translates the assumptions (6.40) (or (6.41)) on ρ in terms of $M(x,y)$. The result is

$$|M^{(k)}(t,y)| \leqslant A'y^{-n-k}, \tag{6.44}$$

$$\int_{\mathbb{R}^n} |t|^{2k}|M^{(k)}(t,y)|^2 dt \leqslant A'y^{-n}. \tag{6.45}$$

In fact, by the definition of M and the condition $|\rho(\xi)| \leqslant A$, it follows that

$$\begin{aligned}|M^{(k)}(x,y)| \leqslant & \left(\frac{|\omega|}{2\pi}\right)^n |\omega|^k \int_{\mathbb{R}^n} |\xi|^k e^{-|\omega\xi|y}\rho(\xi)d\xi \\ \leqslant & A\omega_{n-1}\left(\frac{|\omega|}{2\pi}\right)^n |\omega|^k \int_0^\infty r^k e^{-|\omega|ry} r^{n-1}dr = A'y^{-n-k},\end{aligned}$$

which is (6.44).

To prove (6.45), let us show more particularly that

$$\int_{\mathbb{R}^n} |x^\alpha M^{(k)}(x,y)|^2 dx \leqslant A'y^{-n},$$

where $|\alpha| = k$.

By Plancherel's theorem

$$\|x^\alpha M^{(k)}(x,y)\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{n/2} |\omega|^k \left\| \left(\frac{\partial}{\partial\xi}\right)^\alpha (|\xi|^k \rho(\xi) e^{-|\omega\xi|y}) \right\|_2. \tag{6.46}$$

So we need to evaluate, by using Leibniz' rule,

$$\left(\frac{\partial}{\partial\xi}\right)^\alpha (|\xi|^k \rho(\xi) e^{-|\omega\xi|y}) = \sum_{\beta+\gamma=\alpha} C_{\beta,\gamma} \left(\frac{\partial}{\partial\xi}\right)^\beta (|\xi|^k \rho(\xi)) \left(\frac{\partial}{\partial\xi}\right)^\gamma e^{-|\omega\xi|y}. \tag{6.47}$$

Case I: (6.40) $\Longrightarrow$ (6.45). By the hypothesis (6.40) and Leibniz' rule again, we have

$$\left| \left(\frac{\partial}{\partial\xi}\right)^\beta (|\xi|^k \rho(\xi)) \right| \leqslant A' |\xi|^{k-|\beta|}, \text{ with } |\beta| \leqslant k.$$

Thus,

$$\left| \left(\frac{\partial}{\partial\xi}\right)^\alpha (|\xi|^k \rho(\xi) e^{-|\omega\xi|y}) \right| \leqslant C \sum_{|\beta|+|\gamma|=k} |\xi|^{k-|\beta|} y^{|\gamma|} e^{-|\omega\xi|y} \leqslant C \sum_{0\leqslant r\leqslant k} |\xi|^r y^r e^{-|\omega\xi|y}.$$

Since for $r \geqslant 0$

$$\begin{aligned} y^{2r} \int_{\mathbb{R}^n} |\xi|^{2r} e^{-2|\omega\xi|y} d\xi =& C y^{2r} \int_0^\infty R^{2r} e^{-2|\omega|Ry} R^{n-1} dR \\ =& C y^{-n} \int_0^\infty z^{2r} e^{-2|\omega|z} z^{n-1} dz \leqslant C y^{-n}, \end{aligned}$$

we get

$$\|x^\alpha M^{(k)}(x,y)\|_2^2 \leqslant A' y^{-n}, \quad |\alpha| = k,$$

which proves the assertion (6.45).

Case II: (6.41) $\Longrightarrow$ (6.45). From (6.46) and (6.47), we have, by Leibniz' rule again and (6.41),

$$\begin{aligned} &\|x^\alpha M^{(k)}(x,y)\|_2 \\ \leqslant& C \sum_{|\beta'|+|\beta''|+|\gamma|=k} \left(\int_{\mathbb{R}^n} \left| \left(\frac{\partial}{\partial\xi}\right)^{\beta'} |\xi|^k \right|^2 \left| \left(\frac{\partial}{\partial\xi}\right)^{\beta''} \rho(\xi) \right|^2 e^{-2|\omega\xi|y} y^{2|\gamma|} d\xi \right)^{1/2} \\ \leqslant& C \sum_{|\beta'|+|\beta''|+|\gamma|=k} y^{|\gamma|} \left(\sum_{j\in\mathbb{Z}} \int_{2^j y\leqslant|\xi|\leqslant 2^{j+1}y} |\xi|^{2(k-|\beta'|)} \left| \left(\frac{\partial}{\partial\xi}\right)^{\beta''} \rho(\xi) \right|^2 e^{-2|\omega\xi|y} d\xi \right)^{1/2} \\ \leqslant& C \sum_{|\beta'|+|\beta''|+|\gamma|=k} \sum_{j\in\mathbb{Z}} (2^{j+1}y)^{k-|\beta'|} y^{|\gamma|} e^{-|\omega|2^j y^2} \\ &\qquad \cdot (2^j y)^{-|\beta''|+n/2} \left((2^j y)^{2|\beta''|-n} \int_{2^j y\leqslant|\xi|\leqslant 2^{j+1}y} \left| \left(\frac{\partial}{\partial\xi}\right)^{\beta''} \rho(\xi) \right|^2 d\xi \right)^{1/2} \\ \leqslant& CA \sum_{|\gamma|\leqslant k} \sum_{j\in\mathbb{Z}} (2^j y)^{|\gamma|+n/2} y^{|\gamma|} e^{-|\omega|2^j y^2} = CAy^{-n/2} \sum_{0\leqslant r\leqslant k} \sum_{j\in\mathbb{Z}} (2^j y^2)^r e^{-|\omega|2^j y^2} \\ \leqslant& C y^{-n/2}, \end{aligned}$$

which yields (6.45).

Now, we return to the identity (6.43), and for each y divide the range of integration into two parts, $|t| \leqslant y/2$ and $|t| > y/2$. In the first range, use the estimate (6.44) on $M^{(k)}$ and in the second range, use the estimate (6.45). This together with Schwarz' inequality gives immediately

$$
\begin{aligned}
|U^{(k+1)}(x,y)|^2 \leqslant & Cy^{-n-2k}\int_{|t|\leqslant y/2}|u^{(1)}(x-t,y/2)|^2dt\\
&+Cy^{-n}\int_{|t|>y/2}\frac{|u^{(1)}(x-t,y/2)|^2dt}{|t|^{2k}}\\
=:& I_1(y)+I_2(y).
\end{aligned}
$$

Now

$$
(g_{k+1}(F)(x))^2=\int_0^\infty |U^{(k+1)}(x,y)|^2y^{2k+1}dy\leqslant\sum_{j=1}^2\int_0^\infty I_j(y)y^{2k+1}dy.
$$

However, by a change of variable $y/2\to y$,

$$
\begin{aligned}
\int_0^\infty I_1(y)y^{2k+1}dy\leqslant & C\int_0^\infty\int_{|t|\leqslant y/2}|u^{(1)}(x-t,y/2)|^2y^{-n+1}dtdy\\
\leqslant & C\int_\Gamma |\nabla u(x-t,y)|^2y^{-n+1}dtdy=C(S(f)(x))^2\\
\leqslant & C_\lambda(g_\lambda^*(f)(x))^2.
\end{aligned}
$$

Similarly, with $n\lambda=2k$,

$$
\begin{aligned}
\int_0^\infty I_2(y)y^{2k+1}dy\leqslant & C\int_0^\infty\int_{|t|>y}y^{-n+2k+1}|t|^{-2k}|\nabla u(x-t,y)|^2dtdy\\
\leqslant & C(g_\lambda^*(f)(x))^2.
\end{aligned}
$$

This shows that $g_{k+1}(F)(x)\leqslant C_\lambda g_\lambda^*(f)(x)$. However by Remark 6.4 (iii) of g-functions after Corollary 6.3, we know that $g_1(F)(x)\leqslant C_kg_{k+1}(F)(x)$. Thus, the proof of the lemma is concluded. □

6.3 The partial sums operators

We shall now develop the second main tool in the Littlewood-Paley theory, (the first being the usage of the functions g and g^*).

Let ρ denote an arbitrary rectangle in $\mathbb{R}^n$. By rectangle we shall mean, in the rest of this chapter, a possibly infinite rectangle with sides parallel to the axes, i.e., the Cartesian product of n intervals.

Definition 6.21. For each rectangle ρ denote by S_ρ the *partial sum operator*, that is the multiplier operator with $m=\chi_\rho=$ characteristic function of the rectangle ρ. So

$$
\mathscr{F}(S_\rho(f))=\chi_\rho\hat{f},\quad f\in L^2(\mathbb{R}^n)\cap L^p(\mathbb{R}^n). \tag{6.48}
$$

For this operator, we immediately have the following theorem.

Theorem 6.22.

$$
\|S_\rho(f)\|_p\leqslant A_p\|f\|_p,\quad f\in L^2\cap L^p,
$$

if $1<p<\infty$. *The constant* A_p *does not depend on the rectangle* ρ.

However, we shall need a more extended version of the theorem which arises when we replace complex-valued functions by functions taking their value in a Hilbert space.

Let $\mathscr{H}$ be the sequence Hilbert space,

$$
\mathscr{H}=\{(c_j)_{j=1}^\infty:(\sum_j|c_j|^2)^{1/2}=|c|<\infty\}.
$$

Then we can represent a function $f \in L^p(\mathbb{R}^n, \mathscr{H})$, as sequences

$$f(x) = (f_1(x), \cdots, f_j(x), \cdots),$$

where each f_j is complex-valued and $|f(x)| = (\sum_{j=1}^{\infty} |f_j(x)|^2)^{1/2}$. Let $\Re$ be a sequence of rectangle, $\Re = \{\rho_j\}_{j=1}^{\infty}$. Then we can define the operator $S_{\Re}$, mapping $L^2(\mathbb{R}^n, \mathscr{H})$ to itself, by the rule

$$S_{\Re}(f) = (S_{\rho_1}(f_1), \cdots, S_{\rho_j}(f_j), \cdots), \text{ where } f = (f_1, \cdots, f_j, \cdots). \tag{6.49}$$

We first give a lemma, which will be used in the proof of the theorem or its generalization. Recall the Hilbert transform $f \to H(f)$, which corresponds to the multiplier $-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi)$ in one dimension.

Lemma 6.23. *Let $f(x) = (f_1(x), \cdots, f_j(x), \cdots) \in L^2(\mathbb{R}^n, \mathscr{H}) \cap L^p(\mathbb{R}^n, \mathscr{H})$. Denote $\widetilde{H} f(x) = (Hf_1(x), \cdots, Hf_j(x), \cdots)$. Then*

$$\|\widetilde{H} f\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty,$$

where A_p is the same constant as in the scalar case, i.e., when $\mathscr{H}$ is one-dimensional.

Proof. We use the vector-valued version of the Hilbert transform, as is described more generally in Sec. 4.7. Let the Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ be both identical with $\mathscr{H}$. Take in $\mathbb{R}$, $K(x) = I \cdot 1/\pi x$, where I is the identity mapping on $\mathscr{H}$. Then the kernel $K(x)$ satisfies all the assumptions of Theorem 4.27 and Theorem 4.24. Moreover,

$$\lim_{\varepsilon \to 0} \int_{|y| \geqslant \varepsilon} K(y) f(x-y) dy = \widetilde{H} f(x),$$

and so our lemma is proved. □

The generalization of Theorem 6.22 is then as follows.

Theorem 6.24. *Let $f \in L^2(\mathbb{R}^n, \mathscr{H}) \cap L^p(\mathbb{R}^n, \mathscr{H})$. Then*

$$\|S_{\Re}(f)\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{6.50}$$

where A_p does not depend on the family $\Re$ of rectangles.

Proof. The theorem will be proved in four steps, the first two of which already contain the essence of the matter.

Step 1: $n = 1$, and the rectangles $\rho_1, \rho_2, \cdots, \rho_j, \cdots$ are the semi-infinite intervals $(-\infty, 0)$.

It is clear that $S_{(-\infty,0)} f = \mathscr{F}^{-1} \chi_{(-\infty,0)} \mathscr{F} f = \mathscr{F}^{-1} \frac{1-\operatorname{sgn}(\xi)}{2} \mathscr{F} f$, so

$$S_{(-\infty,0)} = \frac{I - i \operatorname{sgn}(\omega) H}{2}, \tag{6.51}$$

where I is the identity, and $S_{(-\infty,0)}$ is the partial sum operator corresponding to the interval $(-\infty, 0)$.

Now if all the rectangles are the intervals $(-\infty, 0)$, then by (6.51),

$$S_{\Re} = \frac{I - i \operatorname{sgn}(\omega) \widetilde{H}}{2}$$

and so by Lemma 6.23, we have the desired result.

Step 2: $n = 1$, and the rectangles are the intervals $(-\infty, a_1)$, $(-\infty, a_2)$, $\cdots$, $(-\infty, a_j)$, $\cdots$.

Notice that $\mathscr{F}(f(x) e^{-\omega i x \cdot a}) = \hat{f}(\xi + a)$, therefore

$$\mathscr{F}(H(e^{-\omega i x \cdot a} f(x))) = -i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \hat{f}(\xi + a),$$

and hence $\mathscr{F}(e^{\omega i x \cdot a} H(e^{-\omega i x \cdot a} f(x))) = -i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi - a) \hat{f}(\xi)$. From this, we see that

$$(S_{(-\infty,a_j)} f_j)(x) = \frac{f_j - i \operatorname{sgn}(\omega) e^{\omega i x \cdot a_j} H(e^{-\omega i x \cdot a_j} f_j)}{2}. \tag{6.52}$$

If we now write symbolically $e^{-\omega ix\cdot a}f$ for

$$(e^{-\omega ix\cdot a_1}f_1,\cdots,e^{-\omega ix\cdot a_j}f_j,\cdots)$$

with $f=(f_1,\cdots,f_j,\cdots)$, then (6.52) may be written as

$$S_{\Re}f=\frac{f-i\,\mathrm{sgn}\,(\omega)e^{\omega ix\cdot a}\widetilde{H}(e^{-\omega ix\cdot a}f)}{2}, \tag{6.53}$$

and so the result again follows in this case by Lemma 6.23.

Step 3: General n, but the rectangles ρ_j are the half-spaces $x_1<a_j$, i.e., $\rho_j=\{x:x_1<a_j\}$.

Let $S^{(1)}_{(-\infty,a_j)}$ denote the operator defined on $L^2(\mathbb{R}^n)$, which acts only on the x_1 variable, by the action given by $S_{(-\infty,a_j)}$. We claim that

$$S_{\rho_j}=S^{(1)}_{(-\infty,a_j)}. \tag{6.54}$$

This identity is obvious for L^2 functions of the product form

$$f'(x_1)f''(x_2,\cdots,x_n),$$

since their linear span is dense in L^2, the identity (6.54) is established.

We now use the L^p inequality, which is the result of the previous step for each fixed x_2, x_3, $\cdots$, x_n. We raise this inequality to the p^{th} power and integrate w.r.t. x_2, $\cdots$, x_n. This gives the desired result for the present case. Notice that the result holds as well if the half-space $\{x:x_1<a_j\}_{j=1}^{\infty}$, is replaced by the half-space $\{x:x_1>a_j\}_{j=1}^{\infty}$, or if the role of the x_1 axis is taken by the x_2 axis, etc.

Step 4: Observe that every general finite rectangle of the type considered is the intersection of $2n$ half-spaces, each half-space having its boundary hyperplane perpendicular to one of the axes of $\mathbb{R}^n$. Thus a $2n$-fold application of the result of the third step proves the theorem, where the family $\Re$ is made up of finite rectangles. Since the bounds obtained do not depend on the family $\Re$, we can pass to the general case where $\Re$ contains possibly infinite rectangles by an obvious limiting argument. □

We state here the continuous analogue of Theorem 6.24. Let $(\varGamma,d\gamma)$ be a σ-finite measure space,[2] and consider the Hilbert space $\mathscr{H}$ of square integrable functions on $\varGamma$, i.e., $\mathscr{H}=L^2(\varGamma,d\gamma)$. The elements

$$f\in L^p(\mathbb{R}^n,\mathscr{H})$$

are the complex-valued functions $f(x,\gamma)=f_\gamma(x)$ on $\mathbb{R}^n\times\varGamma$, which are jointly measuable, and for which $(\int_{\mathbb{R}^n}(\int_\varGamma|f(x,\gamma)|^2d\gamma)^{p/2}dx)^{1/p}=\|f\|_p<\infty$, if $p<\infty$. Let $\Re=\{\rho_\gamma\}_{\gamma\in\varGamma}$, and suppose that the mapping $\gamma\to\rho_\gamma$ is a measurable function from $\varGamma$ to rectangles; that is, the numerical-valued functions which assign to each γ the components of the vertices of ρ_γ are all measurable.

Suppose $f\in L^2(\mathbb{R}^n,\mathscr{H})$. Then we define $F=S_{\Re}f$ by the rule

$$F(x,\gamma)=S_{\rho_\gamma}(f_\gamma)(x),\quad (f_\gamma(x)=f(x,\gamma)).$$

Theorem 6.25.

$$\|S_{\Re}f\|_p\leqslant A_p\|f\|_p,\quad 1<p<\infty, \tag{6.55}$$

for $f\in L^2(\mathbb{R}^n,\mathscr{H})\cap L^p(\mathbb{R}^n,\mathscr{H})$, where the bound A_p does not depend on the measure space $(\varGamma,d\gamma)$, or on the function $\gamma\to\rho_\gamma$.

---

[2] If μ is measure on a ring R, a set E is said to have σ-finite measure if there exists a sequence $\{E_n\}$ of sets in R such that $E\subset\cup_{n=1}^{\infty}E_n$, and $\mu(E_n)<\infty$, $n=1,2,\cdots$. If the measure of every set E in R is σ-finite, the measure μ is called σ-finite on R.

Proof. The proof of this theorem is an exact repetition of the argument given for Theorem 6.24. The reader may also obtain it from Theorem 6.24 by a limiting argument. □

6.4 The dyadic decomposition

We shall now consider a decomposition of $\mathbb{R}^n$ into rectangles.

First, in the case of $\mathbb{R}$, we decompose it as the union of the "disjoint" intervals (i.e., whose interiors are disjoint) $[2^k, 2^{k+1}]$, $-\infty < k < \infty$, and $[-2^{k+1}, -2^k]$, $-\infty < k < \infty$. This double collection of intervals, one collection for the positive half-line, the other for the negative half-line, will be the dyadic decomposition of $\mathbb{R}$.[3]

Having obtained this decomposition of $\mathbb{R}$, we take the corresponding product decomposition for $\mathbb{R}^n$. Thus we write $\mathbb{R}^n$ as the union of "disjoint" rectangles, which rectangles are products of the intervals which occur for the dyadic decomposition of each of the axes. This is the *dyadic decomposition of* $\mathbb{R}^n$.

The family of resulting rectangles will be denoted by Δ. We recall the partial sum operator S_ρ, defined in (6.48) for each rectangle. Now in an obvious sense, (e.g. L^2 convergence)

$$\sum_{\rho\in\Delta} S_\rho = \text{Identity}.$$

Also in the L^2 case, the different blocks, $S_\rho f$, $\rho \in \Delta$, behave as if they were independent; they are of course mutually orthogonal. To put the matter precisely: The L^2 norm of f can be given exactly in terms of the L^2 norms of $S_\rho f$, i.e.,

$$\sum_{\rho\in\Delta} \|S_\rho f\|_2^2 = \|f\|_2^2, \tag{6.56}$$

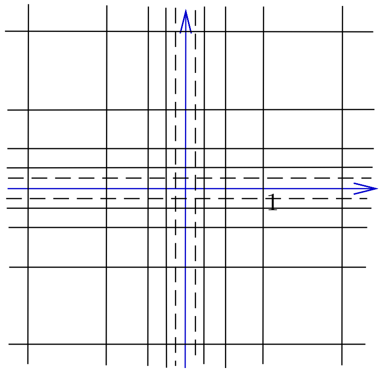


Fig. 6.3 The dyadic decomposition

(and this is true for any decomposition of $\mathbb{R}^n$). For the general L^p case not as much can be hoped for, but the following important theorem can nevertheless be established.

Theorem 6.26 (Littlewood-Paley square function theorem)**.** *Suppose* $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. *Then*

$$\|(\sum_{\rho\in\Delta} |S_\rho f(x)|^2)^{1/2}\|_p \sim \|f\|_p.$$

The Rademacher functions provide a very useful device in the study of L^p norms in terms of quadratic expressions.

These functions, $r_0(t)$, $r_1(t)$, $\cdots$, $r_m(t)$, $\cdots$ are defined on the interval $(0,1)$ as follows:

$$r_0(t) = \begin{cases} 1, & 0 \leqslant t \leqslant 1/2, \\ -1, & 1/2 < t < 1, \end{cases}$$

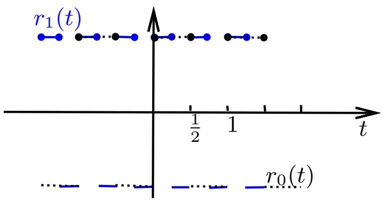


Fig. 6.4 $r_0(t)$ and $r_1(t)$

r_0 is extended outside the unit interval by periodicity, i.e., $r_0(t+1) = r_0(t)$. In general, $r_m(t) = r_0(2^m t)$. The

[3] Strictly speaking, the origin is left out; but for the sake of simplicity of terminology, we still refer to it as the decomposition of $\mathbb{R}$.

sequences of Rademacher functions are orthonormal (and in fact mutually independent) over $[0,1]$. In fact, for $m<k$, the integral

$$\begin{aligned}
&\int_0^1 r_m(t)r_k(t)dt=\int_0^1 r_0(2^m t)r_0(2^k t)dt=2^{-m}\int_0^{2^m} r_0(s)r_0(2^{k-m}s)ds\\
=&\int_0^1 r_0(s)r_0(2^{k-m}s)ds=\int_0^{1/2} r_0(2^{k-m}s)ds-\int_{1/2}^1 r_0(2^{k-m}s)ds\\
=&2^{m-k}\left[\int_0^{2^{k-m-1}} r_0(t)dt-\int_{2^{k-m-1}}^{2^{k-m}} r_0(t)dt\right]\\
=&2^{-1}\left[\int_0^1 r_0(t)dt-\int_0^1 r_0(t)dt\right]=0,
\end{aligned}$$

so, they are orthogonal. It is clear that they are normal since $\int_0^1 (r_m(t))^2dt=1$.

For our purposes, their importance arises from the following fact.

Suppose $\sum_{m=0}^\infty |a_m|^2<\infty$ and set $F(t)=\sum_{m=0}^\infty a_m r_m(t)$. Then for every $1<p<\infty$, $F(t)\in L^p[0,1]$ and

$$A_p\|F\|_p\leqslant \|F\|_2=(\sum_{m=0}^\infty |a_m|^2)^{1/2}\leqslant B_p\|F\|_p, \tag{6.57}$$

for two positive constants A_p and B_p.

Thus, for functions which can be expanded in terms of the Rademacher functions, all the L^p norms, $1<p<\infty$, are comparable.

We shall also need the n-dimensional form of (6.57). We consider the unit cube $Q\subset\mathbb{R}^n$, $Q=\{t=(t_1,t_2,\cdots,t_n): 0\leqslant t_j\leqslant 1\}$. Let m be an n-tuple of non-negative integers $m=(m_1,m_2,\cdots,m_n)$. Define $r_m(t)=r_{m_1}(t_1)r_{m_2}(t_2)\cdots r_{m_n}(t_n)$. Write $F(t)=\sum a_m r_m(t)$. With

$$\|F\|_p=\left(\int_Q |F(t)|^p dt\right)^{1/p},$$

we also have (6.57), whenever $\sum |a_m|^2<\infty$. That is

Lemma 6.27. *Suppose $\sum|a_m|^2<\infty$. Then it holds*

$$\|F\|_p\sim\|F\|_2=\left(\sum_{m=0}^\infty |a_m|^2\right)^{1/2},\quad 1<p<\infty. \tag{6.58}$$

Proof. We split the proof into four steps.

Step 1: Let μ, a_0, a_1, $\cdots$, a_N, be real numbers. Then because the Rademacher functions are mutually independent variables, we have, in view of their definition,

$$\begin{aligned}
\int_0^1 e^{\mu a_m r_m(t)}dt=&\int_0^1 e^{\mu a_m r_0(2^m t)}dt=2^{-m}\int_0^{2^m} e^{\mu a_m r_0(s)}ds=\int_0^1 e^{\mu a_m r_0(s)}ds\\
=&2^{-1}(e^{\mu a_m}+e^{-\mu a_m})=\cosh\mu a_m.
\end{aligned}$$

and for $m<k$

$$\begin{aligned}
&\int_0^1 e^{\mu a_m r_m(t)}e^{\mu a_k r_k(t)}dt=\int_0^1 e^{\mu a_m r_0(2^m t)}e^{\mu a_k r_0(2^k t)}dt\\
=&2^{-m}\int_0^{2^m} e^{\mu a_m r_0(s)}e^{\mu a_k r_0(2^{k-m}s)}ds=\int_0^1 e^{\mu a_m r_0(s)}e^{\mu a_k r_0(2^{k-m}s)}ds\\
=&\int_0^{1/2} e^{\mu a_m}e^{\mu a_k r_0(2^{k-m}s)}ds+\int_{1/2}^1 e^{-\mu a_m}e^{\mu a_k r_0(2^{k-m}s)}ds
\end{aligned}$$

$$=2^{m-k}\left[\int_0^{2^{k-m-1}} e^{\mu a_m} e^{\mu a_k r_0(t)}dt + \int_{2^{k-m-1}}^{2^{k-m}} e^{-\mu a_m} e^{\mu a_k r_0(t)}dt\right]$$
$$=2^{-1}(e^{\mu a_m}+e^{-\mu a_m})\int_0^1 e^{\mu a_k r_0(t)}dt = \int_0^1 e^{\mu a_m r_m(t)}dt \int_0^1 e^{\mu a_k r_k(t)}dt.$$

Thus, by induction, we can verify

$$\int_0^1 e^{\mu \sum_{m=0}^N a_m r_m(t)}dt = \prod_{m=0}^N \int_0^1 e^{\mu a_m r_m(t)}dt.$$

If we now make use of this simple inequality $\cosh x \leqslant e^{x^2}$ (since $\cosh x = \sum_{k=0}^\infty \frac{x^{2k}}{(2k)!} \leqslant \sum_{k=0}^\infty \frac{x^{2k}}{k!} = e^{x^2}$ for $|x| < \infty$ by Taylor expansion), we obtain

$$\int_0^1 e^{\mu F(t)}dt = \prod_{m=0}^N \cosh \mu a_m \leqslant \prod_{m=0}^N e^{\mu^2 a_m^2} = e^{\mu^2 \sum_{m=0}^N a_m^2},$$

with $F(t) = \sum_{m=0}^N a_m r_m(t)$.

Step 2: Let us make the normalizing assumption that $\sum_{n=0}^N a_m^2 = 1$. Then, since $e^{\mu|F|} \leqslant e^{\mu F} + e^{-\mu F}$, we have

$$\int_0^1 e^{\mu|F(t)|}dt \leqslant 2e^{\mu^2}.$$

Recall the distribution function $F_*(\alpha) = m\{t \in [0,1] : |F(t)| > \alpha\}$. If we take $\mu = \alpha/2$ in the above inequality, we have

$$F_*(\alpha) = \int_{|F(t)|>\alpha} dt \leqslant e^{-\frac{\alpha^2}{2}} \int_{|F(t)|>\alpha} e^{\frac{\alpha}{2}|F(t)|}dt \leqslant e^{-\frac{\alpha^2}{2}} 2e^{\frac{\alpha^2}{2^2}} \leqslant 2e^{-\frac{\alpha^2}{4}}.$$

From Theorem 2.16, the above and the formula $\int_0^\infty x^b e^{-ax^2}dx = \Gamma((b+1)/2)/2\sqrt{a^{b+1}}$, it follows immediately that

$$\|F\|_p = \left(p\int_0^\infty \alpha^{p-1}F_*(\alpha)d\alpha\right)^{1/p} \leqslant \left(2p\int_0^\infty \alpha^{p-1}e^{-\frac{\alpha^2}{4}}d\alpha\right)^{1/p} = 2(p\Gamma(p/2))^{1/p},$$

for $1 \leqslant p < \infty$, and so in general

$$\|F\|_p \leqslant A_p \left(\sum_{m=0}^\infty |a_m|^2\right)^{1/2}, \quad 1 \leqslant p < \infty. \tag{6.59}$$

Step 3: We shall now extend the last inequality to several variables. The case of two variables is entirely of the inductive procedure used in the proof of the general case.

We can also limit ourselves to the situation when $p \geqslant 2$, since for the case $p < 2$ the desired inequality is a simple consequence of Hölder's inequality. (Indeed, for $p < 2$ and some $q \geqslant 2$, we have $\|F\|_{L^p(0,1)} \leqslant \|F\|_{L^q(0,1)}\|1\|_{L^{qp/(q-p)}(0,1)} \leqslant \|F\|_{L^q(0,1)}$ by Hölder's inequality.)

We have

$$F(t_1,t_2) = \sum_{m_1=0}^N \sum_{m_2=0}^N a_{m_1m_2} r_{m_1}(t_1) r_{m_2}(t_2) = \sum_{m_1=0}^N F_{m_1}(t_2) r_{m_1}(t_1).$$

By(6.59), it follows

$$\int_0^1 |F(t_1,t_2)|^p dt_1 \leqslant A_p^p \left(\sum_{m_1} |F_{m_1}(t_2)|^2\right)^{p/2}.$$

Integrating this w.r.t. t_2, and using Minkowski's inequlaity with $p/2 \geqslant 1$, we have

$$\int_0^1 \left(\sum_{m_1} |F_{m_1}(t_2)|^2 \right)^{p/2} dt_2 = \left\| \sum_{m_1} |F_{m_1}(t_2)|^2 \right\|_{p/2}^{p/2} \leqslant \left(\sum_{m_1} \| |F_{m_1}(t_2)|^2 \|_{p/2} \right)^{p/2}$$
$$= \left(\sum_{m_1} \|F_{m_1}(t_2)\|_p^2 \right)^{p/2} .$$

However, $F_{m_1}(t_2) = \sum_{m_2} a_{m_1 m_2} r_{m_2}(t_2)$, and therefore the case already proved shows that

$$\|F_{m_1}(t_2)\|_p^2 \leqslant A_p^2 \sum_{m_2} a_{m_1 m_2}^2 .$$

Inserting this in the above gives

$$\int_0^1 \int_0^1 |F(t_1, t_2)|^p dt_1 dt_2 \leqslant A_p^p \left(\sum_{m_1} \sum_{m_2} a_{m_1 m_2}^2 \right)^{p/2} ,$$

which leads to the desired inequality

$$\|F\|_p \leqslant A_p \|F\|_2, \quad 2 \leqslant p < \infty.$$

Step 4: The converse inequality

$$\|F\|_2 \leqslant B_p \|F\|_p, \quad p > 1$$

is a simple consequence of the direct inequality.

In fact, for any $p > 1$, (here we may assume $p < 2$) by Hölder inequality

$$\|F\|_2 \leqslant \|F\|_p^{1/2} \|F\|_{p'}^{1/2} .$$

We already know that $\|F\|_{p'} \leqslant A'_{p'} \|F\|_2$, $p' > 2$. We therefore get

$$\|F\|_2 \leqslant (A'_{p'})^2 \|F\|_p ,$$

which is the required converse inequality. □

Now, let us return to the proof of the Littlewood-Paley square function theorem.

Proof of Theorem 6.26. It will be presented in five steps.

Step 1: We show here that it suffices to prove the inequality

$$\left\| \Big(\sum_{\rho \in \Delta} |S_\rho f(x)|^2 \Big)^{1/2} \right\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{6.60}$$

for $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. To see this sufficiency, let $g \in L^2(\mathbb{R}^n) \cap L^{p'}(\mathbb{R}^n)$, and consider the identity

$$\sum_{\rho \in \Delta} \int_{\mathbb{R}^n} S_\rho f \overline{S_\rho g} dx = \int_{\mathbb{R}^n} f \bar{g} dx$$

which follows from (6.56) by polarization. By Schwarz's inequality and then Hölder's inequality

$$\left| \int_{\mathbb{R}^n} f \bar{g} dx \right| \leqslant \int_{\mathbb{R}^n} \left(\sum_\rho |S_\rho f|^2 \right)^{\frac{1}{2}} \left(\sum_\rho |S_\rho g|^2 \right)^{\frac{1}{2}} dx$$
$$\leqslant \left\| \left(\sum_\rho |S_\rho f|^2 \right)^{\frac{1}{2}} \right\|_p \left\| \left(\sum_\rho |S_\rho g|^2 \right)^{\frac{1}{2}} \right\|_{p'} .$$

Taking the supremum over all such g with the additional restriction that $\|g\|_{p'} \leqslant 1$, gives $\|f\|_p$ for the l.h.s. of the above inequality. The r.h.s. is majorized by

$$A_{p'}\left\|\left(\sum |S_\rho f|^2\right)^{1/2}\right\|_p,$$

since we assume (6.60) for all p. Thus, we have also

$$B_p\|f\|_p \leqslant \left\|\left(\sum_\rho |S_\rho f|^2\right)^{1/2}\right\|_p. \tag{6.61}$$

To dispose of the additional assumption that $f \in L^2$, for $f \in L^p$ take $f_j \in L^2 \cap L^p$ such that $\|f_j - f\|_p \to 0$; use the inequality (6.60) and (6.61) for f_j and $f_j - f_{j'}$; after a simple limiting argument, we get (6.60) and (6.61) for f as well.

Step 2: Here we shall prove the inequality (6.60) for $n = 1$.

We shall need first to introduce a little more notations. We let Δ_1 be the family of dyadic intervals in $\mathbb{R}$, we can enumerate them as $I_0, I_1, \cdots, I_m, \cdots$ (the order is here immaterial). For each $I \in \Delta_1$, we consider the partial sum operator S_I, and a modification of it that we now define. Let $\varphi \in C^1$ be a fixed function with the following properties:

$$\varphi(\xi) = \begin{cases} 1, & 1 \leqslant \xi \leqslant 2, \\ 0, & \xi \leqslant 1/2, \text{ or } \xi \geqslant 4. \end{cases}$$

Suppose I is any dyadic interval, and assume that it is of the form $[2^k, 2^{k+1}]$. Define $\tilde{S}_I$ by

$$\mathscr{F}(\tilde{S}_I f)(\xi) = \varphi(2^{-k}\xi)\hat{f}(\xi) = \varphi_I(\xi)\hat{f}(\xi). \tag{6.62}$$

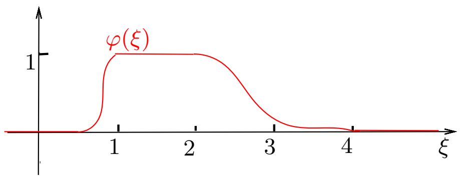


Fig. 6.5 $\varphi(\xi)$

That is, $\tilde{S}_I$, like S_I, is a multiplier transform where the multiplier is equal to one on the interval I; but unlike S_I, the multiplier of $\tilde{S}_I$ is smooth.

A similar definition is made for $\tilde{S}_I$ when $I = [-2^{k+1}, -2^k]$. We observe that

$$S_I\tilde{S}_I = S_I, \tag{6.63}$$

since S_I has as multiplier the characteristic function of I.

Now for each $t \in [0, 1]$, consider the multiplier transform

$$\tilde{T}_t = \sum_{m=0}^{\infty} r_m(t)\tilde{S}_{I_m}.$$

That is, for each t, $\tilde{T}_t$ is the multiplier transform whose multiplier is $m_t(\xi)$, with

$$m_t(\xi) = \sum_{m=0}^{\infty} r_m(t)\varphi_{I_m}(\xi). \tag{6.64}$$

By the definition of φ_{I_m}, it is clear that for any ξ at most three terms in the sum (6.64) can be non-zero. Moreover, we also see easily that

$$|m_t(\xi)| \leqslant B, \quad \left|\frac{dm_t}{d\xi}(\xi)\right| \leqslant \frac{B}{|\xi|}, \tag{6.65}$$

where B is independent of t. Thus, by the Mihlin multiplier theorem (Theorem 6.18)

$$\|\tilde{T}_t f\|_p \leqslant A_p\|f\|_p, \quad \text{for } f \in L^2 \cap L^p, \tag{6.66}$$

and with A_p independent of t. From this, it follows obviously that

$$\left(\int_0^1 \|\tilde{T}_t f\|_p^p dt\right)^{1/p} \leqslant A_p\|f\|_p.$$

However, by Lemma 6.27 about the Rademacher functions,

$$\int_0^1 \|\tilde{T}_t f\|_p^p dt = \int_0^1 \int_{\mathbb{R}^1} \left|\sum r_m(t)(\tilde{S}_{I_m} f)(x)\right|^p dxdt$$

$$\geqslant A'_p \int_{\mathbb{R}^1} \left(\sum_m |\tilde{S}_{I_m} f(x)|^2 \right)^{p/2} dx.$$

Thus, we have

$$\left\| \left(\sum_m |\tilde{S}_{I_m}(f)|^2 \right)^{1/2} \right\|_p \leqslant B_p \|f\|_p. \tag{6.67}$$

Now using (6.63), applying the general theorem about partial sums, Theorem 6.24, with $\Re = \varDelta_1$ here and (6.67), we get, for $F = (\tilde{S}_{I_0} f, \tilde{S}_{I_1} f, \cdots, \tilde{S}_{I_m} f, \cdots)$,

$$\begin{aligned} \left\| \left(\sum_m |S_{I_m}(f)|^2 \right)^{1/2} \right\|_p &= \left\| \left(\sum_m |S_{I_m} \tilde{S}_{I_m}(f)|^2 \right)^{1/2} \right\|_p = \|S_{\varDelta_1} F\|_p \\ \leqslant A_p \|F\|_p = A_p & \left\| \left(\sum_m |\tilde{S}_{I_m}(f)|^2 \right)^{1/2} \right\|_p \leqslant A_p B_p \|f\|_p = C_p \|f\|_p, \end{aligned} \tag{6.68}$$

which is the one-dimensional case of the inequality (6.60), and this is what we had set out to prove.

Step 3: We are still in the one-dimensional case, and we write T_t for the operator

$$T_t = \sum_m r_m(t) S_{I_m}.$$

Our claim is that

$$\|T_t f\|_{L^p_{t,x}} \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{6.69}$$

with A_p independent of t, and $f \in L^2 \cap L^p$.

Write $T_t^N = \sum_{m=0}^N r_m(t) S_{I_m}$, and it suffices to show that (6.69) holds, with T_t^N in place of T_t (and A_p independent of N and t). Since each S_{I_m} is a bounded operator on L^2 and L^p, we have that $T_t^N f \in L^2 \cap L^p$ and so we can apply (6.61) to it, which has already been proved in the case $n = 1$. So

$$B_p \|T_t^N f\|_{L^p_{t,x}} \leqslant \left\| \left(\sum_{m=0}^N |S_{I_m} f|^2 \right)^{1/2} \right\|_p \leqslant C_p \|f\|_p,$$

by using (6.68). Letting $N \to \infty$, we get (6.69).

Step 4: We now turn to the n-dimensional case and define $T_{t_1}^{(1)}$, as the operator T_{t_1} acting only on the x_1 variable. Then, by the inequality (6.69), we get

$$\int_0^1 \int_{\mathbb{R}^1} |T_{t_1}^{(1)} f(x_1, x_2, \cdots, x_n)|^p dx_1 dt_1 \leqslant A_p^p \int_{\mathbb{R}^1} |f(x_1, \cdots, x_n)|^p dx_1, \tag{6.70}$$

for almost every fixed $x_2, x_3, \cdots, x_n$, since $x_1 \to f(x_1, x_2, \cdots, x_n) \in L^2(\mathbb{R}^1) \cap L^p(\mathbb{R}^1)$ for almost every fixed $x_2, \cdots, x_n$, if $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. If we integrate (6.70) w.r.t. $x_2, \cdots, x_n$, we obtain

$$\|T_{t_1}^{(1)} f\|_{L^p_{t_1,x}} \leqslant A_p \|f\|_p, \quad f \in L^2 \cap L^p, \tag{6.71}$$

with A_p independent of t_1. The same inequality of course holds with x_1 replaced by x_2, or x_3, etc.

Step 5: We first describe the additional notation we shall need. With $\varDelta$ representing the collection of dyadic rectangles in $\mathbb{R}^n$, we write any $\rho \in \varDelta$, as $\rho = I_{m_1} \times I_{m_2} \times \cdots \times I_{m_n}$ where $I_0, I_1, \cdots, I_m, \cdots$ represents the arbitrary enumeration of the dyadic intervals used above. Thus if $m = (m_1, m_2, \cdots, m_n)$, with each $m_j \geqslant 0$, we write $\rho_m = I_{m_1} \times I_{m_2} \times \cdots \times I_{m_n}$.

We now apply the operator $T_{t_1}^{(1)}$ for the x_1 variable, and successively its analogues for x_2, x_3, etc. The result is

$$\|T_t f\|_{L^p_{t,x}} \leqslant A_p^n \|f\|_p. \tag{6.72}$$

Here

$$T_t = \sum_{\rho_m \in \Delta} r_m(t) S_{\rho_m}$$

with $r_m(t) = r_{m_1}(t_1)\cdots r_{m_n}(t_n)$ as described in the previous. The inequality holds uniformly for each $(t_1, t_2, \cdots, t_n)$ in the unit cube Q.

We raise this inequality to the p^{th} power and integrate it w.r.t. t, making use of the properties of the Rademacher functions, i.e., Lemma 6.27. We then get, as in the analogous proof of (6.67), that

$$\left\| \left(\sum_{\rho_m \in \Delta} |S_{\rho_m} f|^2 \right)^{1/2} \right\|_p \leqslant A_p \|f\|_p,$$

if $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. This together with the first step concludes the proof of Theorem 6.26. $\square$

6.5 The Marcinkiewicz multiplier theorem

We now present another multiplier theorem which is one of the most important results of the whole theory. For the sake of clarity, we state first the one-dimensional case.

Theorem 6.28. *Let m be a bounded function on $\mathbb{R}^1$, which is of bounded variation on every finite interval not containing the origin. Suppose*

(a) $|m(\xi)| \leqslant B$, $-\infty < \xi < \infty$,

(b) $\int_I |m(\xi)| d\xi \leqslant B$, *for every dyadic interval I.*

Then $m \in M_p$, $1 < p < \infty$; and more precisely, if $f \in L^2 \cap L^p$,

$$\|T_m f\|_p \leqslant A_p \|f\|_p,$$

where A_p depends only on B and p.

To present general theorem, we consider $\mathbb{R}$ as divided into its two half-lines, $\mathbb{R}^2$ as divided into its four quadrants, and generally $\mathbb{R}^n$ as divided into its 2^n "octants". Thus, the first octants in $\mathbb{R}^n$ will be the open "rectangle" of those ξ all of whose coordinates are strictly positive. We shall assume that $m(\xi)$ is defined on each such octant and is there continuous together with its partial derivatives up to and including order n. Thus m may be left undefined on the set of points where one or more coordinate variables vanishes.

For every $k \leqslant n$, we regard $\mathbb{R}^k$ embedded in $\mathbb{R}^n$ in the following obvious way: $\mathbb{R}^k$ is the subspace of all points of the form $(\xi_1, \xi_2, \cdots, \xi_k, 0, \cdots, 0)$.

Theorem 6.29 (Marcinkiewicz' multiplier theorem)**.** *Let m be a bounded function on $\mathbb{R}^n$ that is C^n in all 2^n "octant". Suppose also*

(a) $|m(\xi)| \leqslant B$,

(b) *for each* $0 < k \leqslant n$,

$$\sup_{\xi_{k+1}, \cdots, \xi_n} \int_\rho \left| \frac{\partial^k m}{\partial \xi_1 \partial \xi_2 \cdots \partial \xi_k} \right| d\xi_1 \cdots d\xi_k \leqslant B$$

as ρ ranges over dyadic rectangles of $\mathbb{R}^k$. (If $k = n$, the "sup" sign is omitted.)

(c) *The condition analogous to (b) is valid for every one of the* $n!$ *permutations of the variables* $\xi_1, \xi_2, \cdots, \xi_n$.

Then $m \in M_p$, $1 < p < \infty$*; and more precisely, if* $f \in L^2 \cap L^p$, $\|T_m f\|_p \leqslant A_p \|f\|_p$, *where* A_p *depends only on* B, p *and* n.

Proof. It will be best to prove Theorem 6.29 in the case $n = 2$. This case is already completely typical of the general situation, and in doing only it we can avoid some notational complications.

Let $f \in L^2(\mathbb{R}^2) \cap L^p(\mathbb{R}^2)$, and write $F = T_m f$, that is $\mathscr{F}(F(x)) = m(\xi)\hat{f}(\xi)$.

Let Δ denote the dyadic rectangles, and for each $\rho \in \Delta$, write $f_\rho = S_\rho f$, $F_\rho = S_\rho F$, thus $F_\rho = T_m f_\rho$.

In view of Theorem 6.26, it suffices to show that

$$\Big\| \Big(\sum_{\rho \in \Delta} |F_\rho|^2 \Big)^{1/2} \Big\|_p \leqslant C_p \Big\| \Big(\sum_{\rho \in \Delta} |f_\rho|^2 \Big)^{1/2} \Big\|_p . \tag{6.73}$$

The rectangles in Δ come in four sets, those in the first, the second, the third, and fourth quadrants, respectively. In estimating the l.h.s. of (6.73), consider the rectangles of each quadrant separately, and assume from now on that our rectangles belong to the first quadrant.

We will express F_ρ in terms of an integral involving f_ρ and the partial sum operators. That this is possible is the essential idea of the proof.

Fix ρ and assume $\rho = \{(\xi_1, \xi_2) : 2^k \leqslant \xi_1 \leqslant 2^{k+1}, 2^l \leqslant \xi_2 \leqslant 2^{l+1}\}$. Then, for $(\xi_1, \xi_2) \in \rho$, it is easy to verify the identity

$$\begin{aligned} m(\xi_1, \xi_2) = &\int_{2^l}^{\xi_2} \int_{2^k}^{\xi_1} \frac{\partial^2 m(t_1, t_2)}{\partial t_1 \partial t_2} dt_1 dt_2 + \int_{2^k}^{\xi_1} \frac{\partial}{\partial t_1} m(t_1, 2^l) dt_1 + \int_{2^l}^{\xi_2} \frac{\partial}{\partial t_2} m(2^k, t_2) dt_2 \\ &+ m(2^k, 2^l). \end{aligned}$$

Now let S_t denote the multiplier transform corresponding to the rectangle $\{(\xi_1, \xi_2) : 2^{k+1} > \xi_1 > t_1,\ 2^{l+1} > \xi_2 > t_2\}$. Similarly, let $S_{t_1}^{(1)}$ denote the multiplier corresponding to the interval $2^{k+1} > \xi_1 > t_1$, similarly for $S_{t_2}^{(2)}$. Thus in fact, $S_t = S_{t_1}^{(1)} \cdot S_{t_2}^{(2)}$. Multiplying both sides of the above equation by the function $\chi_\rho \hat{f}$ and taking inverse Fourier transforms yields, by changing the order of integrals in view of Fubini's theorem and the fact that $S_\rho T_m f = F_\rho$, and $S_{t_1}^{(1)} S_\rho = S_{t_1}^{(1)}$, $S_{t_2}^{(2)} S_\rho = S_{t_2}^{(2)}$, $S_t S_\rho = S_t$, we have

$$\begin{aligned} F_\rho = &T_m S_\rho f = \mathscr{F}^{-1} m \chi_\rho \hat{f} \\ = &\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^2} e^{\omega i x \cdot \xi} \Big[\int_{2^l}^{\xi_2} \int_{2^k}^{\xi_1} \frac{\partial^2 m(t_1, t_2)}{\partial t_1 \partial t_2} dt_1 dt_2 \chi_\rho(\xi) \hat{f}(\xi) \Big] d\xi \\ &+ \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^2} e^{\omega i x \cdot \xi} \Big[\int_{2^k}^{\xi_1} \frac{\partial}{\partial t_1} m(t_1, 2^l) dt_1 \chi_\rho(\xi) \hat{f}(\xi) \Big] d\xi \\ &+ \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^2} e^{\omega i x \cdot \xi} \Big[\int_{2^l}^{\xi_2} \frac{\partial}{\partial t_2} m(2^k, t_2) dt_2 \chi_\rho(\xi) \hat{f}(\xi) \Big] d\xi \\ &+ \mathscr{F}^{-1} m(2^k, 2^l) \chi_\rho(\xi) \hat{f}(\xi) \\ = &\left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^2} e^{\omega i x \cdot \xi} \int_{2^l}^{2^{l+1}} \int_{2^k}^{2^{k+1}} \frac{\partial^2 m(t_1, t_2)}{\partial t_1 \partial t_2} \chi_{[2^k, \xi_1]}(t_1) \chi_{[2^l, \xi_2]}(t_2) dt_1 dt_2 \chi_\rho(\xi) \hat{f}(\xi) d\xi \\ &+ \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^2} e^{\omega i x \cdot \xi} \int_{2^k}^{2^{k+1}} \frac{\partial}{\partial t_1} m(t_1, 2^l) \chi_{[2^k, \xi_1]}(t_1) dt_1 \chi_\rho(\xi) \hat{f}(\xi) d\xi \\ &+ \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^2} e^{\omega i x \cdot \xi} \int_{2^l}^{2^{l+1}} \frac{\partial}{\partial t_2} m(2^k, t_2) \chi_{[2^l, \xi_2]}(t_2) dt_2 \chi_\rho(\xi) \hat{f}(\xi) d\xi + m(2^k, 2^l) f_\rho \end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{|\omega|}{2\pi}\right)^n \int_{2^l}^{2^{l+1}}\int_{2^k}^{2^{k+1}}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\chi_{[t_1,2^{k+1}]}(\xi_1)\chi_{[t_2,2^{l+1}]}(\xi_2)\chi_\rho(\xi)\hat f(\xi)d\xi\frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}dt_1dt_2\\
&\quad + \left(\frac{|\omega|}{2\pi}\right)^n \int_{2^k}^{2^{k+1}}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\chi_{[t_1,2^{k+1}]}(\xi_1)\chi_\rho(\xi)\hat f(\xi)d\xi\frac{\partial}{\partial t_1}m(t_1,2^l)dt_1\\
&\quad + \left(\frac{|\omega|}{2\pi}\right)^n \int_{2^l}^{2^{l+1}}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\chi_{[t_2,2^{l+1}]}(\xi_2)\chi_\rho(\xi)\hat f(\xi)d\xi\frac{\partial}{\partial t_2}m(2^k,t_2)dt_2 + m(2^k,2^l)f_\rho\\
&= \int_\rho S_t f_\rho \frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}dt_1dt_2 + \int_{2^k}^{2^{k+1}} S^{(1)}_{t_1} f_\rho\frac{\partial}{\partial t_1}m(t_1,2^l)dt_1\\
&\quad + \int_{2^l}^{2^{l+1}} S^{(2)}_{t_2} f_\rho\frac{\partial}{\partial t_2}m(2^k,t_2)dt_2 + m(2^k,2^l)f_\rho.
\end{aligned}
$$

We apply the Cauchy-Schwarz inequality in the first three terms of the above w.r.t. the measures $|\partial_{t_1}\partial_{t_2}m(t_1,t_2)|dt_1dt_2,\ |\partial_{t_1}m(t_1,2^l)|dt_1,\ |\partial_{t_2}m(2^k,t_2)|dt_2$, respectively, and we use the assumptions of the theorem to deduce

$$
\begin{aligned}
|F_\rho|^2 \lesssim &\Big(\int_\rho |S_t f_\rho|^2\left|\frac{\partial^2 m}{\partial t_1\partial t_2}\right|dt_1dt_2\Big)\Big(\int_\rho\left|\frac{\partial^2 m}{\partial t_1\partial t_2}\right|dt_1dt_2\Big)\\
&+\Big(\int_{2^k}^{2^{k+1}}|S^{(1)}_{t_1}f_\rho|^2\left|\frac{\partial}{\partial t_1}m(t_1,2^l)\right|dt_1\Big)\Big(\int_{2^k}^{2^{k+1}}\left|\frac{\partial}{\partial t_1}m(t_1,2^l)\right|dt_1\Big)\\
&+\Big(\int_{2^l}^{2^{l+1}}|S^{(2)}_{t_2}f_\rho|^2\left|\frac{\partial}{\partial t_2}m(2^k,t_2)\right|dt_2\Big)\Big(\int_{2^l}^{2^{l+1}}\left|\frac{\partial}{\partial t_2}m(2^k,t_2)\right|dt_2\Big)\\
&+|m(2^k,2^l)|^2|f_\rho|^2\\
\leqslant &B'\left\{\int_\rho |S_t f_\rho|^2\left|\frac{\partial^2 m}{\partial t_1\partial t_2}\right|dt_1dt_2 + \int_{I_1}|S^{(1)}_{t_1}f_\rho|^2\left|\frac{\partial m(t_1,2^l)}{\partial t_1}\right|dt_1\right.\\
&\qquad\left.+\int_{I_2}|S^{(2)}_{t_2}f_\rho|^2\left|\frac{\partial m(2^k,t_2)}{\partial t_2}\right|dt_2 + |f_\rho|^2\right\}\\
=&\Im^1_\rho+\Im^2_\rho+\Im^3_\rho+\Im^4_\rho,\ \text{with } \rho = I_1\times I_2.
\end{aligned}
$$

To estimate $\|(\sum_\rho |F_\rho|^2)^{1/2}\|_p$, we estimate separately the contributions of each of the four terms on the r.h.s. of the above inequality by the use of Theorem 6.25. To apply that theorem in the case of $\Im^1_\rho$ we take for Γ the first quadrant, and $d\gamma = |\frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}|dt_1dt_2$, the functions $\gamma\to\rho_\gamma$ are constant on the dyadic rectangles. Since for every rectangle,

$$
\int_\rho d\gamma = \int_\rho\left|\frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}\right|dt_1dt_2 \leqslant B,\ \text{then}\ \left\|\left(\sum_\rho|\Im^1_\rho|\right)^{1/2}\right\|_p \leqslant C_p\left\|\left(\sum_\rho|f_\rho|^2\right)^{1/2}\right\|_p.
$$

Similarly, for $\Im^2_\rho$, $\Im^3_\rho$ and $\Im^4_\rho$, which concludes the proof. □

Chapter 7
Sobolev and Hölder Spaces

7.1 Riesz potentials and fractional integrals

Let f be a sufficiently smooth function which is small at infinity, then the Fourier transform of its Laplacean Δf is

$$\mathscr{F}(-\Delta f)(\xi) = \omega^2 |\xi|^2 \hat{f}(\xi). \tag{7.1}$$

From this, we replace the exponent 2 in $|\xi|^2$ by a general exponent s, and thus to define (at least formally) the fractional power of the Laplacean by

$$(-\Delta)^{s/2} f = \mathscr{F}^{-1}((|\omega||\xi|)^s \hat{f}(\xi)). \tag{7.2}$$

Of special significance will be the negative powers s in the range $-n < s < 0$. In general, with a slight change of notation, we can define

Definition 7.1. Let $s > 0$. The *Riesz potential* of order s is the operator

$$I_s = (-\Delta)^{-s/2}. \tag{7.3}$$

For $0 < s < n$, I_s is actually given in the form

$$I_s f(x) = \frac{1}{\gamma(s)} \int_{\mathbb{R}^n} |x-y|^{-n+s} f(y) dy, \tag{7.4}$$

with

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

The formal manipulations have a precise meaning.

Lemma 7.2. *Let $0 < s < n$.*

(a) *The Fourier transform of the function $|x|^{-n+s}$ is the function $\gamma(s)(|\omega||\xi|)^{-s}$, in the sense that*

$$\int_{\mathbb{R}^n} |x|^{-n+s} \overline{\varphi(x)} dx = \int_{\mathbb{R}^n} \gamma(s)(|\omega||\xi|)^{-s} \overline{\hat{\varphi}(\xi)} d\xi, \tag{7.5}$$

whenever $\varphi \in \mathscr{S}$.

(b) *The identity $\mathscr{F}(I_s f) = (|\omega||\xi|)^{-s} \hat{f}(\xi)$ holds in the sense that*

$$\int_{\mathbb{R}^n} I_s f(x) \overline{g(x)} dx = \int_{\mathbb{R}^n} \hat{f}(\xi)(|\omega||\xi|)^{-s} \overline{\hat{g}(\xi)} d\xi$$

whenever $f, g \in \mathscr{S}$.

Proof. Part (a) is merely a restatement of Lemma 5.14 since $\gamma(s) = |\omega|^s \gamma_{0,s}$.

Part (b) follows immediately from part (a) by writing

$$\begin{aligned}I_s f(x) =&\frac{1}{\gamma(s)}\int_{\mathbb{R}^n} f(x-y)|y|^{-n+s}dy = \int_{\mathbb{R}^n}(|\omega||\xi|)^{-s}\overline{\widehat{\overline{f(x-\cdot)}}}d\xi\\ =&\int_{\mathbb{R}^n}(|\omega||\xi|)^{-s}\hat{f}(\xi)e^{\omega i\xi\cdot x}d\xi = \int_{\mathbb{R}^n}(|\omega||\xi|)^{-s}\hat{f}(\xi)\overline{e^{-\omega i\xi\cdot x}}d\xi,\end{aligned}$$

so

$$\begin{aligned}\int_{\mathbb{R}^n} I_s f(x)\overline{g(x)}dx =& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(|\omega||\xi|)^{-s}\hat{f}(\xi)\overline{e^{-\omega i\xi\cdot x}}d\xi\overline{g(x)}dx\\ =&\int_{\mathbb{R}^n}(|\omega||\xi|)^{-s}\hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi.\end{aligned}$$

This completes the proof. □

Now, we state two further identities which can be obtained from Lemma 7.2 and which reflect essential properties of the potentials I_s.

$$I_s(I_t f) = I_{s+t}f, \quad f\in\mathscr{S},\ s,t>0,\ s+t<n. \tag{7.6}$$

$$\varDelta(I_s f) = I_s(\varDelta f) = -I_{s-2}f, \quad f\in\mathscr{S},\ n\geqslant 3,\ 2\leqslant s\leqslant n. \tag{7.7}$$

The deduction of these two identities have no real difficulties, and these are best left to the interested reader to work out.

A simple consequence of (7.6) is the n-dimensional variant of the beta function,[1]

$$\int_{\mathbb{R}^n}|x-y|^{-n+s}|y|^{-n+t}dy = \frac{\gamma(s)\gamma(t)}{\gamma(s+t)}|x|^{-n+(s+t)} \tag{7.8}$$

with $s,t>0$ and $s+t<n$. Indeed, for any $\varphi\in\mathscr{S}$, we have, by the definition of Riesz potentials and (7.6), that

$$\begin{aligned}&\iint_{\mathbb{R}^n\times\mathbb{R}^n}|x-y|^{-n+s}|y|^{-n+t}dy\varphi(z-x)dx\\ =&\int_{\mathbb{R}^n}|y|^{-n+t}\int_{\mathbb{R}^n}|x-y|^{-n+s}\varphi(z-y-(x-y))dxdy\\ =&\int_{\mathbb{R}^n}|y|^{-n+t}\gamma(s)I_s\varphi(z-y)dy = \gamma(s)\gamma(t)I_t(I_s\varphi)(z) = \gamma(s)\gamma(t)I_{s+t}\varphi(z)\\ =&\frac{\gamma(s)\gamma(t)}{\gamma(s+t)}\int_{\mathbb{R}^n}|x|^{-n+(s+t)}\varphi(z-x)dx.\end{aligned}$$

By the arbitrariness of φ, we have the desired result.

We have considered the Riesz potentials formally and the operation for Schwartz functions. But since the Riesz potentials are integral operators, it is natural to inquire about their actions on the spaces $L^p(\mathbb{R}^n)$.

For this reason, we formulate the following problem. Given $s\in(0,n)$, for what pairs p and q, is the operator $f\to I_s f$ bounded from $L^p(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$? That is, when do we have the inequality

$$\|I_s f\|_q \leqslant A\|f\|_p? \tag{7.9}$$

There is a simple necessary condition, which is merely a reflection of the homogeneity of the kernel $(\gamma(s))^{-1}|y|^{-n+s}$. In fact, we have

Proposition 7.3. *If the inequality (7.9) holds for all $f\in\mathscr{S}$ and a finite constant A, then $1/q = 1/p - s/n$.*

---

[1] The beta function, also called the Euler integral of the first kind, is a special function defined by $B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}dt$ for $\Re x>0$ and $\Re y>0$. It has the relation with $\varGamma$-function: $B(x,y)=\varGamma(x)\varGamma(y)/\varGamma(x+y)$.

Proof. Let us consider the dilation operator δ_ε, defined by $\delta_\varepsilon f(x) = f(\varepsilon x)$ for $\varepsilon > 0$. Then clearly, for $\varepsilon > 0$

$$\begin{aligned}(\delta_{\varepsilon^{-1}} I_s \delta_\varepsilon f)(x) =& \frac{1}{\gamma(s)} \int_{\mathbb{R}^n} |\varepsilon^{-1}x - y|^{-n+s} f(\varepsilon y) dy \\ \overset{z=\varepsilon y}{=\!=}& \varepsilon^{-n} \frac{1}{\gamma(s)} \int_{\mathbb{R}^n} |\varepsilon^{-1}(x-z)|^{-n+s} f(z) dz \\ =& \varepsilon^{-s} I_s f(x).\end{aligned} \tag{7.10}$$

Also

$$\|\delta_\varepsilon f\|_p = \varepsilon^{-n/p} \|f\|_p, \quad \|\delta_{\varepsilon^{-1}} I_s f\|_q = \varepsilon^{n/q} \|I_s f\|_q. \tag{7.11}$$

Thus, by (7.9)

$$\|I_s f\|_q = \varepsilon^s \|\delta_{\varepsilon^{-1}} I_s \delta_\varepsilon f\|_q = \varepsilon^{s+n/q} \|I_s \delta_\varepsilon f\|_q \leqslant A\varepsilon^{s+n/q} \|\delta_\varepsilon f\|_p = A\varepsilon^{s+n/q-n/p} \|f\|_p.$$

If $\|I_s f\|_q \neq 0$, then the above inequality implies

$$1/q = 1/p - s/n. \tag{7.12}$$

If $f \neq 0$ is non-negative, then $I_s f > 0$ everywhere and hence $\|I_s f\|_q > 0$, and we can conclude the desired relations. □

Next, we observe that the inequality must fail at the endpoints $p = 1$ (then $q = n/(n-s)$) and $q = \infty$ (then $p = n/s$).

Let us consider the case $p = 1$. It is not hard to see that the presumed inequality

$$\|I_s f\|_{n/(n-s)} \leqslant A\|f\|_1, \tag{7.13}$$

cannot hold. In fact, we can choose a nice positive function $\varphi \in L^1$ with $\int \varphi = 1$ and a compact support. Then, with $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$, we have that as $\varepsilon \to 0^+$,

$$I_s(\varphi_\varepsilon)(x) \to (\gamma(s))^{-1} |x|^{-n+s}.$$

If $\|I_s \varphi_\varepsilon\|_{n/(n-s)} \leqslant A\|\varphi_\varepsilon\|_1 = A$ were valid uniformly as ε, then Fatou's lemma[2] will imply that

$$\int_{\mathbb{R}^n} |x|^{-n} dx < \infty,$$

and this is a contradiction.

The second atypical case occurs when $q = \infty$. Again the inequality of the type (7.9) cannot hold, and one immediate reason is that this case is dual to the case $p = 1$ just considered. The failure at $q = \infty$ may also be seen directly as follows. Let $f(x) = |x|^{-s}(\ln 1/|x|)^{-(1+\varepsilon)s/n}$, for $|x| \leqslant 1/2$, and $f(x) = 0$, for $|x| > 1/2$, where ε is positive but small. Then $f \in L^{n/s}(\mathbb{R}^n)$, since $\|f\|_{n/s}^{n/s} = \int_{|x|\leqslant 1/2} |x|^{-n}(\ln 1/|x|)^{-1-\varepsilon} dx < \infty$. However, $I_s f$ is essentially unbounded near the origin since

$$I_s f(0) = \frac{1}{\gamma(s)} \int_{|x|\leqslant 1/2} |x|^{-n} (\ln 1/|x|)^{-(1+\varepsilon)s/n} dx = \infty,$$

as long as $(1+\varepsilon)s/n \leqslant 1$.

After these observations, we can formulate the following Hardy-Littlewood-Sobolev theorem of fractional integration. The result was first considered in one dimension on the circle by Hardy and Littlewood. The n-dimensional result was considered by Sobolev.

[2] **Fatou's lamma:** If $\{f_k\}$ is a sequence of nonnegative measurable functions, then

$$\int \liminf_{k\to\infty} f_k d\mu \leqslant \liminf_{k\to\infty} \int f_k d\mu.$$

Theorem 7.4 (Hardy-Littlewood-Sobolev theorem of fractional integrations)**.** *Let* $0 < s < n$, $1 \leqslant p < q < \infty$, $1/q = 1/p - s/n$.

(a) *If* $f \in L^p(\mathbb{R}^n)$, *then the integral* (7.4), *defining* $I_s f$, *converges absolutely for almost every* x.

(b) *If, in addition,* $p > 1$, *then* $\|I_s f\|_q \leqslant A_{p,q}\|f\|_p$.

(c) *If* $f \in L^1(\mathbb{R}^n)$, *then* $m\{x : |I_s f(x)| > \alpha\} \leqslant (A\alpha^{-1}\|f\|_1)^q$, *for all* $\alpha > 0$. *That is, the mapping* $f \to I_s f$ *is of weak type* $(1, q)$, *with* $1/q = 1 - s/n$.

Proof. We first prove parts (a) and (b). Let us write

$$\begin{aligned}\gamma(s)I_s f(x) =& \int_{B(x,\delta)} |x-y|^{-n+s} f(y)dy + \int_{\mathbb{R}^n \setminus B(x,\delta)} |x-y|^{-n+s} f(y)dy \\ =:& L_\delta(x) + H_\delta(x).\end{aligned}$$

Divide the ball $B(x,\delta)$ into the shells $E_j := B(x, 2^{-j}\delta) \setminus B(x, 2^{-(j+1)}\delta)$, $j = 0, 1, 2, ...,$ thus

$$\begin{aligned}|L_\delta(x)| \leqslant& \left|\sum_{j=0}^{\infty} \int_{E_j} |x-y|^{-n+s} f(y)dy\right| \leqslant \sum_{j=0}^{\infty} \int_{E_j} |x-y|^{-n+s}|f(y)|dy \\ \leqslant& \sum_{j=0}^{\infty} \int_{E_j} (2^{-(j+1)}\delta)^{-n+s}|f(y)|dy \leqslant \sum_{j=0}^{\infty} \int_{B(x,2^{-j}\delta)} (2^{-(j+1)}\delta)^{-n+s}|f(y)|dy \\ =& \sum_{j=0}^{\infty} \frac{(2^{-(j+1)}\delta)^{-n+s} m(B(x, 2^{-j}\delta))}{m(B(x, 2^{-j}\delta))} \int_{B(x,2^{-j}\delta)} |f(y)|dy \\ =& \sum_{j=0}^{\infty} \frac{(2^{-(j+1)}\delta)^{-n+s} V_n(2^{-j}\delta)^n}{m(B(x, 2^{-j}\delta))} \int_{B(x,2^{-j}\delta)} |f(y)|dy \\ \leqslant& V_n\delta^s 2^{n-s} \sum_{j=0}^{\infty} 2^{-sj} Mf(x) = \frac{V_n\delta^s 2^n}{2^s - 1} Mf(x).\end{aligned}$$

Now, we derive an estimate for $H_\delta(x)$. By Hölder's inequality and the condition $1/p > s/n$ (i.e., $q < \infty$), we obtain

$$\begin{aligned}|H_\delta(x)| \leqslant& \|f\|_p \left(\int_{\mathbb{R}^n \setminus B(x,\delta)} |x-y|^{(-n+s)p'} dy\right)^{1/p'} \\ =& \|f\|_p \left(\int_{S^{n-1}} \int_\delta^\infty r^{(-n+s)p'} r^{n-1} dr d\sigma\right)^{1/p'} \\ =& \omega_{n-1}^{1/p'} \|f\|_p \left(\int_\delta^\infty r^{(-n+s)p'+n-1} dr\right)^{1/p'} \\ =& \left(\frac{\omega_{n-1}}{(n-s)p' - n}\right)^{1/p'} \delta^{n/p' - (n-s)} \|f\|_p = C(n, s, p)\delta^{s - n/p}\|f\|_p.\end{aligned}$$

By the above two inequalities, we have

$$|\gamma(s)I_s f(x)| \leqslant C(n, s)\delta^s Mf(x) + C(n, s, p)\delta^{s-n/p}\|f\|_p =: F(\delta).$$

Choose $\delta = C(n, s, p)[\|f\|_p/Mf]^{p/n}$, such that the two terms of the r.h.s. of the above are equal, i.e., the minimizer of $F(\delta)$, to get

$$|\gamma(s)I_s f(x)| \leqslant C(Mf)^{1-ps/n}\|f\|_p^{ps/n}.$$

Therefore, by part (i) of Theorem 3.9 for maximal functions, i.e., Mf is finite almost everywhere if $f \in L^p$ $(1 \leqslant p \leqslant \infty)$, it follows that $|I_s f(x)|$ is finite almost everywhere, which proves part (a) of the theorem.

By part (iii) of Theorem 3.9, we know $\|Mf\|_p \leqslant A_p\|f\|_p$ $(1 < p \leqslant \infty)$, thus

$$\|I_s f\|_q \leqslant C\|Mf\|_p^{1-ps/n}\|f\|_p^{ps/n} = C\|f\|_p.$$

This gives the proof of part (b).

Finally, we prove (c). Since we also have $|H_\delta(x)| \leqslant \|f\|_1\delta^{-n+s}$, taking $\alpha = \|f\|_1\delta^{-n+s}$, i.e., $\delta = (\|f\|_1/\alpha)^{1/(n-s)}$, by part (ii) of Theorem 3.9, we get

$$\begin{aligned} &m\{x : |I_s f(x)| > 2(\gamma(s))^{-1}\alpha\} \leqslant m\{x : |L_\delta(x)| > \alpha\} + m\{x : |H_\delta(x)| > \alpha\} \\ \leqslant &m\{x : |C\delta^s Mf(x)| > \alpha\} + 0 \leqslant \frac{C}{\delta^{-s}\alpha}\|f\|_1 = C[\|f\|_1/\alpha]^{n/(n-s)} = C[\|f\|_1/\alpha]^q. \end{aligned}$$

This completes the proof of part (c). □

7.2 Bessel potentials

While the behavior of the kernel $(\gamma(s))^{-1}|x|^{-n+s}$ as $|x| \to 0$ is well suited for their smoothing properties, their decay as $|x| \to \infty$ gets worse as s increases.

We can slightly adjust the Riesz potentials such that we maintain their essential behavior near zero but achieve exponential decay at infinity. The simplest way to achieve this is by replacing the "nonnegative" operator $-\Delta$ by the "strictly positive" operator $I - \Delta$, where I = identity. Here the terms nonnegative and strictly positive, as one may have surmised, refer to the Fourier transforms of these expressions.

Definition 7.5. Let $s > 0$. The *Bessel potential* of order s is the operator

$$J_s = (I - \Delta)^{-s/2}$$

whose action on functions is given by

$$J_s f = \mathscr{F}^{-1}\widehat{G_s}\mathscr{F}f = G_s * f,$$

where

$$G_s(x) = \mathscr{F}^{-1}((1 + \omega^2|\xi|^2)^{-s/2})(x).$$

Now we give some properties of $G_s(x)$ and show why this adjustment yields exponential decay for G_s at infinity.

Proposition 7.6. *Let $s > 0$.*

(a) $G_s(x) = \frac{1}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^\infty e^{-t}e^{-\frac{|x|^2}{4t}}t^{\frac{s-n}{2}}\frac{dt}{t}$.

(b) $G_s(x) > 0$, $\forall x \in \mathbb{R}^n$*; and* $G_s(x) \in L^1(\mathbb{R}^n)$*, precisely,* $\int_{\mathbb{R}^n} G_s(x)dx = 1$.

(c) *There exist two constants* $0 < C(s,n), c(s,n) < \infty$ *such that*

$$G_s(x) \leqslant C(s,n)e^{-|x|/2}, \quad \text{when } |x| \geqslant 2,$$

and such that

$$\frac{1}{c(s,n)} \leqslant \frac{G_s(x)}{H_s(x)} \leqslant c(s,n), \quad \text{when } |x| \leqslant 2,$$

where H_s is a function that satisfies

$$H_s(x) = \begin{cases} |x|^{s-n} + 1 + O(|x|^{s-n+2}), & 0 < s < n, \\ \ln\frac{2}{|x|} + 1 + O(|x|^2), & s = n, \\ 1 + O(|x|^{s-n}), & s > n, \end{cases}$$

as $|x| \to 0$.

(d) $G_s(x) \in L^{p'}(\mathbb{R}^n)$ *for any* $1 \leqslant p \leqslant \infty$ *and* $s > n/p$.

Proof. (a) For $A, s > 0$, we have the Γ-function identity

$$A^{-s/2} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-tA} t^{s/2} \frac{dt}{t},$$

which we use to obtain

$$(1+\omega^2|\xi|^2)^{-s/2} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} e^{-t|\omega\xi|^2} t^{s/2} \frac{dt}{t}.$$

Note that the above integral converges at both ends (as $|\xi| \to 0$, or ∞). Now take the inverse Fourier transform in ξ and use Theorem 1.10 to obtain

$$\begin{aligned} G_s(x) =& \frac{1}{\Gamma(s/2)} \mathscr{F}_\xi^{-1} \int_0^\infty e^{-t} e^{-t|\omega\xi|^2} t^{s/2} \frac{dt}{t} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} \mathscr{F}_\xi^{-1} \left(e^{-t|\omega\xi|^2} \right) t^{s/2} \frac{dt}{t} \\ =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}. \end{aligned}$$

(b) We have easily[3] $\int_{\mathbb{R}^n} G_s(x)dx = \mathscr{F}G_s(0) = 1$. Thus, $G_s \in L^1(\mathbb{R}^n)$.

(c) First, we suppose $|x| \geqslant 2$. Then $t + \frac{|x|^2}{4t} \geqslant t + \frac{1}{t}$ and also $t + \frac{|x|^2}{4t} \geqslant |x|$. This implies that

$$-t - \frac{|x|^2}{4t} \leqslant -\frac{t}{2} - \frac{1}{2t} - \frac{|x|}{2},$$

from which it follows that when $|x| \geqslant 2$

$$G_s(x) \leqslant \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-\frac{t}{2}} e^{-\frac{1}{2t}} t^{\frac{s-n}{2}} \frac{dt}{t} e^{-\frac{|x|}{2}} \leqslant C(s,n) e^{-\frac{|x|}{2}},$$

where $C(s,n) = \frac{2^{|s-n|/2}\Gamma(|s-n|/2)}{(4\pi)^{n/2}\Gamma(s/2)}$ for $s \neq n$, and $C(s,n) = \frac{4}{(4\pi)^{n/2}\Gamma(s/2)}$ for $s = n$ since

$$\begin{aligned} \int_0^\infty e^{-\frac{t}{2}} e^{-\frac{1}{2t}} \frac{dt}{t} \leqslant& \int_0^1 e^{-\frac{1}{2t}} \frac{dt}{t} + \int_1^\infty e^{-\frac{t}{2}} dt = \int_{1/2}^\infty e^{-y} \frac{dy}{y} + 2e^{-1/2} \\ \leqslant& 2\int_{1/2}^\infty e^{-y} dy + 2 \leqslant 4. \end{aligned}$$

Next, suppose that $|x| \leqslant 2$. Write $G_s(x) = G_s^1(x) + G_s^2(x) + G_s^3(x)$, where

$$\begin{aligned} G_s^1(x) =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^{|x|^2} e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}, \\ G_s^2(x) =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_{|x|^2}^4 e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}, \\ G_s^3(x) =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_4^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}. \end{aligned}$$

Since $t|x|^2 \leqslant 16$ in G_s^1, we have $e^{-t|x|^2} = 1 + O(t|x|^2)$ as $|x| \to 0$; thus after changing variables, we can write

---

[3] Or use (a) to show it. From part (a), we know $G_s(x) > 0$. Since $\int_{\mathbb{R}^n} e^{-\pi|x|^2/t} dx = t^{n/2}$, by Fubini's theorem, we have

$$\begin{aligned} \int_{\mathbb{R}^n} G_s(x)dx =& \int_{\mathbb{R}^n} \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t} dx \\ =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} \int_{\mathbb{R}^n} e^{-\frac{|x|^2}{4t}} dx t^{\frac{s-n}{2}} \frac{dt}{t} \\ =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} (4\pi t)^{n/2} t^{\frac{s-n}{2}} \frac{dt}{t} \\ =& \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} t^{\frac{s}{2}-1} dt = 1. \end{aligned}$$

$$\begin{aligned}
G_s^1(x) =&|x|^{s-n}\frac{1}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^1 e^{-t|x|^2}e^{-\frac{1}{4t}}t^{\frac{s-n}{2}}\frac{dt}{t}\\
=&|x|^{s-n}\frac{1}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^1 e^{-\frac{1}{4t}}t^{\frac{s-n}{2}}\frac{dt}{t}+\frac{O(|x|^{s-n+2})}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^1 e^{-\frac{1}{4t}}t^{\frac{s-n}{2}}dt\\
=&\frac{2^{n-s-2}|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)}\int_{1/4}^\infty e^{-y}y^{\frac{s-n}{2}}\frac{dy}{y}+\frac{2^{n-s-4}O(|x|^{s-n+2})}{(4\pi)^{n/2}\Gamma(s/2)}\int_{1/4}^\infty e^{-y}y^{\frac{s-n}{2}}\frac{dy}{y^2}\\
=&c_{s,n}^1|x|^{s-n}+O(|x|^{s-n+2}), \qquad \text{as } |x|\to 0.
\end{aligned}$$

Since $0\leqslant\frac{|x|^2}{4t}\leqslant\frac14$ and $0\leqslant t\leqslant 4$ in G_s^2, we have $e^{-17/4}\leqslant e^{-t-\frac{|x|^2}{4t}}\leqslant 1$, thus as $|x|\to 0$, we obtain

$$G_s^2(x)\sim\int_{|x|^2}^4 t^{(s-n)/2}\frac{dt}{t}=\begin{cases}\frac{|x|^{s-n}}{n-s}-\frac{2^{s-n+1}}{n-s}, & s<n,\\ 2\ln\frac{2}{|x|}, & s=n,\\ \frac{2^{s-n+1}}{s-n}, & s>n.\end{cases}$$

Finally, we have $e^{-1/4}\leqslant e^{-\frac{|x|^2}{4t}}\leqslant 1$ in G_s^3, which yields that $G_s^3(x)$ is bounded above and below by fixed positive constants. Combining the estimates for $G_s^j(x)$, we obtain the desired conclusion.

(d) For $p=1$ and so $p'=\infty$, by part (c), we have $\|G_s(x)\|_\infty\leqslant C$ for $s>n$.

Next, we assume that $1<p\leqslant\infty$ and so $1\leqslant p'<\infty$. Again by part (c), we have, for $|x|\geqslant 2$, that $G_s^{p'}\leqslant Ce^{-p'|x|/2}$, and then the integration over this range $|x|\geqslant 2$ is clearly finite.

On the range $|x|\leqslant 2$, it is clear that $\int_{|x|\leqslant 2}G_s^{p'}(x)dx\leqslant C$ for $s>n$. For the case $s=n$ and $n\neq 1$, we also have $\int_{|x|\leqslant 2}G_s^{p'}(x)dx\leqslant C$ by noticing that

$$\int_{|x|\leqslant 2}\left(\ln\frac{2}{|x|}\right)^q dx=C\int_0^2\left(\ln\frac{2}{r}\right)^q r^{n-1}dr\leqslant C$$

for any $q>0$ since $\lim_{r\to 0}r^\varepsilon\ln(2/r)=0$. For the case $s=n=1$, we have $\int_{|x|\leqslant 2}(\ln\frac{2}{|x|})^q dx=2\int_0^2(\ln 2/r)^q dr=4\int_0^1(\ln 1/r)^q dr=4\Gamma(q+1)$ for $q>0$ by the formula $\int_0^1(\ln 1/x)^{p-1}dx=\Gamma(p)$ for $\Re p>0$. For the final case $s<n$, we have $\int_0^2 r^{(s-n)p'}r^{n-1}dr\leqslant C$ if $(s-n)p'+n>0$, i.e., $s>n/p$.

Thus, we obtain $\|G_s(x)\|_{p'}\leqslant C$ for any $1\leqslant p\leqslant\infty$ and $s>n/p$, which implies the desired result. $\square$

We also have a result analogues to that of Riesz potentials for the operator J_s.

Theorem 7.7. (a) *For all $0<s<\infty$, the operator J_s maps $L^r(\mathbb{R}^n)$ into itself with norm 1 for all $1\leqslant r\leqslant\infty$.*

(b) *Let $0<s<n$ and $1<p<q<\infty$ satisfy $1/q=1/p-s/n$. Then there exists a constant $C_{n,s,p}<\infty$ such that for all $f\in L^p(\mathbb{R}^n)$, we have*

$$\|J_sf\|_q\leqslant C_{n,s,p}\|f\|_p.$$

(c) *If $f\in L^1(\mathbb{R}^n)$, then $m\{x:|J_sf(x)|>\alpha\}\leqslant(C_{n,s}\alpha^{-1}\|f\|_1)^q$, for all $\alpha>0$. That is, the mapping $f\to J_sf$ is of weak type $(1,q)$, with $1/q=1-s/n$.*

Proof. By Young's inequality, we have $\|J_sf\|_r=\|G_s*f\|_r\leqslant\|G_s\|_1\|f\|_r=\|f\|_r$. This proves the result (a).

In the special case $0<s<n$, we have, from the above proposition, that the kernel G_s of J_s satisfies

$$G_s(x)\sim\begin{cases}|x|^{-n+s}, & |x|\leqslant 2,\\ e^{-|x|/2}, & |x|\geqslant 2.\end{cases}$$

Then, we can write

$$\begin{aligned} J_s f(x) \leqslant & C_{n,s} \left[\int_{|y|\leqslant 2} |f(x-y)||y|^{-n+s} dy + \int_{|y|\geqslant 2} |f(x-y)| e^{-|y|/2} dy \right] \\ \leqslant & C_{n,s} \left[I_s(|f|)(x) + \int_{\mathbb{R}^n} |f(x-y)| e^{-|y|/2} dy \right]. \end{aligned}$$

We now use that the function $e^{-|y|/2} \in L^r$ for all $1 \leqslant r \leqslant \infty$, Young's inequality and Theorem 7.4 to complete the proofs of (b) and (c). $\square$

The affinity between the two potentials is given precisely in the following lemma.

Lemma 7.8. *Let $s > 0$.*

(i) *There exists a finite measure μ_s on $\mathbb{R}^n$ such that its Fourier transform $\widehat{\mu_s}$ is given by*

$$\widehat{\mu_s}(\xi) = \frac{|\omega\xi|^s}{(1+|\omega\xi|^2)^{s/2}}.$$

(ii) *There exist a pair of finite measures ν_s and λ_s on $\mathbb{R}^n$ such that*

$$(1+|\omega\xi|^2)^{s/2} = \widehat{\nu_s}(\xi) + |\omega\xi|^s \widehat{\lambda_s}(\xi).$$

Remark 7.9. 1) The first part states in effect that the following formal quotient operator is bounded on every $L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$,

$$\frac{(-\Delta)^{s/2}}{(I-\Delta)^{s/2}}, \quad s > 0. \tag{7.14}$$

2) The second part states also to what extent the same thing is true of the operator inverse to (7.14).

Proof. To prove (i), we use the Taylor expansion

$$(1-t)^{s/2} = 1 + \sum_{m=1}^{\infty} A_{m,s} t^m, \quad |t| < 1, \tag{7.15}$$

where $A_{m,s} = (-1)^m C^m_{s/2} = (-1)^m \frac{\frac{s}{2}(\frac{s}{2}-1)\cdots(\frac{s}{2}-m+1)}{m!} = \frac{(-\frac{s}{2})(1-\frac{s}{2})\cdot(m-\frac{s}{2}-1)}{m!}$. All the $A_{m,s}$ are of same sign for $m > \frac{s}{2}+1$, so $\sum |A_{m,s}| < \infty$, since $(1-t)^{s/2}$ remains bounded as $t \to 1$, if $s \geqslant 0$. Let $t = (1+|\omega\xi|^2)^{-1}$. Then

$$\left(\frac{|\omega\xi|^2}{1+|\omega\xi|^2} \right)^{s/2} = 1 + \sum_{m=1}^{\infty} A_{m,s} (1+|\omega\xi|^2)^{-m}. \tag{7.16}$$

However, $G_{2m}(x) \geqslant 0$ and $\int_{\mathbb{R}^n} G_{2m}(x) e^{-\omega i x\cdot\xi} dx = (1+|\omega\xi|^2)^{-m}$.

We noticed already that $\int G_{2m}(x)dx = 1$ and so $\|G_{2m}\|_1 = 1$.

Thus from the convergence of $\sum |A_{m,s}|$, it follows that if μ_s is defined by

$$\mu_s = \delta_0 + \left(\sum_{m=1}^{\infty} A_{m,s} G_{2m}(x) \right) dx \tag{7.17}$$

with δ_0 the Dirac measure at the origin, then μ_s represents a finite measure. Moreover, by (7.16),

$$\widehat{\mu_s}(\xi) = \frac{|\omega\xi|^s}{(1+|\omega\xi|^2)^{s/2}}. \tag{7.18}$$

For (ii), we now invoke the n-dimensional version of *Wiener's theorem*, to wit: If $\Phi_1 \in L^1(\mathbb{R}^n)$ and $\widehat{\Phi_1}(\xi)+1$ is nowhere zero, then there exists a $\Phi_2 \in L^1(\mathbb{R}^n)$ such that $(\widehat{\Phi_1}(\xi)+1)^{-1} = \widehat{\Phi_2}(\xi)+1$.

For our purposes, we then write

$$\Phi_1(x) = \sum_{m=1}^{\infty} A_{m,s} G_{2m}(x) + G_s(x).$$

Then, by (7.18), we see that

$$\widehat{\Phi_1}(\xi) + 1 = \frac{|\omega\xi|^s + 1}{(1 + |\omega\xi|^2)^{s/2}},$$

which vanishes nowhere. Thus, for an appropriate $\Phi_2 \in L^1$, by Wiener's theorem, we have

$$(1 + |\omega\xi|^2)^{s/2} = (1 + |\omega\xi|^s)[\widehat{\Phi_2}(\xi) + 1],$$

and so we obtain the desired conclusion with $\nu_s = \lambda_s = \delta_0 + \Phi_2(x)dx$. □

7.3 Sobolev spaces

We start by weakening the notation of partial derivatives by the theory of distributions. The appropriate definition is stated in terms of the space $\mathscr{D}(\mathbb{R}^n)$.

Let ∂^α be a differential monomial, whose total order is $|\alpha|$. Suppose we are given two locally integrable functions on $\mathbb{R}^n$, f and g. Then we say that $\partial^\alpha f = g$ (in the weak sense), if

$$\int_{\mathbb{R}^n} f(x)\partial^\alpha \varphi(x)dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} g(x)\varphi(x)dx, \quad \forall \varphi \in \mathscr{D}. \tag{7.19}$$

Integration by parts shows us that this is indeed the relation that we would expect if f had continuous partial derivatives up to order $|\alpha|$, and $\partial^\alpha f = g$ had the usual meaning.

Of course, it is not true that every locally integrable function has partial derivatives in this sense: consider, for example, $f(x) = c^{i/|x|^n}$. However, when the partial derivatives exist, they are determined almost everywhere by the defining relation (7.19).

In this section, we study a quantitative way of measuring smoothness of functions. Sobolev spaces serve exactly this purpose. They measure the smoothness of a given function in terms of the integrability of its derivatives. We begin with the classical definition of Sobolev spaces.

Definition 7.10. Let k be a nonnegative integer and let $1 \leqslant p \leqslant \infty$. The *Sobolev space* $W^{k,p}(\mathbb{R}^n)$ is defined as the space of functions f in $L^p(\mathbb{R}^n)$ all of whose distributional derivatives $\partial^\alpha f$ are also in $L^p(\mathbb{R}^n)$ for all multi-indices α that satisfies $|\alpha| \leqslant k$. This space is normed by the expression

$$\|f\|_{W^{k,p}} = \sum_{|\alpha| \leqslant k} \|\partial^\alpha f\|_p, \tag{7.20}$$

where $\partial^{(0,\dots,0)} f = f$.

The index k indicates the "degree" of smoothness of a given function in $W^{k,p}$. As k increases, the functions become smoother. Equivalently, these spaces form a decreasing sequence

$$L^p \supset W^{1,p} \supset W^{2,p} \supset \cdots$$

meaning that each $W^{k+1,p}(\mathbb{R}^n)$ is a subspace of $W^{k,p}(\mathbb{R}^n)$ in view of the Sobolev norms.

We next observe that *the space* $W^{k,p}(\mathbb{R}^n)$ *is complete*. Indeed, if $\{f_m\}$ is a Cauchy sequence in $W^{k,p}$, then for each α, $\{\partial^\alpha f_m\}$ is a Cauchy sequence in L^p, $|\alpha| \leqslant k$. By the completeness of L^p, there exist functions $f^{(\alpha)}$ such that $f^{(\alpha)} = \lim_m \partial^\alpha f_m$ in L^p, then clearly

$$(-1)^{|\alpha|}\int_{\mathbb{R}^n} f_m \partial^\alpha \varphi dx = \int_{\mathbb{R}^n} \partial^\alpha f_m \varphi dx \to \int_{\mathbb{R}^n} f^{(\alpha)} \varphi dx,$$

for each $\varphi \in \mathscr{D}$. Since the first expression converges to

$$(-1)^{|\alpha|}\int_{\mathbb{R}^n} f \partial^\alpha \varphi dx,$$

it follows that the distributional derivative $\partial^\alpha f$ is $f^{(\alpha)}$. This implies that $f_j \to f$ in $W^{k,p}(\mathbb{R}^n)$ and proves the completeness of this space.

First, we generalize Riesz and Bessel potentials to any $s \in \mathbb{R}$ by

$$\begin{aligned} I^s f =& \mathscr{F}^{-1}|\omega\xi|^s \mathscr{F} f, \quad f \in \mathscr{S}'(\mathbb{R}^n),\ 0 \notin \operatorname{supp} \hat{f}, \\ J^s f =& \mathscr{F}^{-1}(1+|\omega\xi|^2)^{s/2} \mathscr{F} f, \quad f \in \mathscr{S}'(\mathbb{R}^n). \end{aligned}$$

It is clear that $I^{-s} = I_s$ and $J^{-s} = J_s$ for $s > 0$ are exactly Riesz and Bessel potentials, respectively. we also note that $J^s \cdot J^t = J^{s+t}$ for any $s, t \in \mathbb{R}$ from the definition.

Next, we shall extend the spaces $W^{k,p}(\mathbb{R}^n)$ to the case where the number k is real.

Definition 7.11. Let $s \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$. We write

$$\|f\|_{\dot{H}^s_p} = \|I^s f\|_p, \quad \|f\|_{H^s_p} = \|J^s f\|_p.$$

Then, the *homogeneous Sobolev space* $\dot{H}^s_p(\mathbb{R}^n)$ is defined by

$$\dot{H}^s_p(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) :\ \hat{f} \in L^1_{loc}(\mathbb{R}^n), \text{ and } \|f\|_{\dot{H}^s_p} < \infty \right\}, \tag{7.21}$$

The *nonhomogeneous Sobolev space* $H^s_p(\mathbb{R}^n)$ is defined by

$$H^s_p(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) :\ \|f\|_{H^s_p} < \infty \right\}. \tag{7.22}$$

If $p = 2$, we denote $\dot{H}^s_2(\mathbb{R}^n)$ by $\dot{H}^s(\mathbb{R}^n)$ and $H^s_2(\mathbb{R}^n)$ by $H^s(\mathbb{R}^n)$ for simplicity.

It is clear that the space $H^s_p(\mathbb{R}^n)$ is a normed linear space with the above norm. Moreover, it is complete and therefore Banach space. To prove the completeness, let $\{f_m\}$ be a Cauchy sequence in H^s_p. Then, by the completeness of L^p, there exists a $g \in L^p$ such that

$$\|f_m - J^{-s} g\|_{H^s_p} = \|J^s f_m - g\|_p \to 0, \quad \text{as } m \to \infty.$$

Clearly, $J^{-s} g \in \mathscr{S}'$ and thus H^s_p is complete.

We give some elementary results about Sobolev spaces.

Theorem 7.12. *Let $s \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$, then we have*

(a) *$\mathscr{S}$ is dense in H^s_p, $1 \leqslant p < \infty$.*

(b) *$H^{s+\varepsilon}_p \subset H^s_p$, $\forall \varepsilon > 0$.*

(c) *$H^s_p \subset L^\infty$, $\forall s > n/p$.*

(d) *Suppose $1 < p < \infty$ and $s \geqslant 1$. Then $f \in H^s_p(\mathbb{R}^n)$ if and only if $f \in H^{s-1}_p(\mathbb{R}^n)$ and for each j, $\frac{\partial f}{\partial x_j} \in H^{s-1}_p(\mathbb{R}^n)$. Moreover, the two norms are equivalent:*

$$\|f\|_{H^s_p} \sim \|f\|_{H^{s-1}_p} + \sum_{j=1}^n \left\| \frac{\partial f}{\partial x_j} \right\|_{H^{s-1}_p}.$$

(e) *$H^k_p(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$, $1 < p < \infty$, $\forall k \in \mathbb{N}$.*

Proof. (a) Take $f \in H^s_p$, i.e., $J_s f \in L^p$. Since $\mathscr{S}$ is dense in L^p $(1 \leqslant p < \infty)$, there exists a $g \in \mathscr{S}$ such that

$$\|f - J^{-s} g\|_{H^s_p} = \|J^s f - g\|_p$$

is smaller than any given positive number. Since $J^{-s} g \in \mathscr{S}$, therefore $\mathscr{S}$ is dense in H^s_p.

(b) Suppose that $f \in H^{s+\varepsilon}_p$. By part (a) in Theorem 7.7, we see that J_ε maps L^p into L^p with norm 1 for $\varepsilon > 0$. Form this, we get the result since

$$\|f\|_{H^s_p} = \|J^s f\|_p = \|J^{-\varepsilon} J^{s+\varepsilon} f\|_p = \|J_\varepsilon J^{s+\varepsilon} f\|_p \leqslant \|J^{s+\varepsilon} f\|_p = \|f\|_{H^{s+\varepsilon}_p}.$$

(c) By Young's inequality, the definition of the kernel $G_s(x)$ and part (d) of Proposition 7.6, we get for $s > 0$

$$\begin{aligned}\|f\|_\infty =&\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{-s/2}(1+|\omega\xi|^2)^{s/2}\mathscr{F} f\|_\infty = \|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{-s/2} * J^s f\|_\infty \\ \leqslant&\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{-s/2}\|_{p'}\|J^s f\|_p = \|G_s(x)\|_{p'}\|f\|_{H^s_p} \leqslant C\|f\|_{H^s_p}.\end{aligned}$$

(d) From the Mihlin multiplier theorem, we can get $(\omega\xi_j)(1+|\omega\xi|^2)^{-1/2} \in M_p$ for $1 < p < \infty$ (or use part (i) of Lemma 7.8 and properties of Riesz transforms), and thus

$$\begin{aligned}\left\|\frac{\partial f}{\partial x_j}\right\|_{H^{s-1}_p} =&\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{(s-1)/2}(\omega i\xi_j)\mathscr{F} f\|_p \\ =&\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{-1/2}(\omega\xi_j)(1+|\omega\xi|^2)^{s/2}\mathscr{F} f\|_p \\ =&\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{-1/2}(\omega\xi_j) * J^s f\|_p \leqslant C\|J^s f\|_p = C\|f\|_{H^s_p}.\end{aligned}$$

Combining with $\|f\|_{H^{s-1}_p} \leqslant \|f\|_{H^s_p}$, we get

$$\|f\|_{H^{s-1}_p} + \sum_{j=1}^n \left\|\frac{\partial f}{\partial x_j}\right\|_{H^{s-1}_p} \leqslant C\|f\|_{H^s_p}.$$

Now, we prove the converse inequality. We use the Mihlin multiplier theorem once more and an auxiliary function χ on $\mathbb{R}$, infinitely differentiable, non-negative and with $\chi(x) = 1$ for $|x| > 2$ and $\chi(x) = 0$ for $|x| < 1$. We obtain

$$(1+|\omega\xi|^2)^{1/2}(1+\sum_{j=1}^n \chi(\xi_j)|\xi_j|)^{-1} \in M_p, \quad \chi(\xi_j)|\xi_j|\xi_j^{-1} \in M_p, \quad 1 < p < \infty.$$

Thus,

$$\begin{aligned}\|f\|_{H^s_p} =&\|J^s f\|_p = \|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{1/2}\mathscr{F} J^{s-1} f\|_p \\ \leqslant& C\|\mathscr{F}^{-1}(1+\sum_{j=1}^n \chi(\xi_j)|\xi_j|)\mathscr{F} J^{s-1} f\|_p \\ \leqslant& C\|f\|_{H^{s-1}_p} + C\sum_{j=1}^n \|\mathscr{F}^{-1}\chi(\xi_j)|\xi_j|\xi_j^{-1}\mathscr{F} J^{s-1}\frac{\partial f}{\partial x_j}\|_p \\ \leqslant& C\|f\|_{H^{s-1}_p} + \sum_{j=1}^n \left\|\frac{\partial f}{\partial x_j}\right\|_{H^{s-1}_p}.\end{aligned}$$

Thus, we have obtained the desired result.

(e) It is obvious that $W^{0,p} = H^0_p = L^p$ for $k = 0$. However, from part (d), if $k \geqslant 1$, then $f \in H^k_p$ if and only if f and $\frac{\partial f}{\partial x_j} \in H^{k-1}_p$, $j = 1, ..., n$. Thus, we can extends the identity of $W^{k,p} = H^k_p$ from $k = 0$ to $k = 1, 2, ...$ □

We continue with the Sobolev embedding theorem.

Theorem 7.13 (Sobolev embedding theorem)**.** *Let $1 < p \leqslant p_1 < \infty$ and $s, s_1 \in \mathbb{R}$. Assume that $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$. Then the following conclusions hold*

$$H^s_p \subset H^{s_1}_{p_1}, \quad \dot{H}^s_p \subset \dot{H}^{s_1}_{p_1}.$$

Proof. It is trivial for the case $p = p_1$ since we also have $s = s_1$ in this case. Now, we assume that $p < p_1$. Since $\frac{1}{p_1} = \frac{1}{p} - \frac{s-s_1}{n}$, by part (b) of Theorem 7.7, we get

$$\|f\|_{H^{s_1}_{p_1}} = \|J^{s_1} f\|_{p_1} = \|J^{s_1-s} J^s f\|_{p_1} = \|J_{s-s_1} J^s f\|_{p_1} \leqslant C\|J^s f\|_p = C\|f\|_{H^s_p}.$$

Similarly, we can show the homogeneous case. Therefore, we complete the proof. □

Theorem 7.14. *Let $s, \sigma \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$. Then J^σ is an isomorphism between H_p^s and $H_p^{s-\sigma}$.*

Proof. It is clear from the definition. □

Corollary 7.15. *Let $s \in \mathbb{R}$ and $1 \leqslant p < \infty$. Then*

$$(H_p^s)' = H_{p'}^{-s}.$$

Proof. It follows from the above theorem and the fact that $(L^p)' = L^{p'}$, if $1 \leqslant p < \infty$. □

Next we give the connection between the homogeneous and the nonhomogeneous spaces, whose proof will be postponed to next chapter.

Theorem 7.16. *Suppose that $f \in \mathscr{S}'(\mathbb{R}^n)$ and $0 \notin \operatorname{supp} \hat{f}$. Then*

$$f \in \dot{H}_p^s \Leftrightarrow f \in H_p^s, \quad \forall s \in \mathbb{R},\ 1 \leqslant p \leqslant \infty.$$

Moreover, for $1 \leqslant p \leqslant \infty$, we have

$$\begin{aligned} H_p^s =& L^p \cap \dot{H}_p^s, \quad \forall s > 0,\\ H_p^s =& L^p + \dot{H}_p^s, \quad \forall s < 0,\\ H_p^0 =& L^p = \dot{H}_p^0. \end{aligned}$$

7.4 Hölder spaces

Definition 7.17. Let $0 < \alpha < 1$. Define the *Hölder (or Lipschitz) space* C^α as

$$C^\alpha = \{f \in L^\infty(\mathbb{R}^n) : \|f(x-t) - f(x)\|_\infty \leqslant A|t|^\alpha\}.$$

The C^α norm is then given by

$$\|f\|_{C^\alpha} = \|f\|_\infty + \sup_{|t|>0} \frac{\|f(x-t) - f(x)\|_\infty}{|t|^\alpha}. \tag{7.23}$$

The first thing to observe is that the functions in C^α may be taken to be continuous, and so the relation $|f(x-t) - f(x)| \leqslant A|t|^\alpha$ holds for every x. More precisely,

Proposition 7.18. *Every $f \in C^\alpha$ may be modified on a set of measure zero such that it becomes continuous.*

Proof. The proof can be carried out by using the device of regularization. Any smooth regularization will do, and we shall use here that of the Poisson integral. Thus, consider

$$u(x,y) = \int_{\mathbb{R}^n} P_y(t) f(x-t)dt, \quad P_y(t) = \frac{c_n y}{(|t|^2 + y^2)^{(n+1)/2}}, \quad y > 0.$$

Then, since $\int_{\mathbb{R}^n} P_y(t)dt = 1$,

$$u(x,y) - f(x) = \int_{\mathbb{R}^n} P_y(t)[f(x-t) - f(x)]dt,$$

and so, for $0 < \alpha < 1$,

$$\begin{aligned} \|u(x,y) - f(x)\|_\infty \leqslant& \int_{\mathbb{R}^n} P_y(t) \|f(x-t) - f(x)\|_\infty dt \leqslant A c_n y \int_{\mathbb{R}^n} \frac{|t|^\alpha}{(|t|^2 + y^2)^{(n+1)/2}} dt\\ \overset{t=ys}{=\!=}& A c_n y^\alpha \int_{\mathbb{R}^n} \frac{|s|^\alpha}{(|s|^2+1)^{(n+1)/2}} ds = A' y^\alpha. \end{aligned}$$

In particular, $\|u(x,y_1)-u(x,y_2)\|_\infty \to 0$, as y_1 and $y_2 \to 0$, and since $u(x,y)$ is continuous in x, then $u(x,y)$ converges uniformly to $f(x)$ as $y\to 0$. Therefore, $f(x)$ may be taken to be continuous. □

We begin by giving a characterization of $f\in C^\alpha$ in terms of their Poisson integrals $u(x,y)$.

Proposition 7.19. *Suppose $f\in L^\infty(\mathbb{R}^n)$ and $0<\alpha<1$. Then $f\in C^\alpha(\mathbb{R}^n)$ if and only if*

$$\left\|\frac{\partial u(x,y)}{\partial y}\right\|_\infty \leqslant Ay^{-1+\alpha}. \tag{7.24}$$

Remark 7.20. If A_1 is the smallest constant A for which (7.24) holds, then $\|f\|_\infty + A_1$ and $\|f\|_{C^\alpha}$ give equivalent norms.

Proof. For Poisson kernel, we have

$$\begin{aligned}\frac{\partial P_y(x)}{\partial y} =& c_n\frac{(|x|^2+y^2)^{(n+1)/2} - y\frac{n+1}{2}(|x|^2+y^2)^{(n-1)/2}\cdot 2y}{(|x|^2+y^2)^{n+1}}\\ =& c_n\frac{|x|^2+y^2-(n+1)y^2}{(|x|^2+y^2)^{(n+1)/2+1}} = c_n\frac{|x|^2-ny^2}{(|x|^2+y^2)^{(n+1)/2+1}},\end{aligned} \tag{7.25}$$

so

$$\left|\frac{\partial P_y(x)}{\partial y}\right| \leqslant \frac{c}{(|x|^2+y^2)^{(n+1)/2}}, \quad y>0. \tag{7.26}$$

Differentiating $\int_{\mathbb{R}^n} P_y(x)dx = 1$ w.r.t. y, we obtain

$$\int_{\mathbb{R}^n}\frac{\partial P_y(x)}{\partial y}dx = 0, \quad y>0. \tag{7.27}$$

Thus,

$$\frac{\partial u}{\partial y}(x,y) = \int_{\mathbb{R}^n}\frac{\partial P_y(t)}{\partial y}f(x-t)dt = \int_{\mathbb{R}^n}\frac{\partial P_y(t)}{\partial y}[f(x-t)-f(x)]dt.$$

Hence, by changing variables, we have

$$\begin{aligned}\left\|\frac{\partial u(\cdot,y)}{\partial y}\right\|_\infty \leqslant& \|f\|_{C^\alpha}\int_{\mathbb{R}^n}\left|\frac{\partial P_y(t)}{\partial y}\right||t|^\alpha dt \leqslant c\|f\|_{C^\alpha}\int_{\mathbb{R}^n}\frac{|t|^\alpha}{(|t|^2+y^2)^{(n+1)/2}}dt\\ \leqslant& c\|f\|_{C^\alpha}\int_{\mathbb{R}^n}\frac{1}{(|t|^2+y^2)^{(n+1-\alpha)/2}}dt\\ \overset{t=ys}{=\!=\!=}& c\|f\|_{C^\alpha}y^{-1+\alpha}\int_{\mathbb{R}^n}\frac{1}{(|s|^2+1)^{(n+1-\alpha)/2}}ds \leqslant C\|f\|_{C^\alpha}y^{-1+\alpha}.\end{aligned}$$

This proves the necessariness part.

For the sufficiency part, it is far more enlightening, as it reveals an essential feature of the spaces in question, although it is not much more difficult. This insight is contained in the lemma below and the comments that follow. So we shall return to the proof of the second part. (to be continued) □

Lemma 7.21. *Suppose $f\in L^\infty(\mathbb{R}^n)$ and $0<\alpha<1$. Then the single condition (7.24) is equivalent with the n conditions*

$$\left\|\frac{\partial u(x,y)}{\partial x_j}\right\|_\infty \leqslant A'y^{-1+\alpha}, \quad j=1,...,n. \tag{7.28}$$

Remark 7.22. The smallest A in (7.24) is comparable to the smallest A' in (7.28).

Proof. From the Poisson kernel, we can derive

$$\frac{\partial P_y(x)}{\partial x_j} = -\frac{(n+1)c_n y x_j}{(|x|^2+y^2)^{(n+1)/2+1}}, \quad \left|\frac{\partial P_y(x)}{\partial x_j}\right| \leqslant \frac{C}{(|x|^2+y^2)^{(n+1)/2}}, \quad y>0. \tag{7.29}$$

Since we also have $P_y(x) = (\mathscr{F}^{-1}e^{-|\omega\xi|y})(x)$, so for $y = y_1+y_2$, it follows that $P_y = \mathscr{F}^{-1}e^{-|\omega\xi|(y_1+y_2)} = \mathscr{F}^{-1}(e^{-|\omega\xi|y_1}e^{-|\omega\xi|y_1}) = P_{y_1} * P_{y_2}$, with $y_1, y_2 > 0$. Thus,

$$u(x,y) = P_y * f = P_{y_1} * P_{y_2} * f = P_{y_1} * u(x, y_2),$$

and therefore, with $y_1 = y_2 = y/2$, we get

$$\frac{\partial^2 u}{\partial y \partial x_j} = \frac{\partial P_{y/2}}{\partial x_j} * \frac{\partial u(x,y/2)}{2\partial y_2}.$$

By Young's inequality, (7.29) and (7.24), we get

$$\begin{aligned}\left\|\frac{\partial^2 u}{\partial y\partial x_j}\right\|_\infty &\leqslant \left\|\frac{\partial P_{y/2}}{\partial x_j}\right\|_1 \left\|\frac{\partial u(x,y/2)}{2\partial y_2}\right\|_\infty \leqslant C\int_{\mathbb{R}^n}\frac{dx}{(|x|^2+y^2/4)^{(n+1)/2}}\cdot 2^{-\alpha}Ay^{-1+\alpha}\\ &\overset{x=yt/2}{=\!=\!=} CAy^{-2+\alpha}\int_{\mathbb{R}^n}\frac{dt}{(|t|^2+1)^{(n+1)/2}} = A_1 y^{-2+\alpha}.\end{aligned} \tag{7.30}$$

However, by Young's inequality and (7.29),

$$\left\|\frac{\partial}{\partial x_j}u(x,y)\right\|_\infty = \left\|\frac{\partial P_y}{\partial x_j} * f\right\|_\infty \leqslant \left\|\frac{\partial P_y}{\partial x_j}\right\|_1 \|f\|_\infty \leqslant \frac{c\|f\|_\infty}{y}.$$

So

$$\frac{\partial}{\partial x_j}u(x,y) \to 0, \quad \text{as } y\to\infty,$$

and therefore,

$$\frac{\partial}{\partial x_j}u(x,y) = -\int_y^\infty \frac{\partial^2 u(x,y')}{\partial y'\partial x_j}dy'.$$

Then, for $\alpha < 1$, (7.30) gives that

$$\left\|\frac{\partial u}{\partial x_j}\right\|_\infty \leqslant A_1\int_y^\infty y'^{-2+\alpha}dy' \leqslant A_2 y^{-1+\alpha}.$$

Conversely, suppose that (7.28) is satisfied. Reasoning as before, we get that $\|\frac{\partial^2 u}{\partial x_j^2}\|_\infty \leqslant A_3 y^{-2+\alpha}$, $j = 1, ..., n$. However, since u is harmonic, that is because $\frac{\partial^2 u}{\partial y^2} = -\sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2}$, we have $\|\frac{\partial^2 u}{\partial y^2}\|_\infty \leqslant A_4 y^{-2+\alpha}$. Then, a similar integration argument shows that $\|\frac{\partial u}{\partial y}\|_\infty \leqslant A_5 y^{-1+\alpha}$. □

We can now prove the converse part of Proposition 7.19.

Proof of Proposition 7.19 (continue). Suppose $\|\frac{\partial}{\partial y}u(x,y)\|_\infty \leqslant Ay^{-1+\alpha}$. Then Lemma 7.21 also shows that $\|\frac{\partial}{\partial x_j}u(x,y)\|_\infty \leqslant A'y^{-1+\alpha}$. We write

$$f(x+t) - f(x) = [u(x+t,y) - u(x,y)] + [f(x+t) - u(x+t,y)] - [f(x) - u(x,y)].$$

Here y does not necessarily depend on t but it is best to choose $y = |t|$. Now $|u(x+t,y) - u(x,y)| \leqslant \int_L |\nabla_x u(x+s,y)|ds$ where L is the line segment (of length $|t|$) joining x with $x+t$. Thus

$$|u(x+t,y) - u(x,y)| \leqslant |t|\sum_{j=1}^n \|u_{x_j}(x,y)\|_\infty \leqslant C|t||t|^{-1+\alpha} = C|t|^\alpha.$$

Also

$$f(x+t) - u(x+t,y) = -\int_0^y \frac{\partial}{\partial y'}u(x+t,y')dy',$$

and so

$$|f(x+t)-u(x+t,y)| \leqslant \int_0^y \left\|\frac{\partial u}{\partial y'}\right\|_\infty dy' \leqslant Cy^\alpha = C|t|^\alpha.$$

With a similar estimate for $f(x)-u(x,y)$, the proof of the proposition is concluded. □

Similar to Lemma 7.21, we can prove the following lemma, and remaind the proof to interested readers.

Lemma 7.23. *Suppose $f \in L^\infty(\mathbb{R}^n)$, and $\alpha > 0$. Let k and l be two integers, both greater than α. Then the two conditions*

$$\left\|\frac{\partial^k u(x,y)}{\partial y^k}\right\|_\infty \leqslant A_k y^{-k+\alpha}, \quad \text{and} \quad \left\|\frac{\partial^l u(x,y)}{\partial y^l}\right\|_\infty \leqslant A_l y^{-l+\alpha}$$

are equivalent. Moreover, the smallest A_k and A_l holding in the above inequalities are comparable.

The utility of this lemma will be apparent soon.

We now can define the space $C^\alpha(\mathbb{R}^n)$ for any $\alpha > 0$. Suppose that k is the smallest integer greater than α, i.e., the ceiling function of α. We set

$$C^\alpha = \left\{ f \in L^\infty(\mathbb{R}^n) : \left\|\frac{\partial^k}{\partial y^k} u(x,y)\right\|_\infty \leqslant Ay^{-k+\alpha} \right\}. \tag{7.31}$$

If A_k denotes the smallest A appearing in the inequality in (7.31), then we can define the C^α norm by

$$\|f\|_{C^\alpha} = \|f\|_\infty + A_k. \tag{7.32}$$

According to Proposition 7.19, when $0 < \alpha < 1$, this definition is equivalent with the previous one and the resulting norms are also equivalent. Lemma 7.23 also shows us that we could have replaced the $\frac{\partial^k u(x,y)}{\partial y^k}$ by the corresponding estimate for $\frac{\partial^l u(x,y)}{\partial y^l}$ where l is any integer greater than α.

A remark about the condition in (7.31) is in order. The estimate

$$\left\|\frac{\partial^k}{\partial y^k} u(x,y)\right\|_\infty \leqslant Ay^{-k+\alpha}$$

is of interest only for y near zero, since the inequality $\left\|\frac{\partial^k}{\partial y^k} u(x,y)\right\|_\infty \leqslant Ay^{-k}$ (which is stronger away from zero) follows already from the fact that $f \in L^\infty$, (as the argument of Lemma 7.21 shows). This observation allows us to assert the inclusion $C^\alpha \subset C^{\alpha'}$, if $\alpha > \alpha'$.

In the case of $0 < \alpha < 1$, we considered the first order difference, next, we will consider the case $0 < \alpha < 2$, it would be better to use the second order differences. In general, the *m-th order difference operator* $\triangle_t^m$ is defined by

$$\triangle_t^m f(x) = \sum_{k=0}^m C_m^k (-1)^k f(x+kt).$$

Thus, $\triangle_t^2 f(x) = f(x) - 2f(x+t) + f(x+2t)$. But for simplicity, we denote

$$\triangle_t^2 f(x) = f(x-t) - 2f(x) + f(x+t)$$

in this section.

Proposition 7.24. *Suppose $0 < \alpha < 2$. Then $f \in C^\alpha$ if and only if $f \in L^\infty(\mathbb{R}^n)$ and $\|f(x-t)-2f(x)+f(x+t)\|_\infty \leqslant A|t|^\alpha$. The expression*

$$\|f\|_\infty + \sup_{|t|>0} \frac{\|f(x-t)-2f(x)+f(x+t)\|_\infty}{|t|^\alpha}$$

is equivalent with the C^α norm.

Proof. Differentiating $\int_{\mathbb{R}^n} P_y(t)dt = 1$ twice w.r.t. y, we obtain

$$\int_{\mathbb{R}^n} \frac{\partial^2 P_y(t)}{\partial y^2} dt = 0, \quad y > 0. \tag{7.33}$$

From (7.25), we have

$$\frac{\partial^2 P_y(t)}{\partial y^2} = -\frac{c_n(n+1)(3|t|^2 - ny^2)y}{(|t|^2 + y^2)^{(n+5)/2}},$$

and then

$$\frac{\partial^2 P_y(t)}{\partial y^2} = \frac{\partial^2 P_y(-t)}{\partial y^2}, \quad \left|\frac{\partial^2 P_y(t)}{\partial y^2}\right| \leqslant \frac{c}{(|t|^2 + y^2)^{(n+2)/2}}. \tag{7.34}$$

Thus, we get

$$\frac{\partial^2}{\partial y^2} u(x,y) = \frac{1}{2}\int_{\mathbb{R}^n} \frac{\partial^2}{\partial y^2} P_y(t)[f(x-t) - 2f(x) + f(x+t)]dt,$$

and so, for $\alpha < 2$,

$$\begin{aligned}
\left\|\frac{\partial^2}{\partial y^2} u(x,y)\right\|_\infty \leqslant & \frac{Ac}{2}\left[y^{-n-2}\int_{|t|\leqslant y} |t|^\alpha dt + \int_{|t|>y} |t|^{-n-2+\alpha} dt\right] \\
\leqslant & C\left[y^{-n-2}\int_0^y r^{\alpha+n-1} dr + \int_y^\infty r^{-3+\alpha} dr\right] \\
\leqslant & Cy^{-2+\alpha}.
\end{aligned}$$

To prove the converse, we observe that if F has two orders continuous derivatives, then

$$\triangle_t^2 F(x) = \int_0^{|t|}\int_{-s}^s \frac{d^2}{d\tau^2} F(x + t'\tau) d\tau ds, \quad \text{where } t' = t/|t|.$$

It follows immediately that

$$\| \triangle_t^2 F(x)\|_\infty \leqslant |t|^2 \sum_{i,j} \left\|\frac{\partial^2 F}{\partial x_i \partial x_j}\right\|_\infty. \tag{7.35}$$

By the definition (7.31), it is clear that $f \in C^\alpha \Rightarrow f \in C^{\alpha'}$ where $\alpha' < \alpha$. If we choose an $\alpha' < 1$, then by the results in Propositions 7.18 and 7.19, we get

$$\|u(x,y) - f(x)\|_\infty \to 0, \quad \text{and } y\|u_y(x,y)\|_\infty \to 0, \text{ as } y \to 0. \tag{7.36}$$

Thus, the identity

$$f(x) = u(x,0) = \int_0^y y' \frac{\partial^2}{\partial y'^2} u(x,y')dy' - y\frac{\partial u}{\partial y}(x,y) + u(x,y) \tag{7.37}$$

is obtained by noticing that the derivative w.r.t. y of the extreme r.h.s. vanishes, and by the use of the end-point conditions (7.36). However, the arguments of Lemma 7.21 and 7.23 show that the inequality $\|\frac{\partial^2 u(x,y)}{\partial y^2}\|_\infty \leqslant Ay^{-2+\alpha}$ implies the estimates

$$\left\|\frac{\partial^2 u(x,y)}{\partial x_i \partial x_j}\right\|_\infty \leqslant A'y^{-2+\alpha}, \quad \left\|\frac{\partial^3 u(x,y)}{\partial y \partial x_i \partial x_j}\right\|_\infty \leqslant A'y^{-3+\alpha}.$$

Thus, by using(7.35) to the last two terms of the r.h.s. of (7.37),

$$\begin{aligned}
\| \triangle_t^2 f\|_\infty \leqslant & \| \triangle_t^2 \int_0^y y' \frac{\partial^2}{\partial y'^2} u(x,y')dy'\|_\infty + y\| \triangle_t^2 \frac{\partial u}{\partial y}(x,y)\|_\infty + \| \triangle_t^2 u(x,y)\|_\infty \\
\leqslant & 4\int_0^y y' \|\frac{\partial^2}{\partial y'^2} u(x,y')\|_\infty dy' + |t|^2 \sum_{i,j}\left[y\left\|\frac{\partial^3 u}{\partial y \partial x_i \partial x_j}\right\|_\infty + \left\|\frac{\partial^2 u}{\partial x_i \partial x_j}\right\|_\infty\right] \\
\leqslant & C\int_0^y y' y'^{-2+\alpha} dy' + C|t|^2[yy^{-3+\alpha} + y^{-2+\alpha}] \leqslant Cy^\alpha + C|t|^2 y^{-2+\alpha}.
\end{aligned}$$

Taking $y = |t|$ gives

$$\|\Delta_t^2 f\|_\infty \leqslant C|t|^\alpha, \qquad \text{if } \alpha > 0,$$

which is the desired result. □

Proposition 7.25. *Suppose $\alpha > 1$. Then $f \in C^\alpha$ if and only if $f \in L^\infty$ and $\frac{\partial f}{\partial x_j} \in C^{\alpha-1}$, $j = 1, ..., n$. The norms $\|f\|_{C^\alpha}$ and $\|f\|_\infty + \sum_{j=1}^n \|\frac{\partial f}{\partial x_j}\|_{C^{\alpha-1}}$ are equivalent.*

Proof. Let us suppose for simplicity that $1 < \alpha \leqslant 2$, the other cases can be argued similarly.

We first prove that $\frac{\partial f}{\partial x_j} \in L^\infty$. We have $\|\frac{\partial^3 u}{\partial y^3}\|_\infty \leqslant Ay^{-3+\alpha}$ since $f \in C^\alpha$, which implies, as we know, $\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\|_\infty \leqslant Ay^{-3+\alpha}$. Equivalently, we see that $\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\|_\infty \leqslant Ay^{-1-\beta}$, where $0 \leqslant \beta < 1$ since $1 < \alpha \leqslant 2$. We restrict to $0 < y \leqslant 1$, then an integration in y gives

$$\int_y^1 \frac{\partial^3 u}{\partial y^2 \partial x_j} dy = \left[\frac{\partial^2 u}{\partial y \partial x_j}\right]_{y=1} - \frac{\partial^2 u}{\partial y \partial x_j},$$

and then

$$\left\|\frac{\partial^2 u}{\partial y \partial x_j}\right\|_\infty \leqslant \int_y^1 \left\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\right\|_\infty dy + \left\|\left[\frac{\partial^2 u}{\partial y \partial x_j}\right]_{y=1}\right\|_\infty \leqslant Cy^{-\beta} + C.$$

Another integration,

$$\int_{y_1}^{y_2} \frac{\partial^2 u}{\partial y \partial x_j} dy = \frac{\partial}{\partial x_j} u(x, y_2) - \frac{\partial}{\partial x_j} u(x, y_1),$$

then shows that

$$\left\|\frac{\partial}{\partial x_j} u(x, y_2) - \frac{\partial}{\partial x_j} u(x, y_1)\right\|_\infty \leqslant \int_{y_1}^{y_2} [Cy^{-\beta} + C] dy = C(y_2^{1-\beta} - y_1^{1-\beta}) + C(y_2 - y_1).$$

Thus, $\{\frac{\partial}{\partial x_j} u(x, y)\}$ is Cauchy in the L^∞ norm as $y \to 0$, and so its limit can be taken to be $\frac{\partial f}{\partial x_j}$. The argument also gives the bound

$$\left\|\frac{\partial f}{\partial x_j}\right\|_\infty \leqslant \left\|\frac{\partial u}{\partial x_j}\right\|_\infty \leqslant CAy^{-1+\alpha} \leqslant CA \leqslant C\|f\|_{C^\alpha},$$

since $0 < y \leqslant 1$ and $\alpha > 1$.

Since the (weak) derivative of f is $\frac{\partial f}{\partial x_j}$, the Poisson integral of the latter is $\frac{\partial u}{\partial x_j}$. But $\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\|_\infty \leqslant Ay^{-3+\alpha}$. Therefore, $\frac{\partial f}{\partial x_j} \in C^{\alpha-1}$. The converse implication is proved in the same way. □

The last proposition reduces the study of the spaces C^α to those α such that $0 < \alpha \leqslant 1$. Concerning the space C^α, $0 < \alpha \leqslant 1$, the following additional remark is in order.

Remark 7.26. When $0 < \alpha < 1$, Proposition 7.24 shows that if $f \in L^\infty$, the two conditions $\|f(x+t) - f(x)\|_\infty \leqslant A|t|^\alpha$ and $\|f(x-t) - 2f(x) + f(x+t)\|_\infty \leqslant A'|t|^\alpha$ are equivalent. However, this is not the case when $\alpha = 1$.

Example 7.27. There exists $f \in L^\infty(\mathbb{R}^n)$ such that

$$\|f(x-t) - 2f(x) + f(x+t)\|_\infty \leqslant A|t|, \quad |t| > 0,$$

but $\|f(x+t) - f(x)\|_\infty \leqslant A'|t|$ fails for all A'.

Solution. One can construct such f by lacunary series, and more particularly as *Hardy-Weierstrass non-differentiable functions.*[4] To do this, we consider the function of one vari-

[4] In mathematics, the *Weierstrass function* is a pathological example of a real-valued function on the real line. The function has the property that it is continuous everywhere but differentiable nowhere. It is named after its discoverer Karl Weierstrass.

Historically, the Weierstrass function is important because it was the first published (1872) to challenge the notion that every continuous function was differentiable except on a set of isolated points.

able x, given by $f(x) = \sum_{k=1}^{\infty} a^{-k} e^{2\pi i a^k x}$. Here $a > 1$, for simplicity, we take a to be an integer and this makes f periodic.[5] Now

$$f(x-t) - 2f(x) + f(x+t) = \sum_{k=1}^{\infty} a^{-k}[e^{2\pi i a^k(x-t)} - 2e^{2\pi i a^k x} + e^{2\pi i a^k(x+t)}]$$

$$= \sum_{k=1}^{\infty} a^{-k}[e^{-2\pi i a^k t} - 2 + e^{2\pi i a^k t}]e^{2\pi i a^k x} = 2\sum_{k=1}^{\infty} a^{-k}[\cos 2\pi a^k t - 1]e^{2\pi i a^k x}.$$

Therefore, (assume $|t| < 1$ without loss of generalities)

$$\|f(x-t) - 2f(x) + f(x+t)\|_\infty \leqslant 2\sum_{a^k|t|\leqslant 1} a^{-k}B(a^k t)^2 + 4\sum_{a^k|t|>1} a^{-k}$$

$$\leqslant 2B|t|^2 \sum_{k\leqslant[\log_a |t|^{-1}]} a^k + 4\sum_{k>[\log_a |t|^{-1}]} a^{-k} \leqslant 2B|t|^2 \frac{a^{[\log_a |t|^{-1}]+1} - 1}{a-1} + 4\frac{a^{-[\log_a |t|^{-1}]-1}}{1-a^{-1}}$$

$$\leqslant A|t|.$$

We have used merely the fact that $|\cos 2\pi a^k t - 1| \leqslant \min(B(a^k t)^2, 2)$ with $B = 2\pi^2$ since $\sin x \leqslant x$ for any $x \geqslant 0$.

However, if we had $\|f(x+t) - f(x)\|_\infty \leqslant A'|t|$, then by Bessel's inequality[6] for L^2 periodic functions we would get

$$(A'|t|)^2 \geqslant \int_0^1 |f(x+t) - f(x)|^2 dx = \int_0^1 \left|\sum_{k=1}^{\infty} a^{-k}[e^{2\pi i a^k t} - 1]e^{2\pi i a^k x}\right|^2 dx$$

$$\geqslant \sum_{k=1}^{\infty} a^{-2k}|e^{2\pi i a^k t} - 1|^2 \geqslant \sum_{a^k|t|<1/2} a^{-2k}|e^{2\pi i a^k t} - 1|^2.$$

In the range $a^k|t| < 1/2$, we have $|e^{2\pi i a^k t} - 1|^2 > 16(a^k t)^2$ due to the inequality $\frac{\sin x}{x} > \frac{2}{\pi}$ for any $x \in (-\pi/2, \pi/2)$, and so we would arrive at the contradiction

$$(A'|t|)^2 \geqslant 16|t|^2 \sum_{a^k|t|<1/2} 1 \geqslant 16|t|^2 \sum_{1\leqslant k\leqslant[\log_a |t|^{-1}/2]-1} 1,$$

which implies that $A'^2 \geqslant 16\left[\log_a |t|^{-1}/2\right] - 16$ and so $A' \to \infty$ as $|t| \to 0$. □

---

[5] The result also holds if a is non-integral.

[6] Let $S = \{x_\alpha : \alpha \in A\}$ be a complete orthonormal system of a Hilbert space X. For any $f \in X$, we define its Fourier coefficients $f_\alpha = (f, x_\alpha)$. Then we have *Bessel's inequality* $\sum_{\alpha\in A} |f_\alpha|^2 \leqslant \|f\|_2$. Moreover, we have *Parseval's relation* $\|f\|^2 = \sum_{\alpha\in A} |f_\alpha|^2$.

Chapter 8
Besov and Triebel-Lizorkin Spaces

8.1 The dyadic decomposition: the smooth version

In this section, we will introduce another Littlewood-Paley dyadic decomposition, which is also a very basic way to carve up the phase space.

The dyadic decomposition with rectangles is very intuitionistic for the statement, but it is not convenient to do some operations such as differentiation, multiplier and so on. Therefore, we use a smooth form of this decomposition. Throughout, we shall call a *ball* any set $\{\xi \in \mathbb{R}^n : |\xi| \leqslant R\}$ with $R > 0$ and an *annulus* any set $\{\xi \in \mathbb{R}^n : R_1 \leqslant |\xi| \leqslant R_2\}$ with $0 < R_1 < R_2$.

Let $\alpha \in (1, \sqrt{2})$ and $\psi : \mathbb{R}^n \to [0,1]$ be a real radial smooth bump function, e.g.

$$\psi(\xi) = \begin{cases} 1, & |\xi| \leqslant \alpha^{-1}, \\ \text{smooth}, & \alpha^{-1} < |\xi| < \alpha, \\ 0, & |\xi| \geqslant \alpha. \end{cases} \tag{8.1}$$

Let $\varphi(\xi)$ be the function

$$\varphi(\xi) := \psi(\xi/2) - \psi(\xi). \tag{8.2}$$

Thus, φ is a bump function supported on the annulus

$$\mathbb{A} = \left\{\xi : \alpha^{-1} \leqslant |\xi| \leqslant 2\alpha\right\}. \tag{8.3}$$

By construction, we have

$$\sum_{k\in\mathbb{Z}} \varphi(2^{-k}\xi) = 1$$

for all $\xi \neq 0$. Thus, we can partition unity into the functions $\varphi(2^{-k}\xi)$ for integers k, each of which is supported on an annulus of the form $|\xi| \sim 2^k$.

For convenience, we define the following functions

$$\begin{cases} \psi_k(\xi) = \psi(2^{-k}\xi), \quad k \in \mathbb{Z}, \\ \varphi_k(\xi) = \varphi(2^{-k}\xi) = \psi_{k+1}(\xi) - \psi_k(\xi), \quad k \in \mathbb{Z}. \end{cases} \tag{8.4}$$

Since $\operatorname{supp}\varphi \subset \mathbb{A}$, we have

$$\begin{aligned} &\operatorname{supp}\varphi_k \subset 2^k\mathbb{A} := \left\{\xi : 2^k\alpha^{-1} \leqslant |\xi| \leqslant 2^{k+1}\alpha\right\}, \quad k \in \mathbb{Z}, \\ &\operatorname{supp}\psi_k \subset \left\{\xi : |\xi| \leqslant 2^k\alpha\right\}, \quad k \in \mathbb{Z}. \end{aligned} \tag{8.5}$$

We now define the *k-th homogeneous dyadic blocks* $\dot\triangle_k$ and the *homogeneous low-frequency cut-off operators* $\dot S_k$ by

$$\dot\triangle_k f = \mathscr{F}^{-1}\varphi_k\mathscr{F} f, \quad \dot S_k f = \mathscr{F}^{-1}\psi_k\mathscr{F} f = \sum_{j\leqslant k-1} \dot\triangle_j f, \quad k \in \mathbb{Z}. \tag{8.6}$$

Informally, $\dot{\triangle}_k$ is a frequency projection[1] to the annulus $\{\xi : 2^k \alpha^{-1} \leqslant |\xi| \leqslant 2^{k+1}\alpha\}$, while $\dot{S}_k$ is a frequency projection to the ball $\{\xi : |\xi| \leqslant 2^k \alpha\}$. The nonhomogeneous dyadic blocks $\triangle_k$ are defined by

$$\triangle_k f = 0 \text{ if } k \leqslant -2, \quad \triangle_{-1} f = \dot{S}_0 f, \quad \text{and } \triangle_k f = \dot{\triangle}_k f \text{ if } k \geqslant 0.$$

The nonhomogeneous low-frequency cut-off operator S_k is defined by

$$S_k f = \sum_{j \leqslant k-1} \triangle_j f.$$

Obviously, $S_k f = 0$ if $k \leqslant -1$, and $S_k f = \dot{S}_k f$ if $k \geqslant 0$.

Observe that $\dot{S}_{k+1} = \dot{S}_k + \dot{\triangle}_k$ from (8.4). Also, if f is an L^2 function, then $\dot{S}_k f \to 0$ in L^2 as $k \to -\infty$, and $\dot{S}_k f \to f$ in L^2 as $k \to +\infty$ (this is an easy consequence of Parseval's theorem). By telescoping the series, we thus can write the following (formal) Littlewood-Paley (or dyadic) decomposition[2]

$$\mathrm{Id} = \sum_{k \in \mathbb{Z}} \dot{\triangle}_k \quad \text{and} \quad \mathrm{Id} = \sum_{k \in \mathbb{Z}} \triangle_k.$$

The homogeneous decomposition takes a single function and writes it as a superposition of a countably infinite family of functions $\dot{\triangle}_k f$, each one of which has frequency of magnitude roughly 2^k. Lower values of k represent low frequency components of f; higher values represent high frequency components.

Both decompositions have advantages and drawbacks. The nonhomogeneous one is more suitable for characterizing the usual functional spaces whereas the properties of invariance by dilation of the homogeneous decomposition may be more adapted for studying certain PDEs or stating optimal functional inequalities having some scaling invariance.

In the nonhomogeneous cases, the above decomposition makes sense in $\mathscr{S}'(\mathbb{R}^n)$.

Proposition 8.1. *Let $f \in \mathscr{S}'(\mathbb{R}^n)$, then $f = \lim_{k \to +\infty} S_k f$ in $\mathscr{S}'(\mathbb{R}^n)$.*

Proof. Note that $\langle f - S_k f, g\rangle = \langle f, g - S_k g\rangle$ for all $f \in \mathscr{S}'(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$, so it suffices to prove that $g = \lim_{k \to +\infty} S_k g$ in $\mathscr{S}(\mathbb{R}^n)$. Because the Fourier transform is an automorphism of $\mathscr{S}(\mathbb{R}^n)$, we can alternatively prove that $\psi(2^{-k}\cdot)\hat{g}$ tends to $\hat{g}$ in $\mathscr{S}(\mathbb{R}^n)$. This can easily be verified, so we left it to the interested reader. $\square$

For the operators $\dot{\triangle}_k$ and $\dot{S}_k$, we can easily verify the following result:

Proposition 8.2. *Let $\alpha \in (1, \sqrt{2})$, $k, l \in \mathbb{Z}$, and $\dot{\triangle}_k, \dot{S}_k$ be defined as in (8.6). For any $f \in \mathscr{S}'(\mathbb{R}^n)$ and $g \in \mathscr{S}'(\mathbb{R}^n)$, we have the following properties:*

$$\begin{aligned} \dot{S}_k \dot{\triangle}_{k+l} f &\equiv 0, \quad \text{if } l \geqslant 1, \\ \dot{\triangle}_k \dot{\triangle}_l f &\equiv 0, \quad \text{if } |k-l| \geqslant 2, \\ \dot{\triangle}_k (\dot{S}_{l-1} f \dot{\triangle}_l g) &\equiv 0, \quad \text{if } l - k \geqslant 1 + \log_2 \frac{2\alpha^2}{2-\alpha^2}, \text{ or } k - l \geqslant -1 + \log_2 5\alpha^2. \end{aligned}$$

[1] Strictly speaking, these are not quite projections, even though they are self-adjoint. They do not quite square to themselves because we choose ψ to be a smooth cut-off rather than a rough one. However, the operator $\dot{\triangle}_k \dot{\triangle}_k$ is of the same form as $\dot{\triangle}_k$, and similarly for $\dot{S}_k$, and so it is still quite reasonable to think of these operators as (smoothed out) projection operators.

[2] Actually, this decomposition works for just about any locally integrable function which has some decay at infinity, and one usually has all the convergence properties of the summation that one needs. In many applications one can make the *a priori* assumption that f is Schwartz, in which case the convergence is uniform. However, if f does not decay, the this formula fails. For instance, if $f \equiv 1$, then all the projections $\dot{\triangle}_k f$ vanish because $\dot{\triangle}_k 1 = (|\omega|/(2\pi))^n \int e^{\omega i x \xi} \varphi_k(\xi)(|\omega|/(2\pi))^{-n}\delta(\xi)d\xi = \varphi_k(0) = \varphi(0) = 0$.

Remark 8.3. In these properties, we need the condition $\alpha^2 < 2$ which is the reason that we requires $\alpha < \sqrt{2}$ in the beginning of the section.

When dealing with the Littlewood-Paley decomposition, it is convenient to introduce the functions

$$\tilde{\psi}(\xi) = \psi(\xi/2), \quad \tilde{\varphi}(\xi) = \varphi_{-1}(\xi) + \varphi_0(\xi) + \varphi_1(\xi) = \psi(\xi/4) - \psi(4\xi).$$

as well as the operators

$$\tilde{S}_k = \mathscr{F}^{-1}\tilde{\psi}(2^{-k}\xi)\mathscr{F} = \dot{S}_{k+1}, \quad \tilde{\triangle}_k = \mathscr{F}^{-1}\tilde{\varphi}(2^{-k}\xi)\mathscr{F}.$$

It is clear that $\dot{S}_k = \tilde{S}_k\dot{S}_k$ due to $\alpha < \sqrt{2}$, and $\dot{\triangle}_k = \tilde{\triangle}_k\dot{\triangle}_k$ from Proposition 8.2.

Now, we give the Bernstein inequalities:

Proposition 8.4 (Bernstein inequalities)**.** *Let $k \in \mathbb{N}_0$, $1 \leqslant p \leqslant q \leqslant \infty$, $\mathbb{A}$ be an annulus and B be a ball. Then, for any $f \in L^p(\mathbb{R}^n)$, we have*

$$\operatorname{supp}\hat{f} \subset \lambda B \quad \Longrightarrow \quad \|D^k f\|_q := \sup_{|\alpha|=k} \|\partial^\alpha f\|_q \leqslant C^{k+1}\lambda^{k+n(\frac{1}{p}-\frac{1}{q})}\|f\|_p,$$

$$\operatorname{supp}\hat{f} \subset \lambda \mathbb{A} \quad \Longrightarrow \quad C^{-k-1}\lambda^k\|f\|_p \leqslant \|D^k f\|_p \leqslant C^{k+1}\lambda^k\|f\|_p.$$

Proof. Let ϕ be a function of $\mathscr{D}(\mathbb{R}^n)$ with value 1 near B and denote $\phi_\lambda(\xi) = \phi(\xi/\lambda)$. As $\hat{f}(\xi) = \phi_\lambda(\xi)\hat{f}(\xi)$, we have

$$\partial^\alpha f = \partial^\alpha g_\lambda * f \quad \text{with} \quad g_\lambda = \mathscr{F}^{-1}\phi_\lambda.$$

Thus, $g_\lambda(x) = \lambda^n(\mathscr{F}^{-1}\phi)(\lambda x) = \lambda^n g(\lambda x)$ in view of Proposition 1.3, where we denote $g := g_1$.

Applying Young's inequality with $\frac{1}{r} := 1 - \frac{1}{p} + \frac{1}{q}$, we get

$$\begin{aligned}
\|\partial^\alpha f\|_q =& \|\partial^\alpha g_\lambda * f\|_q \leqslant \|\partial^\alpha g_\lambda\|_r\|f\|_p \\
=& \lambda^{n+k}\|(\partial^\alpha g)(\lambda x)\|_r\|f\|_p = \lambda^{k+n/r'}\|\partial^\alpha g\|_r\|f\|_p \\
=& \lambda^{k+n(\frac{1}{p}-\frac{1}{q})}\|\partial^\alpha g\|_r\|f\|_p.
\end{aligned}$$

The first assertion follows via

$$\begin{aligned}
\|\partial^\alpha g\|_r \leqslant& \|\partial^\alpha g\|_\infty + \|\partial^\alpha g\|_1 \\
\leqslant& \|\partial^\alpha g\|_\infty + \int_{\mathbb{R}^n} |\partial^\alpha g|(1+|\omega x|^2)^n \frac{1}{(1+|\omega x|^2)^n} dx \\
\leqslant& \|\partial^\alpha g\|_\infty + \|(1+|\omega x|^2)^n \partial^\alpha g\|_\infty \int_{\mathbb{R}^n} \frac{1}{(1+|\omega x|^2)^n} dx \\
\leqslant& C_n\|(1+|\omega x|^2)^n\partial^\alpha g\|_\infty = C_n\|\mathscr{F}^{-1}\mathscr{F}((1+|\omega x|^2)^n\partial^\alpha g)\|_\infty \\
\leqslant& C_n\|\mathscr{F}((1+|\omega x|^2)^n\partial^\alpha g)\|_1 = C_n\|(1-\Delta)^n((\omega i\xi)^\alpha\phi(\xi))\|_1 \\
=& C_n\left\|\sum_{j=0}^n C_n^j(-1)^j\Delta^j(\xi^\alpha\phi(\xi))\right\|_1 \leqslant C_n\sum_{j=0}^n C_n^j\|\Delta^j(\xi^\alpha\phi(\xi))\|_1 \\
\leqslant& C_n \sup_{0\leqslant|\beta|\leqslant|\alpha|,\, 0\leqslant|\sigma|\leqslant 2n-|\beta|} \|\partial^\beta(\xi^\alpha)\partial^\sigma\phi\|_1 \\
\leqslant& C_n \sup_{0\leqslant|\beta|\leqslant|\alpha|,\, 0\leqslant|\sigma|\leqslant 2n-|\beta|} \|\xi^\beta\partial^\sigma\phi\|_1 \\
\leqslant& C_n C^k \sup_{0\leqslant|\sigma|\leqslant 2n} \|\partial^\sigma\phi\|_1 \quad \text{(since ϕ is compactly supported)} \\
\leqslant& C_n^{k+1}.
\end{aligned}$$

To prove the second assertion, we consider a function $\tilde{\phi} \in \mathscr{D}(\mathbb{R}^n \setminus \{0\})$ with value 1 on a neighborhood of $\mathbb{A}$. From the algebraic identity

$$|\xi|^{2k} = \sum_{1\leqslant j_1,\cdots,j_k\leqslant n} \xi_{j_1}^2 \cdots \xi_{j_k}^2 = \sum_{|\alpha|=k} a_\alpha (i\xi)^\alpha(-i\xi)^\alpha,$$

for some integer constants a_α and the fact that $\hat{f} = \tilde{\phi}\hat{f}$, we deduce that there exists a family of integers $(a_\alpha)_{\alpha\in\mathbb{N}_0^n}$ such that

$$f = \sum_{|\alpha|=k} h_\alpha * \partial^\alpha f, \quad \text{with} \quad h_\alpha := a_\alpha \mathscr{F}^{-1}(-\omega i\xi)^\alpha |\omega\xi|^{-2k}\tilde{\phi}(\xi),$$

and the result follows. □

Remark 8.5. When the frequency is localized, one can upgrade low Lebesgue integrability to high Lebesgue integrability, at the cost of some powers of λ; when the frequency λ is very slow, this cost is in fact a gain, and it becomes quite suitable to use Bernstein's inequality whenever the opportunity arises.

The following lemma describes the action of Fourier multipliers which behave like homogeneous functions of degree m.

Lemma 8.6. *Let $\mathbb{A}$ be an annulus, $m \in \mathbb{R}$, and $k = 2[1+n/2]$ where $[s]$ denotes the integer part of $s \in \mathbb{R}$. Let σ be a k-times differentiable function on $\mathbb{R}^n \setminus \{0\}$ satisfying that for any $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leqslant k$, there exists a constant C_α such that*

$$|\partial^\alpha \sigma(\xi)| \leqslant C_\alpha |\xi|^{m-|\alpha|}, \quad \forall \xi \in \mathbb{R}^n.$$

Then, there exists a constant C, depending only on the constants C_α, such that for any $p \in [1,\infty]$ and any $\lambda > 0$, we have, for any function $f \in L^p$ with $\operatorname{supp} \hat{f} \subset \lambda\mathbb{A}$,

$$\|\sigma(D)f\|_p \leqslant C\lambda^m \|f\|_p, \quad \textit{with } \sigma(D)f := \mathscr{F}^{-1}(\sigma\hat{f}).$$

Proof. Consider a smooth function θ supported in an annulus and such that $\theta = 1$ on $\mathbb{A}$. It is clear that we have

$$\sigma(D)f = \mathscr{F}^{-1}\sigma * f = \mathscr{F}^{-1}(\theta(\xi/\lambda)\sigma(\xi)) * f = \lambda^n \mathscr{F}^{-1}(\theta(\xi)\sigma(\lambda\xi))(\lambda\cdot) * f.$$

Thus, in view of Young's inequality, we only need to prove $\lambda^n\mathscr{F}^{-1}(\theta(\xi)\sigma(\lambda\xi))(\lambda\cdot) \in L^1(\mathbb{R}^n)$, or equivalently, $\mathscr{F}^{-1}(\theta(\xi)\sigma(\lambda\xi)) \in L^1(\mathbb{R}^n)$. In fact, for $M := [1+n/2]$, we have

$$\begin{aligned}(1+|\omega x|^2)^M \mathscr{F}^{-1}(\theta(\xi)\sigma(\lambda\xi))(x) &= \mathscr{F}^{-1}(1-\Delta_\xi)^M(\theta(\xi)\sigma(\lambda\xi)) \\ =&\mathscr{F}^{-1}\left(\sum_{|\alpha|+|\beta|\leqslant 2M} C_{\alpha,\beta}\lambda^{|\beta|}\partial^\alpha\theta(\xi)\partial^\beta\sigma(\lambda\xi)\right)\end{aligned}$$

for some integers $C_{\alpha,\beta}$ (whose exact values do not matter). The integration may be restricted to $\operatorname{supp}\theta$. On this set, we have $|\partial^\beta\sigma(\lambda\xi)| \leqslant C_\beta\lambda^{m-|\beta|}$ by the assumption. Thus, we get

$$(1+|\omega x|^2)^M|\mathscr{F}^{-1}(\theta(\xi)\sigma(\lambda\xi))(x)| \leqslant C_M\lambda^m.$$

As $k = 2M > n$, we may conclude that $\|\mathscr{F}^{-1}(\theta(\xi)\sigma(\lambda\xi))\|_1 \leqslant C\lambda^m$. This completes the proof. □

By Young's inequality, we can easily prove the following crucial properties of the operators $\dot{\triangle}_k$ and $\dot{S}_k$:

Proposition 8.7 (Boundedness of the operators)**.** *For any $1 \leqslant p \leqslant \infty$ and $k \in \mathbb{Z}$, it holds*

$$\|\dot{\triangle}_k f\|_p \leqslant C\|f\|_p, \quad \|\dot{S}_k f\|_p \leqslant C\|f\|_p,$$

for some constant C independent of p.

We now study how the Littlewood-Paley pieces $\dot{\triangle}_k f$ (or $\dot{S}_k f$) of a function are related to the function itself. Specifically, we are interested in how the L^p behavior of the $\dot{\triangle}_k f$ relate to the L^p behavior of f. One can already see this when $p = 2$, in which case we have

$$\|f\|_2 \sim \left(\sum_{k\in\mathbb{Z}} \|\dot{\triangle}_k f\|_2^2\right)^{1/2}. \tag{8.7}$$

In fact, we square both sides and take Plancherel to obtain

$$\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 d\xi \sim \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^n} |\varphi_k(\xi)|^2 |\hat{f}(\xi)|^2 d\xi.$$

Observe that for each $\xi \neq 0$ there are only three values of $\varphi_k(\xi)$ which does not vanish. That is, for $|\xi| \sim 2^\ell$,

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}} |\varphi_k(\xi)|^2 = \varphi_{\ell-1}^2(\xi) + \varphi_\ell^2(\xi) + \varphi_{\ell+1}^2(\xi)\\
=&(\varphi_{\ell-1}(\xi) + \varphi_\ell(\xi) + \varphi_{\ell+1}(\xi))^2 - 2(\varphi_{\ell-1}(\xi)\varphi_\ell(\xi) + \varphi_{\ell-1}(\xi)\varphi_{\ell+1}(\xi) + \varphi_\ell(\xi)\varphi_{\ell+1}(\xi))\\
=&1 - 2(\varphi_{\ell-1}(\xi) + \varphi_{\ell+1}(\xi))\varphi_\ell(\xi) = 1 - 2(1 - \varphi_\ell(\xi))\varphi_\ell(\xi) = 1 - 2\varphi_\ell(\xi) + 2\varphi_\ell^2(\xi)\\
=&\frac{1}{2} + 2\left(\frac{1}{2} - \varphi_\ell(\xi)\right)^2,
\end{aligned}$$

which yields

$$\frac{1}{2} \leqslant \sum_{k\in\mathbb{Z}} |\varphi_k(\xi)|^2 \leqslant 1, \quad \forall \xi \neq 0.$$

The claim follows.

Another way to rewrite (8.7) is

$$\|f\|_2 \sim \left\|\left(\sum_{k\in\mathbb{Z}} |\dot{\triangle}_k f|^2\right)^{1/2}\right\|_2. \tag{8.8}$$

The quantity $\left(\sum_{k\in\mathbb{Z}} |\dot{\triangle}_k f|^2\right)^{1/2}$ is also known as the *Littlewood-Paley square function.* More generally, the Littlewood-Paley square function theorem, i.e., Theorem 6.26 is valid for this smooth type decomposition:

Theorem 8.8 (Littlewood-Paley square function theorem)**.** *For any $1 < p < \infty$, we have*

$$\left\|\left(\sum_{k\in\mathbb{Z}} |\dot{\triangle}_k f|^2\right)^{1/2}\right\|_p \sim \|f\|_p$$

with the implicit constant depending on p.

The proof of this theorem is very similar to that of Theorem 6.26, so we remain it to the interested reader.

Now, we can give the proof of Theorem 7.16.

Proof of Theorem 7.16. Since $0 \notin \operatorname{supp} \hat{f}$, we have $\hat{f}(\xi) = 0$ in a neighborhood of $\xi = 0$. Then there is some integer k_0 such that $f = \sum_{k\geqslant k_0} \dot{\triangle}_k f$. Noting that

$$(1 + |\omega\xi|^2)^{s/2} |\omega\xi|^{-s} \sum_{k\geqslant k_0} \varphi_k(\xi) \in M_1$$

by the Mihlin multiplier theorem, we see that for $f \in \dot{H}_p^s(\mathbb{R}^n)$

$$\|f\|_{H_p^s} = \|\mathscr{F}^{-1}(1 + |\omega\xi|^2)^{s/2} |\omega\xi|^{-s} \sum_{k\geqslant k_0} \varphi_k(\xi) \mathscr{F} I^s f\|_p \leqslant C\|f\|_{\dot{H}_p^s}.$$

Conversely, if $f \in H_p^s$, then we note that $|\omega\xi|^s(1+|\omega\xi|^2)^{-s/2}\sum_{k\geqslant k_0}\varphi_k(\xi) \in M_1$ in view of Mihlin multiplier theorem. Thus,

$$\|f\|_{\dot{H}_p^s} = \|\mathscr{F}^{-1}|\omega\xi|^s(1+|\omega\xi|^2)^{-s/2}\sum_{k\geqslant k_0}\varphi_k(\xi)\mathscr{F}J^s f\|_p \leqslant C\|f\|_{H_p^s}.$$

If $f \in L^p \cap \dot{H}_p^s$, then we obtain as above for $s \in \mathbb{R}$

$$\begin{aligned}\|f\|_{H_p^s} \leqslant&\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\sum_{k\geqslant 0}\varphi_k(\xi)\mathscr{F}I^s f\|_p\\ &+\|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{s/2}\sum_{k<0}\varphi_k(\xi)\mathscr{F}f\|_p\\ \leqslant&C(\|f\|_{\dot{H}_p^s}+\|f\|_p).\end{aligned}$$

Conversely, if $s>0$ and $f \in H_p^s$, then clearly $f \in L^p$ and

$$\begin{aligned}\|f\|_{\dot{H}_p^s} \leqslant&\|\mathscr{F}^{-1}|\omega\xi|^s(1+|\omega\xi|^2)^{-s/2}\sum_{k\geqslant 0}\varphi_k(\xi)\mathscr{F}J^s f\|_p\\ &+|\omega|^s\sum_{k<0}2^{ks}\|\mathscr{F}^{-1}(2^{-k}|\xi|)^s\varphi(2^{-k}\xi)\mathscr{F}f\|_p\\ \leqslant&C(\|f\|_{H_p^s}+\|f\|_p) \leqslant C\|f\|_{H_p^s}.\end{aligned}$$

Now, we consider the case $s<0$. If $f \in L^p + \dot{H}_p^s$, i.e., $f = f_1 + f_2$ for some $f_1 \in L^p$ and $f_2 \in \dot{H}_p^s$, then

$$\begin{aligned}\|f\|_{H_p^s} \leqslant&\|f_1\|_{H_p^s}+\|f_2\|_{H_p^s} = \|J^s f_1\|_p + \|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\mathscr{F}I_s f_2\|_p\\ \leqslant&\|f_1\|_p + C\|f_2\|_{\dot{H}_p^s},\end{aligned}$$

by Theorem 7.7 and the fact that $(1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s} \in M_1$ for $s<0$ by the Mihlin multiplier theorem. Conversely, if $f \in H_p^s$, then $f = \sum_{k<0}\dot{\triangle}_k f + \sum_{k\geqslant 0}\dot{\triangle}_k f$ where $\|\sum_{k<0}\dot{\triangle}_k f\|_p \leqslant \|f\|_p$ and $\|\sum_{k\geqslant 0}\dot{\triangle}_k f\|_{\dot{H}_p^s} \leqslant \|f\|_{H_p^s}$ by the first conclusion since $0 \notin \operatorname{supp}\mathscr{F}(\sum_{k\geqslant 0}\dot{\triangle}_k f)$.

For the case $s=0$, it is obviously from the definitions. □

8.2 Besov spaces and Triebel-Lizorkin spaces

The Littlewood-Paley decomposition is very useful. For example, we can define (independently of the choice of the initial function ψ) the following Besov spaces. In order to define the homogeneous Besov spaces, we first modify the Schwartz space $\mathscr{S}$ and its dual $\mathscr{S}'$. Denote

$$\dot{\mathscr{S}}(\mathbb{R}^n) = \left\{f \in \mathscr{S}(\mathbb{R}^n):\ (\partial^\alpha \hat{f})(0) = 0,\ \forall \alpha \in \mathbb{N}_0^n\right\}, \tag{8.9}$$

which is a subspace of $\mathscr{S}(\mathbb{R}^n)$ with the same topology. We denote its dual space by $\dot{\mathscr{S}}'(\mathbb{R}^n)$ which can be also identified by the quotient space of $\mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n)$ with the polynomials space $\mathscr{P}(\mathbb{R}^n)$. Then we can give the following definition.

Definition 8.9. Let $s \in \mathbb{R}$, $1 \leqslant p,\ r \leqslant \infty$. We write

$$\|f\|_{\dot{B}_{p,r}^s} = \left(\sum_{k=-\infty}^{\infty}\left(2^{sk}\|\dot{\triangle}_k f\|_p\right)^r\right)^{\frac{1}{r}}, \quad \forall f \in \dot{\mathscr{S}}'(\mathbb{R}^n),$$

$$\|f\|_{B^s_{p,r}} = \|S_0 f\|_p + \left(\sum_{k=0}^{\infty} \left(2^{sk}\|\triangle_k f\|_p\right)^r\right)^{\frac{1}{r}}, \quad \forall f \in \mathscr{S}'(\mathbb{R}^n).$$

The homogeneous Besov space $\dot{B}^s_{p,r}$ is defined by

$$\dot{B}^s_{p,r} = \left\{f \in \dot{\mathscr{S}}'(\mathbb{R}^n) : \|f\|_{\dot{B}^s_{p,r}} < \infty\right\},$$

and the nonhomogeneous Besov space $B^s_{p,r}$ is defined by

$$B^s_{p,r} = \left\{f \in \mathscr{S}'(\mathbb{R}^n) : \|f\|_{B^s_{p,r}} < \infty\right\}.$$

For the sake of completeness, we also define the Triebel-Lizorkin spaces.

Definition 8.10. Let $s \in \mathbb{R}$, $1 \leqslant p < \infty$, $1 \leqslant r \leqslant \infty$. We write

$$\|f\|_{\dot{F}^s_{p,r}} = \left\|\left(\sum_{k=-\infty}^{\infty} \left(2^{sk}|\dot{\triangle}_k f|\right)^r\right)^{\frac{1}{r}}\right\|_p, \quad \forall f \in \dot{\mathscr{S}}'(\mathbb{R}^n),$$

$$\|f\|_{F^s_{p,r}} = \|S_0 f\|_p + \left\|\left(\sum_{k=0}^{\infty} \left(2^{sk}|\triangle_k f|\right)^r\right)^{\frac{1}{r}}\right\|_p, \quad \forall f \in \mathscr{S}'(\mathbb{R}^n).$$

The homogeneous Triebel-Lizorkin space $\dot{F}^s_{p,r}$ is defined by

$$\dot{F}^s_{p,r} = \left\{f \in \dot{\mathscr{S}}'(\mathbb{R}^n) : \|f\|_{\dot{F}^s_{p,r}} < \infty\right\},$$

and the nonhomogeneous Triebel-Lizorkin space $F^s_{p,r}$ is defined by

$$F^s_{p,r} = \left\{f \in \mathscr{S}'(\mathbb{R}^n) : \|f\|_{F^s_{p,r}} < \infty\right\}.$$

Remark 8.11. It is easy to see that the above quantities define a quasi-norm and a norm in general, with the usual convention that $r = \infty$ in both cases corresponds to the usual L^∞ norm. On the other hand, we have not included the case $r = \infty$ in the definition of Triebel-Lizorkin space because the L^∞ norm has to be replaced here by a more complicated Carleson measure.

Besov space and Triebel-Lizorkin space were constructed between 1960's and 1980's. Recently, they are widely applied to study PDEs. Roughly speaking, these spaces are products of the function spaces $\ell^r(L^p)$ or $L^p(\ell^r)$ by combining the Littlewood-Paley decomposition of phase space. The index s in the definition, describes the regularity of the space.

From Theorem 8.8, we immediately have the following relations involving Sobolev spaces and Triebel-Lizorkin spaces:

Theorem 8.12. *Let $s \in \mathbb{R}$ and $1 < p < \infty$. Then*

$$H^s_p = F^s_{p,2}, \quad \dot{H}^s_p = \dot{F}^s_{p,2}, \tag{8.10}$$

with equivalent norms.

For simplicity, we use X to denote B or F in the spaces, that is, $X^s_{p,r}$ ($\dot{X}^s_{p,r}$, resp.) denotes $B^s_{p,r}$ ($\dot{B}^s_{p,r}$, resp.) or $F^s_{p,r}$ ($\dot{F}^s_{p,r}$, resp.). But it will denote only one of them in the same formula. We always assume that $1 \leqslant p \leqslant \infty$ for $B^s_{p,r}$ ($\dot{B}^s_{p,r}$, resp.) and $1 \leqslant p < \infty$ for $F^s_{p,r}$ ($\dot{F}^s_{p,r}$, resp.) if no other statement is declared. We have the following embedding relations:

Theorem 8.13. *Let X denote B or F. Then, we have the following embedding:*

$$X^s_{p,r_1} \subset X^s_{p,r_2}, \quad \dot{X}^s_{p,r_1} \subset \dot{X}^s_{p,r_2}, \quad \text{if } r_1 \leqslant r_2,$$

$$\begin{aligned} X^{s+\varepsilon}_{p,r_1} \subset X^s_{p,r_2}, &\quad \textit{if } \varepsilon > 0,\\ B^s_{p,\min(p,r)} \subset F^s_{p,r} \subset B^s_{p,\max(p,r)}, &\quad \textit{if } 1 \leqslant p < \infty,\\ \dot{B}^s_{p,\min(p,r)} \subset \dot{F}^s_{p,r} \subset \dot{B}^s_{p,\max(p,r)}, &\quad \textit{if } 1 \leqslant p < \infty. \end{aligned}$$

Proof. It is clear that the first one is valid because of $\ell^r \subset \ell^{r+a}$ for any $a \geqslant 0$. For the second one, we notice that

$$\left(\sum_{k=0}^{\infty} 2^{skr_2}|a_k|^{r_2}\right)^{\frac{1}{r_2}} \leqslant \sup_{k\geqslant 0} 2^{(s+\varepsilon)k}|a_k| \left(\sum_{k=0}^{\infty} 2^{-\varepsilon k r_2}\right)^{\frac{1}{r_2}} \lesssim \sup_{k\geqslant 0} 2^{(s+\varepsilon)k}|a_k|.$$

Taking $a_k = \|\triangle_k f\|_p$ or $a_k = |\triangle_k f|$, we can get

$$X^{s+\varepsilon}_{p,\infty} \subset X^s_{p,r_2},$$

which yields the second result in view of the first one.

For the third and last one, we separate into two cases and denote $b_k = 2^{sk}|\dot{\triangle}_k f|$ and $j = 0$ for the third or $j = -\infty$ for the last one.

Case I: $r \leqslant p$. In this case, we have $\ell^r \subset \ell^p$ and

$$\begin{aligned} \sum_{k=j}^{\infty} \|b_k\|_p^p =& \sum_{k=j}^{\infty} \int_{\mathbb{R}^n} |b_k(x)|^p dx = \int_{\mathbb{R}^n} \sum_{k=j}^{\infty} |b_k(x)|^p dx\\ =& \int_{\mathbb{R}^n} \|(b_k)\|_{\ell^p}^p dx \lesssim \int_{\mathbb{R}^n} \|(b_k)\|_{\ell^r}^p dx, \end{aligned}$$

which yields the second parts of embedding relations. Moreover, by Minkowski's inequality,[3] we get

$$\left\| \left(\sum_{k=j}^{\infty} b_k^r\right)^{\frac{1}{r}} \right\|_p^r = \left\| \sum_{k=j}^{\infty} b_k^r \right\|_{\frac{p}{r}} \leqslant \sum_{k=j}^{\infty} \|b_k^r\|_{\frac{p}{r}} = \sum_{k=j}^{\infty} \|b_k\|_p^r,$$

which yields the first parts of embedding relations.

Case II: $p < r$. By Minkowski's inequality, we have

$$\sum_{k=j}^{\infty} \|b_k\|_p^r = \sum_{k=j}^{\infty} \|b_k^r\|_{\frac{p}{r}} \leqslant \left\| \sum_{k=j}^{\infty} b_k^r \right\|_{\frac{p}{r}} = \left\| \left(\sum_{k=j}^{\infty} b_k^r\right)^{\frac{1}{r}} \right\|_p^r,$$

which yields the second parts of embedding relations.

In this case, we have $\ell^p \subset \ell^r$ and

$$\| \|(b_k)\|_{\ell^r} \|_p^p \lesssim \| \|(b_k)\|_{\ell^p} \|_p^p = \left\| \sum_{k=j}^{\infty} b_k^p \right\|_1 = \sum_{k=j}^{\infty} \|b_k\|_p^p,$$

which yields the first parts of embedding relations. Thus, we complete the proof. □

From Theorems 8.12 and 8.13, we can get the following corollary.

Corollary 8.14. *Let $s \in \mathbb{R}$. Then we have*

i) *For $1 < p < \infty$, $B^s_{p,\min(p,2)} \subset H^s_p \subset B^s_{p,\max(p,2)}$ and $\dot{B}^s_{p,\min(p,2)} \subset \dot{H}^s_p \subset \dot{B}^s_{p,\max(p,2)}$. In particular, $H^s = B^s_{2,2} = F^s_{2,2}$ and $\dot{H}^s = \dot{B}^s_{2,2} = \dot{F}^s_{2,2}$.*

ii) *For $1 \leqslant p \leqslant \infty$, $B^s_{p,1} \subset H^s_p \subset B^s_{p,\infty}$ and $\dot{B}^s_{p,1} \subset \dot{H}^s_p \subset \dot{B}^s_{p,\infty}$.*

---

[3] Minkowski's inequalities read
i) $\|\sum_{j=0}^{\infty} f_j\|_p \leqslant \sum_{j=0}^{\infty} \|f_j\|_p$, for any $p \in [1,\infty]$;
ii) $\sum_{j=0}^{\infty} \|f_j\|_p \leqslant \|\sum_{j=0}^{\infty} f_j\|_p$, for any $p \in (0,1)$ and $f_j \geqslant 0$.

Proof. It obviously follows from Theorems 8.12 and 8.13 except the endpoint cases $p=1$ or ∞ in ii). For the proof of the endpoint cases, one can see [BL76, Chapter 6]. $\square$

Theorem 8.15. *Let X denote B or F. Then,*

i) *$X^s_{p,r}$ is a Banach space;*

ii) *$\mathscr{S}(\mathbb{R}^n) \subset X^s_{p,r} \subset \mathscr{S}'(\mathbb{R}^n)$, $\dot{\mathscr{S}}(\mathbb{R}^n) \subset \dot{X}^s_{p,r} \subset \dot{\mathscr{S}}'(\mathbb{R}^n)$;*

iii) *$\mathscr{S}(\mathbb{R}^n)$ is dense in $X^s_{p,r}$, if $1 \leqslant p,r < \infty$; $\dot{\mathscr{S}}(\mathbb{R}^n)$ is dense in $\dot{X}^s_{p,r}$, if $1 \leqslant p,r < \infty$.*

Proof. We only show the nonhomogeneous cases and leave the homogeneous cases to the interested reader. Clearly, $X^s_{p,r}$ is a normed linear space with the norm $\|\cdot\|_{X^s_{p,r}}$ since either $\ell^r(L^p)$ or $L^p(\ell^r)$ is a normed linear space. Moreover, it is complete and therefore Banach space which will be proved in the final. Let's first prove the second result. We divide the proofs into two steps.

Step 1: To prove $\mathscr{S} \subset B^s_{p,\infty}$. In fact, for $\sigma = \max(s,0)$ and sufficiently large[4] L, $N \in \mathbb{N}_0$, we have for any $f \in \mathscr{S}$, from Proposition 8.4 and 8.7, that

$$\begin{aligned}\|f\|_{B^s_{p,\infty}} =&\|S_0 f\|_p + \sup_{k\geqslant 0} 2^{sk}\|\triangle_k f\|_p \leqslant C\|f\|_p + \sup_{k\geqslant 0} 2^{sk}\|(\sqrt{-\Delta})^{-\sigma}\triangle_k(\sqrt{-\Delta})^{\sigma} f\|_p\\ \lesssim&\sum_{\alpha,\beta}|f|_{\alpha,\beta} + \sup_{k\geqslant 0} 2^{sk}2^{-\sigma k}\|(\sqrt{-\Delta})^{\sigma} f\|_p\\ \lesssim&\sum_{\alpha,\beta}|f|_{\alpha,\beta} + \|(1+|x|^2)^L(\sqrt{-\Delta})^{\sigma} f\|_\infty \lesssim \sum_{\alpha,\beta}|f|_{\alpha,\beta}.\end{aligned}$$

where $|f|_{\alpha,\beta}$ is one of the semi-norm sequence of $\mathscr{S}$. Thus, we obtain the result.

Step 2: To prove $\mathscr{S} \subset X^s_{p,r}$. From Step 1, we know $\mathscr{S} \subset B^{s+\varepsilon}_{p,\infty}$ for any $\varepsilon > 0$. From Theorem 8.13, we get $B^{s+\varepsilon}_{p,\infty} \subset B^s_{p,\min(p,r)} \subset B^s_{p,r} \cap F^s_{p,r}$. Therefore, $\mathscr{S} \subset X^s_{p,r}$.

Finally, let us prove the completeness of $B^s_{p,r}$. The completeness of $F^s_{p,r}$ can be proved in a similar way. Let $\{f_l\}_1^\infty$ be a Cauchy sequence in $B^s_{p,r}$. So it does in $\mathscr{S}'$ in view of ii). Because $\mathscr{S}'$ is a complete local convex topological linear space, there exists a $f \in \mathscr{S}'$ such that $f_l \to f$ according to the strong topology of $\mathscr{S}'$. On the other hand, that $\{f_l\}_1^\infty$ is a Cauchy sequence implies that $\{\triangle_k f_l\}_{l=1}^\infty$ is a Cauchy sequence in L^p. From the completeness of L^p, there is a $g_k \in L^p$ such that

$$\|\triangle_k f_l - g_k\|_p \to 0, \quad l \to \infty. \tag{8.11}$$

Since $L^p \subset \mathscr{S}'$ and $\triangle_k f_l \to \triangle_k f$ as $l \to \infty$ in $\mathscr{S}'$, we get $g_k = \triangle_k f$. Hence, (8.11) implies

$$\|\triangle_k(f_l - f)\|_p \to 0, \quad l \to \infty.$$

which yields $\sup_{k\geqslant 0} 2^{(s+\varepsilon)k}\|\triangle_k(f_l - f)\|_p \to 0$ as $l \to \infty$ for any $\varepsilon > 0$.

Similarly, we have

$$\|S_0(f_l - f)\|_p \to 0, \quad l \to \infty.$$

Therefore,

$$\|f_l - f\|_{B^s_{p,r}} \lesssim \|f_l - f\|_{B^{s+\varepsilon}_{p,\infty}} \to 0, \quad l \to \infty.$$

In a similar way, we can obtain the density statement in iii). We omit the details. $\square$

---

[4] It is enough to assume that $L > \frac{n}{2p}$ and $N \geqslant \max(2L,\sigma)$. In fact,

$$\begin{aligned}\|(1+|x|^2)^{-L}\|_p =& C\left(\int_0^\infty r^{n-1}(1+r^2)^{-pL}dr\right)^{1/p} \leqslant C2^L\left(\int_0^\infty r^{n-1}(1+r)^{-2pL}dr\right)^{1/p}\\ \leqslant& C2^L\left(\int_0^\infty (1+r)^{-2pL+n-1}dr\right)^{1/p} \leqslant C2^L(2pL-n)^{-1/p},\end{aligned}$$

where we assume that $2pL > n$.

8.3 Embedding theorems and Gagliardo-Nirenberg inequalities

Theorem 8.16 (The embedding theorem)**.** *Let* $1 \leqslant p, p_1, r, r_1 \leqslant \infty$ *and* $s, s_1 \in \mathbb{R}$. *Assume that* $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$. *Then the following conclusions hold*

$$\begin{aligned} &B^s_{p,r} \subset B^{s_1}_{p_1,r_1}, \quad \dot{B}^s_{p,r} \subset \dot{B}^{s_1}_{p_1,r_1}, \quad \forall p \leqslant p_1 \text{ and } r \leqslant r_1. \\ &F^s_{p,r} \subset F^{s_1}_{p_1,r_1}, \quad \dot{F}^s_{p,r} \subset \dot{F}^{s_1}_{p_1,r_1}, \quad \forall p < p_1 < \infty. \end{aligned}$$

Proof. We only give the proof of the nonhomogeneous cases, the homogeneous cases can be treated in a similar way.

Let us prove the first conclusion. From the Bernstein inequality in Proposition 8.4, we immediately have

$$\|\triangle_k f\|_{p_1} \lesssim 2^{kn(\frac{1}{p}-\frac{1}{p_1})}\|\triangle_k f\|_p, \quad \|S_0 f\|_{p_1} \lesssim \|S_0 f\|_p, \tag{8.12}$$

since $1 \leqslant p \leqslant p_1 \leqslant \infty$. Thus, with the help of the embedding $B^s_{p,r} \subset B^s_{p,r_1}$ for $r \leqslant r_1$ in Theorem 8.13, we get

$$\begin{aligned} \|f\|_{B^{s_1}_{p_1,r_1}} =& \|S_0 f\|_{p_1} + \left(\sum_{k=0}^{\infty} \left(2^{s_1 k}\|\triangle_k f\|_{p_1}\right)^{r_1}\right)^{\frac{1}{r_1}} \\ \lesssim& \|S_0 f\|_p + \left(\sum_{k=0}^{\infty} \left(2^{sk}\|\triangle_k f\|_p\right)^{r_1}\right)^{\frac{1}{r_1}} = \|f\|_{B^s_{p,r_1}} \lesssim \|f\|_{B^s_{p,r}}. \end{aligned}$$

This gives the first conclusion.

Next, we prove the second conclusion. In view of Theorem 8.13, we need only prove

$$F^s_{p,\infty} \subset F^{s_1}_{p_1,1}.$$

Without loss of generality, we assume $\|f\|_{F^s_{p,\infty}} = 1$ and consider the norm

$$\|f\|_{F^{s_1}_{p_1,1}} = \|S_0 f\|_{p_1} + \left\|\sum_{k=0}^{\infty} 2^{s_1 k}|\triangle_k f|\right\|_{p_1}.$$

We use the following equivalent norm (i.e., Theorem 2.16) on L^p for $1 \leqslant p < \infty$:

$$\|f\|_p^p = p\int_0^\infty t^{p-1}\mathrm{m}(\{x : |f(x)| > t\})dt.$$

Thus, we have

$$\begin{aligned} \left\|\sum_{k=0}^{\infty} 2^{s_1 k}|\triangle_k f|\right\|_{p_1}^{p_1} =& p_1\int_0^A t^{p_1-1}\mathrm{m}\left(\left\{x : \sum_{k=0}^{\infty} 2^{s_1 k}|(\triangle_k f)(x)| > t\right\}\right)dt \\ &+ p_1\int_A^\infty t^{p_1-1}\mathrm{m}\left(\left\{x : \sum_{k=0}^{\infty} 2^{s_1 k}|(\triangle_k f)(x)| > t\right\}\right)dt \\ =&: I + II, \end{aligned}$$

where $A \gg 1$ is a constant which can be chosen as below. Noticing that $p < p_1$ and $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$ imply $s > s_1$, we have

$$\sum_{k=K}^{\infty} 2^{s_1 k}|\triangle_k f| \lesssim 2^{K(s_1-s)}\sup_{k\geqslant 0} 2^{sk}|\triangle_k f|, \quad \forall K \in \mathbb{N}_0. \tag{8.13}$$

By taking $K = 0$ and noticing $p < p_1$ (which implies that $t^{p_1-1} \leqslant A^{p_1-p}t^{p-1}$ for $t \leqslant A$), we get

$$I \lesssim \int_0^A t^{p_1-1}\mathrm{m}\left(\left\{x : \sup_{k\geqslant 0} 2^{sk}|(\triangle_k f)(x)| > ct\right\}\right)dt$$

$$\lesssim \int_0^{cA} \tau^{p-1} \mathfrak{m}\left(\left\{x : \sup_{k\geqslant 0} 2^{sk}|(\triangle_k f)(x)| > \tau\right\}\right) d\tau \lesssim \left\|\sup_{k\geqslant 0} 2^{sk}|\triangle_k f|\right\|_p^p \lesssim 1,$$

where the implicit constant depends on A, but it is a fixed constant.

Now we estimate II. By the Bernstein inequality in Proposition 8.4, we have

$$\|\triangle_k f\|_\infty \lesssim 2^{kn/p}\|\triangle_k f\|_p \lesssim 2^{k(n/p-s)} \left\|\sup_{k\geqslant 0} 2^{sk}|\triangle_k f|\right\|_p .$$

Hence, for $K \in \mathbb{N}$, we obtain

$$\begin{aligned}\sum_{k=0}^{K-1} 2^{s_1 k}|\triangle_k f| \lesssim &\sum_{k=0}^{K-1} 2^{k(s_1-s+n/p)} \left\|\sup_{k\geqslant 0} 2^{sk}|\triangle_k f|\right\|_p \\ \lesssim &2^{Kn/p_1} \left\|\sup_{k\geqslant 0} 2^{sk}|\triangle_k f|\right\|_p \lesssim 2^{Kn/p_1}.\end{aligned} \tag{8.14}$$

Taking K to be the largest natural number satisfying $C2^{Kn/p_1} \leqslant t/2$, we have $2^K \sim t^{p_1/n}$. It is easy to see that such a K exists if $t \geqslant A \gg 1$. Thus, for $t \geqslant A$ and $\sum_{k=0}^\infty 2^{s_1 k}|(\triangle_k f)(x)| \geqslant t$, we have, from (8.13) and (8.14), that

$$C2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk}|\triangle_k f| \geqslant \sum_{k=K}^\infty 2^{s_1 k}|\triangle_k f| > t/2. \tag{8.15}$$

Hence, from (8.14) and (8.15), we get

$$\begin{aligned} II =&p_1 \int_A^\infty t^{p_1-1}\mathfrak{m}\left(\left\{x : \sum_{k=0}^\infty 2^{s_1 k}|(\triangle_k f)(x)| > t\right\}\right) dt \\ \lesssim &\int_A^\infty t^{p_1-1}\mathfrak{m}\left(\left\{x : \sum_{k=0}^{K-1} 2^{s_1 k}|(\triangle_k f)(x)| > t/2\right\}\right) dt \\ &\quad + \int_A^\infty t^{p_1-1}\mathfrak{m}\left(\left\{x : \sum_{k=K}^{\infty} 2^{s_1 k}|(\triangle_k f)(x)| > t/2\right\}\right) dt \\ \lesssim &\int_A^\infty t^{p_1-1}\mathfrak{m}\left(\left\{x : C2^{Kn/p_1} > t/2\right\}\right) dt \\ &\quad + \int_A^\infty t^{p_1-1}\mathfrak{m}\left(\left\{x : C2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk}|(\triangle_k f)(x)| > t/2\right\}\right) dt \\ \lesssim &\int_A^\infty t^{p_1-1}\mathfrak{m}\left(\left\{x : \sup_{k\geqslant 0} 2^{sk}|(\triangle_k f)(x)| > ct^{p_1/p}\right\}\right) dt \\ \lesssim &\int_{A'}^\infty \tau^{p-1}\mathfrak{m}\left(\left\{x : \sup_{k\geqslant 0} 2^{sk}|(\triangle_k f)(x)| > \tau\right\}\right) d\tau \\ \lesssim &\|\sup_{k\geqslant 0} 2^{sk}|\triangle_k f|\|_p^p \lesssim 1.\end{aligned}$$

That is,

$$\left\|\sum_{k=0}^\infty 2^{s_1 k}|\triangle_k f|\right\|_{p_1} \lesssim 1.$$

On the other hand, from (8.12), we have $\|S_0 f\|_{p_1} \lesssim 1$. Therefore, we have obtained $\|f\|_{B^{s_1}_{p_1,1}} \lesssim 1$ under the assumption $\|f\|_{B^s_{p,\infty}} = 1$. This completes the proof. □

Theorem 8.17. *Let $1 \leqslant p < \infty$, $s > n/p$ and $1 \leqslant r \leqslant \infty$. Let $X^s_{p,r}$ denote $B^s_{p,r}$ or $F^s_{p,r}$. Then it holds*

$$X^s_{p,r} \subset B^0_{\infty,1} \subset L^\infty.$$

Proof. By Bernstein's inequality and Theorem 8.13, we have

$$\|f\|_\infty \leqslant \sum_{k=-1}^{\infty} \|\triangle_k f\|_\infty \lesssim \sum_{k=-1}^{\infty} 2^{kn/p}\|\triangle_k f\|_p \lesssim \left(\sum_{k=-1}^{\infty} 2^{k(n/p-s)}\right) \|f\|_{B^s_{p,\infty}} \lesssim \|f\|_{X^s_{p,r}}.$$

□

Now, we give some fractional Gagliardo-Nirenberg inequalities in homogeneous Besov, Triebel-Lizorkin and Sobolev spaces without proofs (cf. [HMOW10]).

Theorem 8.18. *Let* $0 < p, p_0, p_1, r, r_0, r_1 \leqslant \infty$, $s, s_0, s_1 \in \mathbb{R}$, $0 \leqslant \theta \leqslant 1$. *Assume that*

$$s - \frac{n}{p} = (1-\theta)\left(s_0 - \frac{n}{p_0}\right) + \theta\left(s_1 - \frac{n}{p_1}\right),$$

then the fractional Gagliardo-Nirenberg inequality of the following type

$$\|f\|_{\dot{B}^s_{p,r}} \lesssim \|f\|^{1-\theta}_{\dot{B}^{s_0}_{p_0,r_0}} \|f\|^{\theta}_{\dot{B}^{s_1}_{p_1,r_1}}$$

holds for all $f \in \dot{B}^{s_0}_{p_0,r_0} \cap \dot{B}^{s_1}_{p_1,r_1}$ *if and only if one of the following conditions holds:*

i) $s \leqslant (1-\theta)s_0 + \theta s_1$ *and* $\dfrac{1}{r} \leqslant \dfrac{1-\theta}{r_0} + \dfrac{\theta}{r_1}$;

ii) $p_0 = p_1$ *and* $s = (1-\theta)s_0 + \theta s_1$ *but* $s_0 \neq s_1$;

iii) $s_0 - \dfrac{n}{p_0} \neq s - \dfrac{n}{p}$ *and* $s < (1-\theta)s_0 + \theta s_1$.

For homogeneous Triebel-Lizorkin spaces $\dot{F}^s_{p,r}$, we have the following:

Theorem 8.19. *Let* $1 \leqslant p, p_i, r < \infty$, $s, s_0, s_1 \in \mathbb{R}$, $0 < \theta < 1$. *Then the fractional Gagliardo-Nirenberg inequality of the following type*

$$\|f\|_{\dot{F}^s_{p,r}} \lesssim \|f\|^{1-\theta}_{\dot{F}^{s_0}_{p_0,\infty}} \|f\|^{\theta}_{\dot{F}^{s_1}_{p_1,\infty}}$$

holds if and only if

$$s - \frac{n}{p} = (1-\theta)\left(s_0 - \frac{n}{p_0}\right) + \theta\left(s_1 - \frac{n}{p_1}\right),$$
$$s \leqslant (1-\theta)s_0 + \theta s_1,$$
$$s_0 \neq s_1 \quad \textit{if} \quad s = (1-\theta)s_0 + \theta s_1.$$

As a corollary, the following is the Gagliardo-Nirenberg inequality with fractional derivatives in Sobolev spaces.

Corollary 8.20. *Let* $1 < p, p_0, p_1 < \infty$, s, $s_1 \in \mathbb{R}$, $0 \leqslant \theta \leqslant 1$. *Then the fractional Gagliardo-Nirenberg inequality of the following type*

$$\|f\|_{\dot{H}^s_p} \lesssim \|f\|^{1-\theta}_{L^{p_0}} \|f\|^{\theta}_{\dot{H}^{s_1}_{p_1}}$$

holds if and only if

$$s - \frac{n}{p} = (1-\theta)\frac{n}{p_0} + \theta\left(s_1 - \frac{n}{p_1}\right), \quad s \leqslant \theta s_1.$$

At the end of this section, we state the following interpolation inequalities of nonhomogeneous Besov spaces.

Theorem 8.21. *Let* $s_1, s_2 \in \mathbb{R}$ *with* $s_1 < s_2$, $0 < \theta < 1$, *and* $1 \leqslant p, r \leqslant \infty$. *Let* $s = (1-\theta)\theta s_1 + \theta s_2$, *then there sxists a constant* $C > 0$ *such that*

$$\|f\|_{B^s_{p,r}} \leqslant \|f\|^{1-\theta}_{B^{s_1}_{p,r}} \|f\|^{\theta}_{B^{s_2}_{p,r}},$$
$$\|f\|_{B^s_{p,1}} \leqslant \frac{C}{s_2 - s_1}\left(\frac{1}{\theta} + \frac{1}{1-\theta}\right) \|f\|^{1-\theta}_{B^{s_1}_{p,\infty}} \|f\|^{\theta}_{B^{s_2}_{p,\infty}}.$$

8.4 Differential-difference norm on Besov spaces

The next theorem points to an alternative definition of the Besov spaces $B^s_{p,r}$ $(s > 0)$ in terms of derivatives and moduli of continuity. The modulus of continuity is defined by

$$\omega_p^m(t, f) = \sup_{|y|<t} \| \triangle_y^m f\|_p,$$

where $\triangle_y^m$ is the m-th order difference operator:

$$\triangle_y^m f(x) = \sum_{k=0}^m C_m^k (-1)^k f(x + ky).$$

Theorem 8.22. *Assume that $s > 0$, and let m and N be integers, such that $m + N > s$ and $0 \leqslant N < s$. Then, with $1 \leqslant p, r \leqslant \infty$,*

$$\|f\|_{B^s_{p,r}} \sim \|f\|_p + \sum_{j=1}^n \left(\int_0^\infty \left(t^{N-s} \omega_p^m \left(t, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \frac{dt}{t} \right)^{1/r}.$$

Proof. We note that ω_p^m is an increasing function of t. Therefore it is sufficient to prove that

$$\|f\|_{B^s_{p,r}} \sim \|f\|_p + \sum_{j=1}^n \left(\sum_{\ell=-\infty}^{\infty} \left(2^{\ell(s-N)} \omega_p^m \left(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \right)^{1/r}.$$

First, we assume that $f \in B^s_{p,r}$. It is clear that

$$\begin{aligned}
\omega_p^m(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N}) &= \sup_{|y|<2^{-\ell}} \left\| \triangle_y^m \frac{\partial^N f}{\partial x_j^N} \right\|_p = \sup_{|y|<2^{-\ell}} \left\| \sum_{k=0}^m C_m^k (-1)^k \frac{\partial^N f}{\partial x_j^N}(x + ky) \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \sum_{k=0}^m C_m^k (-1)^k \mathscr{F}^{-1} e^{\omega iky\xi} \mathscr{F} \frac{\partial^N f}{\partial x_j^N}(x) \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \left(\mathscr{F}^{-1} \sum_{k=0}^m C_m^k (-1)^k e^{\omega iky\xi} \right) * \frac{\partial^N f}{\partial x_j^N} \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \left(\mathscr{F}^{-1} (1 - e^{\omega iy\xi})^m \right) * \frac{\partial^N f}{\partial x_j^N} \right\|_p.
\end{aligned}$$

Thus, we denote $\rho_y(\xi) = (1 - e^{\omega iy\xi})^m$. By the Littlewood-Paley decomposition, we have

$$\begin{aligned}
\omega_p^m(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N}) &= \sup_{|y|<2^{-\ell}} \left\| (\mathscr{F}^{-1} \rho_y) * (S_0 + \sum_{k=0}^\infty \triangle_k) \frac{\partial^N f}{\partial x_j^N} \right\|_p \\
&\leqslant \sup_{|y|<2^{-\ell}} \left\| (\mathscr{F}^{-1} \rho_y) * S_0 \frac{\partial^N f}{\partial x_j^N} \right\|_p + \sup_{|y|<2^{-\ell}} \sum_{k=0}^\infty \left\| (\mathscr{F}^{-1} \rho_y) * \triangle_k \frac{\partial^N f}{\partial x_j^N} \right\|_p \\
&\lesssim \sup_{|y|<2^{-\ell}} \left\| (\mathscr{F}^{-1} \rho_y) * S_0 f \right\|_p + \sup_{|y|<2^{-\ell}} \sum_{k=0}^\infty 2^{kN} \left\| (\mathscr{F}^{-1} \rho_y) * \triangle_k f \right\|_p.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sum_{j=1}^n \left(\sum_{\ell=-\infty}^{\infty} \left(2^{\ell(s-N)} \omega_p^m \left(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \right)^{1/r} \\
\lesssim &\Big(\sum_{\ell=-\infty}^{\infty} \Big(2^{\ell(s-N)} \sup_{|y|<2^{-\ell}} \|(\mathscr{F}^{-1} \rho_y) * S_0 f\|_p
\end{aligned}$$

$$+ \sup_{|y|<2^{-\ell}} \sum_{k=0}^{\infty} 2^{(\ell-k)(s-N)} 2^{ks} \|(\mathscr{F}^{-1}\rho_y) * \triangle_k f\|_p \Big)^r \Big)^{1/r}.$$

If we can prove that for all integers k

$$\|\mathscr{F}^{-1}\rho_y * S_0 f\|_p \lesssim \min(1, |y|^m) \|S_0 f\|_p, \tag{8.16}$$

and that

$$\|\mathscr{F}^{-1}\rho_y * \triangle_k f\|_p \lesssim \min(1, |y|^m 2^{mk}) \|\triangle_k f\|_p. \tag{8.17}$$

Then,

$$\begin{aligned}
&\sum_{j=1}^{n} \left(\sum_{\ell=-\infty}^{\infty} \left(2^{\ell(s-N)} \omega_p^m \left(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \right)^{1/r} \\
\lesssim &\Big(\sum_{\ell=-\infty}^{\infty} \Big(2^{\ell(s-N)} \sup_{|y|<2^{-\ell}} \min(1, |y|^m) \|S_0 f\|_p \\
&\qquad + \sup_{|y|<2^{-\ell}} \sum_{k=0}^{\infty} 2^{(\ell-k)(s-N)} 2^{ks} \min(1, |y|^m 2^{mk}) \|\triangle_k f\|_p \Big)^r \Big)^{1/r} \\
\lesssim &\Big(\sum_{\ell=-\infty}^{\infty} \Big(2^{\ell(s-N)} \min(1, 2^{-\ell m}) \|S_0 f\|_p \\
&\qquad + \sum_{k=0}^{\infty} 2^{(\ell-k)(s-N)} \min(1, 2^{-(\ell-k)m}) 2^{ks} \|\triangle_k f\|_p \Big)^r \Big)^{1/r} \\
\lesssim &\|(2^{k(s-N)} \min(1, 2^{-km})) * (\alpha_k)\|_{\ell^r} \\
\lesssim &\|(2^{k(s-N)} \min(1, 2^{-km}))\|_{\ell^1} \|(\alpha_k)\|_{\ell^r} \lesssim \|f\|_{B^s_{p,r}},
\end{aligned}$$

where the sequence $(\alpha_k)_{k=-\infty}^{\infty}$ with $\alpha_k = 2^{sk}\|\triangle_k f\|_p$ if $k \geqslant 0$, $\alpha_{-1} = \|S_0 f\|_p$ and $\alpha_k = 0$ if $k < -1$, and we have used the Young inequality for a convolution of two sequences. In addition, we have

$$\begin{aligned}
\|f\|_p \lesssim &\|S_0 f\|_p + \sum_{k=0}^{\infty} \|\triangle_k f\|_p \\
\lesssim &\|S_0 f\|_p + \left(\sum_{k=0}^{\infty} 2^{-skr'} \right)^{1/r'} \left(\sum_{k=0}^{\infty} (2^{sk}\|\triangle_k f\|_p)^r \right)^{1/r} \lesssim \|f\|_{B^s_{p,r}},
\end{aligned}$$

which implies the desired conclusion.

Now, we turn to prove (8.16) and (8.17). We only need to show $\rho_y \in M_p$ and $\rho_y(\cdot)\langle y, \cdot\rangle^{-m} \in M_p$ for $p \in [1, \infty]$ and

$$\|\rho_y\|_{M_p} \leqslant C, \quad \|\rho_y(\cdot)\langle y, \cdot\rangle^{-m}\|_{M_p} \leqslant C, \quad \forall y \neq 0. \tag{8.18}$$

In fact, from the definition of ρ_y, we get

$$\|\rho_y\|_{M_p} = \sup_{f \in \mathscr{S}} \frac{\|\mathscr{F}^{-1}\rho_y * f\|_p}{\|f\|_p} = \sup_{f \in \mathscr{S}} \frac{\| \sum_{k=0}^m C_m^k (-1)^k f(x+ky)\|_p}{\|f\|_p} \leqslant \sum_{k=0}^{m} C_m^k \leqslant C.$$

By Theorem 6.16, we have

$$\begin{aligned}
\|\rho_y(\xi)\langle y, \xi\rangle^{-m}\|_{M_p(\mathbb{R}^n)} =&\|(1 - e^{\omega i \langle y, \xi\rangle})^m \langle y, \xi\rangle^{-m}\|_{M_p(\mathbb{R}^n)} \\
=&\|((1 - e^{\omega i \eta})/\eta)^m\|_{M_p(\mathbb{R})}.
\end{aligned}$$

In view of Theorem 6.17, we only need to show $((1 - e^{\omega i\eta})/\eta)^m \in L^2(\mathbb{R})$ and $\partial_\eta((1 - e^{\omega i\eta})/\eta)^m \in L^2(\mathbb{R})$. We split the L^2 integral into two parts $|\eta| < 1$ and $|\eta| \geqslant 1$. For $|\eta| < 1$, we can use Taylor's expansion to get the bound of the integral. For $|\eta| \geqslant 1$, it is easy to get the bound by noting that $m \geqslant 1$ in view of the conditions $m + N > s$ and

$0 \leqslant N < s$. Thus, $\|((1-e^{\omega i\eta})/\eta)^m\|_{M_p(\mathbb{R})} \leqslant C$ by Theorem 6.17, which completes the proof of (8.18).

Similarly, we can prove

$$\|\langle y/|y|, \cdot\rangle^m \psi(\cdot)\|_{M_p} \leqslant C, \text{ and } \|\langle y/|y|, \cdot\rangle^m \varphi(\cdot)\|_{M_p} \leqslant C,$$

which implies

$$\|\langle y, \cdot\rangle^m \psi(\cdot)\|_{M_p} \leqslant C|y|^m, \quad \|\langle y, \cdot\rangle^m \varphi(2^{-k}\cdot)\|_{M_p} \leqslant C|y|^m 2^{mk}.$$

Thus clearly,

$$\begin{aligned}
&\|\mathscr{F}^{-1}\rho_y * S_0 f\|_p \lesssim \|S_0 f\|_p,\\
&\|\mathscr{F}^{-1}\rho_y * S_0 f\|_p = \|(\mathscr{F}^{-1}\rho_y(\xi)\langle y,\xi\rangle^{-m}) * \mathscr{F}^{-1}\langle y,\xi\rangle^m \mathscr{F} S_0 f\|_p \lesssim |y|^m \|S_0 f\|_p,
\end{aligned}$$

which yields (8.16). In the same way, we have

$$\begin{aligned}
&\|\mathscr{F}^{-1}\rho_y * \triangle_k f\|_p \lesssim \|\triangle_k f\|_p,\\
&\|\mathscr{F}^{-1}\rho_y * \triangle_k f\|_p = \|(\mathscr{F}^{-1}\rho_y(\xi)\langle y,\xi\rangle^{-m}) * \mathscr{F}^{-1}\langle y,\xi\rangle^m \mathscr{F} \triangle_k f\|_p \lesssim |y|^m 2^{mk} \|\triangle_k f\|_p,
\end{aligned}$$

which yields (8.17).

The converse inequality will follow if we can prove the estimate

$$\|\triangle_k f\|_p \leqslant C 2^{-Nk} \sum_{j=1}^n \left\| (\mathscr{F}^{-1}\rho_{jk}) * \frac{\partial^N f}{\partial_{x_j}^N} \right\|_p, \tag{8.19}$$

where $\rho_{jk} = \rho_{(2^{-k}e_j)}$ with e_j being the unit vector in the direction of the ξ_j-axis and ρ_y defined as the previous. In fact, if (8.19) is valid, we have, by noting $\psi \in M_1$, that

$$\begin{aligned}
\|f\|_{B^s_{p,r}} &\lesssim \|f\|_p + \left(\sum_{k=0}^\infty \left(2^{k(s-N)} \sum_{j=1}^n \left\| (\mathscr{F}^{-1}\rho_{jk}) * \frac{\partial^N f}{\partial_{x_j}^N} \right\|_p \right)^r \right)^{1/r}\\
&\lesssim \|f\|_p + \sum_{j=1}^n \left(\sum_{k=0}^\infty \left(2^{k(s-N)} \omega_p^m \left(2^{-k}, \frac{\partial^N f}{\partial_{x_j}^N} \right) \right)^r \right)^{1/r},
\end{aligned}$$

which implies the desired inequality.

In order to prove (8.19), we need the following lemma.

Lemma 8.23. *Assume that $n \geqslant 2$ and take φ as in (8.2). Then there exist functions $\chi_j \in \mathscr{S}(\mathbb{R}^n)$ $(1 \leqslant j \leqslant n)$, such that*

$$\begin{aligned}
&\sum_{j=1}^n \chi_j = 1 \quad \text{on } \operatorname{supp}\varphi \subset \left\{\xi : \alpha^{-1} \leqslant |\xi| \leqslant 2\alpha\right\},\\
&\operatorname{supp}\chi_j \subset \left\{\xi \in \mathbb{R}^n : |\xi_j| \geqslant (3\sqrt{n})^{-1}\right\}, \quad 1 \leqslant j \leqslant n.
\end{aligned}$$

Proof. Choose $\kappa \in \mathscr{S}(\mathbb{R})$ with $\operatorname{supp}\kappa = \{\xi \in \mathbb{R} : |\xi| \geqslant (3\sqrt{n})^{-1}\}$ and with positive values in the interior of $\operatorname{supp}\kappa$. Moreover, choose $\ell \in \mathscr{S}(\mathbb{R}^{n-1})$ with $\operatorname{supp}\ell = \left\{\xi \in \mathbb{R}^{n-1} : |\xi| \leqslant 3\right\}$ and positive in the interior. Writing $\bar{\xi}^j = (\xi_1, \cdots, \xi_{j-1}, \xi_{j+1}, \cdots, \xi_n)$ and

$$\chi_j(\xi) = \kappa(\xi_j)\ell(\bar{\xi}^j) / \sum_{j=1}^n \kappa(\xi_j)\ell(\bar{\xi}^j), \quad 1 \leqslant j \leqslant n,$$

where $\sum_{j=1}^n \kappa(\xi_j)\ell(\bar{\xi}^j) > 0$ on $\operatorname{supp}\varphi$, only routine verifications remain to complete the proof of the lemma. □

We now complete the proof of the theorem, i.e., we prove (8.19). By the previous lemma, we obtain the formula

$$\begin{aligned}\|\triangle_k f\|_p \lesssim & \sum_{j=1}^{n} \left\| \mathscr{F}^{-1}\rho_{jk}^{-m}\chi_j(2^{-k}\cdot)\xi_j^{-N}\varphi(2^{-k}\cdot) * \mathscr{F}^{-1}\rho_{jk}^{m}\mathscr{F}\frac{\partial^N f}{\partial_{x_j}^N}\right\|_p \\ \lesssim & 2^{-kN}\sum_{j=1}^{n} \left\| \mathscr{F}^{-1}\rho_{jk}^{-m}\chi_j(2^{-k}\cdot)(2^{-k}\xi_j)^{-N}\varphi(2^{-k}\cdot) * \mathscr{F}^{-1}\rho_{jk}^{m}\mathscr{F}\frac{\partial^N f}{\partial_{x_j}^N}\right\|_p \\ \lesssim & 2^{-kN}\sum_{j=1}^{n} \|\rho_{j0}^{-m}\chi_j\xi_j^{-N}\varphi\|_{M_p(\mathbb{R})}\left\| \mathscr{F}^{-1}\rho_{jk}^{m} * \frac{\partial^N f}{\partial_{x_j}^N}\right\|_p \\ \lesssim & 2^{-kN}\sum_{j=1}^{n} \left\| \mathscr{F}^{-1}\rho_{jk}^{m} * \frac{\partial^N f}{\partial_{x_j}^N}\right\|_p,\end{aligned}$$

since, by Theorem 6.16 and 6.17, we have

$$(1-e^{\omega i\xi_j})^{-m}\chi_j(\xi)\xi_j^{-N}\varphi(\xi) \in M_p,$$

for $1 \leqslant j \leqslant n$ and $1 \leqslant p \leqslant \infty$. □

Now, we give a relation between Hölder spaces and Besov spaces.

Corollary 8.24. *Let $s > 0$. Then we have $B^s_{\infty,\infty} = C^s$.*

Proof. By Theorem 8.22 with $p = r = \infty$ and $m = 2$, and Proposition 7.24, for $0 < s \leqslant 1$, we can take $N = 0$ and then

$$\|f\|_{B^s_{\infty,\infty}} \sim \|f\|_\infty + \sup_{t>0} t^{-s}\omega^2_\infty(t,f) = \|f\|_{C^s}.$$

By Proposition 7.25, we can extend to any $s > 1$. This completes the proof. □

Now we give a corollary which is very convenient for nonlinear estimates in PDEs.

Corollary 8.25. *Assume that $s > 0$ and $s \notin \mathbb{N}$. Let $1 \leqslant p$, $r \leqslant \infty$, then*

$$\|f\|_{B^s_{p,r}} \sim \|f\|_p + \sum_{j=1}^{n}\left(\int_0^\infty \left(t^{[s]-s}\sup_{|h|\leqslant t}\left\|\triangle_h\, \partial^{[s]}_{x_j} f\right\|_p\right)^r \frac{dt}{t}\right)^{1/r},$$

where $[s]$ denotes the integer part of the real number s and $\triangle_h$ denotes the first order difference operator.

Similarly, we can get a equivalent norm for the homogeneous Besov space.

Theorem 8.26. *Assume that $s > 0$, and let m and N be integers, such that $m + N > s$ and $0 \leqslant N < s$. Then, with $1 \leqslant p$, $r \leqslant \infty$,*

$$\|f\|_{\dot B^s_{p,r}} \sim \sum_{j=1}^{n}\left(\int_0^\infty \left(t^{N-s}\omega_p^m\left(t, \frac{\partial^N f}{\partial x_j^N}\right)\right)^r \frac{dt}{t}\right)^{1/r}.$$

In particular, if $s > 0$ and $s \notin \mathbb{N}$, then

$$\|f\|_{\dot B^s_{p,r}} \sim \sum_{j=1}^{n}\left(\int_0^\infty \left(t^{[s]-s}\sup_{|h|\leqslant t}\left\|\triangle_h\, \partial^{[s]}_{x_j} f\right\|_p\right)^r \frac{dt}{t}\right)^{1/r},$$

One of the following result is a straightforward consequence of Theorem 8.22 and Theorem 8.26, which indicates the relation between homogeneous and nonhomogeneous spaces.

Theorem 8.27. *Suppose that $f \in \mathscr{S}'$ and $0 \notin \operatorname{supp} \hat f$. Then*

$$f \in B^s_{p,r} \Leftrightarrow f \in \dot B^s_{p,r}, \quad \forall s \in \mathbb{R},\ 1 \leqslant p, r \leqslant \infty.$$

Moreover,

$$B^s_{p,r} = L^p \cap \dot B^s_{p,r}, \quad \forall s > 0,\ 1 \leqslant p, r \leqslant \infty,$$

$$B^s_{p,r} = L^p + \dot{B}^s_{p,r}, \quad \forall s < 0,\ 1 \leqslant p, r \leqslant \infty.$$

Proof. One can see [BL76, Chapter 6]. □

Now, we give the duality theorem:

Theorem 8.28 (The duality theorem)**.** *Let* $s \in \mathbb{R}$*. Then we have*

i) $(B^s_{p,r})' = B^{-s}_{p',r'}$*, if* $1 \leqslant p,\ r < \infty$*.*

ii) $(F^s_{p,r})' = F^{-s}_{p',r'}$*, if* $1 < p,\ r < \infty$*.*

Proof. Please read [BL76, Tri83] for details. □

Chapter 9
BMO Spaces

9.1 Functions of bounded mean oscillation and BMO spaces

Functions of bounded mean oscillation were introduced by F. John and L. Nirenberg [JN61], in connection with differential equations.

Definition 9.1. The *mean oscillation* of a locally integrable function f (i.e. a function belonging to $L^1_{loc}(\mathbb{R}^n)$) over a cube Q in $\mathbb{R}^n$ is defined as the following integral:

$$\tilde{f}_Q = \frac{1}{\mathfrak{m}(Q)} \int_Q |f(x) - f_Q| dx,$$

where f_Q is the average value of f on the cube Q, i.e.

$$f_Q = \frac{1}{\mathfrak{m}(Q)} \int_Q f(x) dx.$$

Definition 9.2. A *BMO function* is any function f belonging to $L^1_{loc}(\mathbb{R}^n)$ whose mean oscillation has a finite supremum over the set of all cubes[1] Q contained in $\mathbb{R}^n$. Let

$$M^{\#} f(x) = \sup_{x \in Q} \tilde{f}_Q,$$

be called the *maximal BMO function* or the *sharp maximal function*, $M^{\#}$ is called the *sharp maximal operator*. Then we denote the norm of f in this space by $\|f\|_{\mathrm{BMO}} = \|M^{\#} f\|_\infty$. The set

$$\mathrm{BMO}(\mathbb{R}^n) = \{f \in L^1_{loc}(\mathbb{R}^n) : \|f\|_{\mathrm{BMO}} < \infty\}$$

is called the *function space of bounded mean oscillation* or the *BMO space*.

Remark 9.3. 1) Note first that the null elements in the BMO norm are the constants, so that a function in BMO is, strictly speaking, defined only up to an additive constant. Although $\|\cdot\|_{\mathrm{BMO}}$ is only a seminorm, we will occasionally refer to it as a norm with no reason to cause confusion. Observe also that f would be still in BMO if the definition of $\tilde{f}_Q$ were extended to allow arbitrary constants c_Q in place of the mean values f_Q. Indeed, if $\|f\|_{\mathrm{BMO}} \leqslant A$ in our definition, we would then have $|c_Q - f_Q| \leqslant A$, and $\|f\|_{\mathrm{BMO}} \leqslant 2A$ follows. Similar reason shows that an equivalent definition of BMO arises if we replace the family of all cubes by, say, the family of all balls.

2) It is trivial that any bounded function is in BMO, i.e., $L^\infty \subset \mathrm{BMO}$. The converse is false. A simple example that already typifies some of the essential properties of BMO

[1] The use of cubes Q in $\mathbb{R}^n$ as the integration domains on which the mean oscillation is calculated, is not mandatory: Wiegerinck (2001) uses balls instead and, as remarked by Stein ([Ste93], p. 140), in doing so a perfectly equivalent of definition of functions of bounded mean oscillation arises.

is given by the function $f(x) = \ln|x|$. To check that this function is in BMO, note that the scaling transforms $f(x) \to f(\delta x)$, $\delta > 0$, map BMO functions to BMO functions, and in fact do not change their norms. Under these scaling, $\ln|x|$ is changed by at most an additive constant. Thus to verify $\|f\|_{\mathrm{BMO}} \leqslant C$ it suffices to check the alternative assertions

$$\int_Q |\ln|x||\,dx \leqslant \int_{|x|\leqslant 1+\sqrt{2}} |\ln|x||\,dx \leqslant C,$$

$$\begin{aligned}\int_Q |\ln|x| - \ln(1/|x_0|)|\,dx = \int_Q |\ln(|x|\,|x_0|)|\,dx \leqslant \int_{|x-x_0|\leqslant\sqrt{2}} |\ln(|x|\,|x_0|)|\,dx\\ \leqslant |x_0|^{-n}\int_{(\sqrt{2}-1)\leqslant|z|\leqslant(1+\sqrt{2})} |\ln|z||\,dz \leqslant C\end{aligned}$$

where Q is a cube of length 2 centered at x_0. The first inequality holds when $|x_0| \leqslant 1$; the second holds when $|x_0| \geqslant 1$. This proves that $\ln|x| \in \mathrm{BMO}$.

3) It is a simple but useful fact that the space of real-valued BMO function forms a lattice. If f and g belong to BMO, then so do $\max(f, g)$ and $\min(f, g)$. This follows from the observation that $|f|$ is in BMO whenever f is, which in turn is a consequence of the fact that $||f| - |f_Q|| \leqslant |f - f_Q|$.

An element of BMO is a function that is "nearly bounded". We shall see below the precise sense in which this is true. A useful related fact is the following, which describes the behavior of BMO functions at infinity.

Theorem 9.4. *Let $f \in \mathrm{BMO}$, then $f(x)(1 + |x|^{n+1})^{-1}$ is integrable on $\mathbb{R}^n$, and we have*

$$I = \int_{\mathbb{R}^n} \frac{|f(x) - f_{Q_0}|}{1 + |x|^{n+1}}\,dx \leqslant C\|f\|_{\mathrm{BMO}}, \tag{9.1}$$

where C is independent of f, and $Q_0 = Q(0, 1)$.

Proof. Let $Q_k = Q(0, 2^k)$, $S_k = Q_k \setminus Q_{k-1}$ for $k \in \mathbb{N}$, $S_0 = Q_0$, and

$$I_k = \int_{S_k} \frac{|f(x) - f_{Q_0}|}{1 + |x|^{n+1}}\,dx, \quad k \in \mathbb{N}_0.$$

Then, we have

$$I = I_0 + \sum_{k=1}^{\infty} I_k.$$

Since

$$I_0 = \int_{Q_0} \frac{|f(x) - f_{Q_0}|}{1 + |x|^{n+1}}\,dx \leqslant \int_{Q_0} |f(x) - f_{Q_0}|dx \leqslant \mathrm{m}(Q_0)\|f\|_{\mathrm{BMO}},$$

it suffices to prove $I_k \leqslant C_k\|f\|_{\mathrm{BMO}}$ and $\sum_k C_k < \infty$. For $x \in S_k$, we have $|x| > 2^{k-2}$ and then

$$1 + |x|^{n+1} > 1 + 2^{(k-2)(n+1)} > 4^{-(n+1)}2^{k(n+1)}.$$

Hence,

$$\begin{aligned}I_k \leqslant& 4^{n+1}2^{-k(n+1)}\int_{Q_k} |f(x) - f_{Q_0}|dx\\ \leqslant& 4^{n+1}2^{-k(n+1)}\int_{Q_k} [|f(x) - f_{Q_k}| + |f_{Q_k} - f_{Q_0}|]dx\\ \leqslant& 4^{n+1}2^{-k(n+1)}\mathrm{m}(Q_k)(\|f\|_{\mathrm{BMO}} + |f_{Q_k} - f_{Q_0}|)\\ =& 4^{n+1}2^{-k(n+1)}2^{kn}(\|f\|_{\mathrm{BMO}} + |f_{Q_k} - f_{Q_0}|).\end{aligned}$$

The second term can be controlled as follows:

$$\begin{aligned}
|f_{Q_k} - f_{Q_0}| \leqslant & \sum_{i=1}^{k} |f_{Q_i} - f_{Q_{i-1}}| \\
\leqslant & \sum_{i=1}^{k} \frac{1}{\mathrm{m}(Q_{i-1})} \int_{Q_{i-1}} |f(x) - f_{Q_i}| dx \\
\leqslant & \sum_{i=1}^{k} \frac{2^n}{\mathrm{m}(Q_i)} \int_{Q_i} |f(x) - f_{Q_i}| dx \\
\leqslant & k \cdot 2^n \|f\|_{\mathrm{BMO}}.
\end{aligned}$$

Therefore,

$$I_k \leqslant 4^{n+1} 2^{-k(n+1)} 2^{nk} (1 + k2^n) \|f\|_{\mathrm{BMO}},$$

where $C_k = Ck2^{-k}$ and $\sum_k C_k < \infty$. This completes the proof. □

9.2 The John-Nirenberg theorem

Having set down some basic facts about BMO, we now turn to a deeper property of BMO functions: their exponential integrability. We begin with a preliminary example.

Example 9.5. Let $f(x) = \ln|x|$, $I = (0, b)$, and

$$E_\alpha = \{x \in I : |\ln x - f_I| > \alpha\},$$

then we have

$$\begin{aligned}
E_\alpha = & \{x \in I : \ln x - f_I > \alpha\} \cup \{x \in I : \ln x - f_I < -\alpha\} \\
= & \{x \in I : x > e^{\alpha + f_I}\} \cup \{x \in I : x < e^{-\alpha + f_I}\}.
\end{aligned}$$

When α is large enough, the first set is an empty set and the second one is $(0, e^{-\alpha + f_I})$. Thus

$$\mathrm{m}(E_\alpha) = e^{-\alpha + f_I}.$$

By Jensen's inequality, we get

$$e^{f_I} \leqslant \frac{1}{|I|} \int_I e^{\ln t} dt = \frac{|I|}{2}.$$

Therefore,

$$\mathrm{m}(E_\alpha) \leqslant \frac{1}{2} |I| e^{-\alpha}.$$

Although the above relation is obtained from the function $\ln|x|$ over $(0, b)$, it indeed reflects an essential property for any BMO function in the BMO space.

Theorem 9.6 (The John-Nirenberg theorem). *For all $f \in \mathrm{BMO}(\mathbb{R}^n)$, for all cubes Q, and all $\alpha > 0$, we have*

$$\mathrm{m}(\{x \in Q : |f(x) - f_Q| > \alpha\}) \leqslant e\mathrm{m}(Q) e^{-A\alpha/\|f\|_{\mathrm{BMO}}} \tag{9.2}$$

with $A = (2^n e)^{-1}$.

Proof. Since inequality (9.2) is not altered when we multiply both f and α by the same constant, it suffices to assume that $\|f\|_{\mathrm{BMO}} = 1$. Let us now fix a closed cube Q and a constant $b > 1$ to be chosen later.

We apply the Calderón-Zygmund decomposition to the function $f - f_Q$ inside the cube Q. We introduce the following selection criterion for a cube R:

$$\frac{1}{\mathrm{m}(R)}\int_R |f(x)-f_Q|dx > b. \tag{9.3}$$

Since

$$\frac{1}{\mathrm{m}(Q)}\int_Q |f(x)-f_Q|dx \leqslant \|f\|_{\mathrm{BMO}} = 1 < b,$$

the cube Q does not satisfy the selection criterion (9.3). Set $Q^{(0)}=Q$ and subdivide $Q^{(0)}$ into 2^n equal closed subcubes of side length equal to half of the side length of Q. Select such a subcube R if it satisfies the selection criterion (9.3). Now subdivide all nonselected cubes into 2^n equal subcubes of half their side length by bisecting the sides, and select among these subcubes those that satisfy (9.3). Continue this process indefinitely. We obtain a countable collection of cubes $\{Q_j^{(1)}\}_j$ satisfying the following properties:

(A-1) The interior of every $Q_j^{(1)}$ is contained in $Q^{(0)}$.

(B-1) $b < \mathrm{m}(Q_j^{(1)})^{-1}\int_{Q_j^{(1)}} |f(x)-f_{Q^{(0)}}|dx \leqslant 2^n b$.

(C-1) $\left|f_{Q_j^{(1)}} - f_{Q^{(0)}}\right| \leqslant 2^n b$.

(D-1) $\sum_j \mathrm{m}(Q_j^{(1)}) \leqslant \frac{1}{b}\sum_j \int_{Q_j^{(1)}} |f(x)-f_{Q^{(0)}}|dx \leqslant \frac{1}{b}\mathrm{m}(Q^{(0)})$.

(E-1) $|f-f_{Q^{(0)}}| \leqslant b$ a.e. on the set $Q^{(0)}\setminus \cup_j Q_j^{(1)}$.

We call the cubes $Q_j^{(1)}$ of first generation. Note that the second inequality in (D-1) requires (B-1) and the fact that $Q^{(0)}$ does not satisfy (9.3). We now fix a selected first-generation cube $Q_j^{(1)}$ and we introduce the following selection criterion for a cube R:

$$\frac{1}{\mathrm{m}(R)}\int_R |f(x)-f_{Q_j^{(1)}}|dx > b. \tag{9.4}$$

Observe that $Q_j^{(1)}$ does not satisfy the selection criterion (9.4). We apply a similar Calderón - Zygmund decomposition to the function

$$f - f_{Q_j^{(1)}}$$

inside the cube $Q_j^{(1)}$. Subdivide $Q_j^{(1)}$ into 2n equal closed subcubes of side length equal to half of the side length of $Q_j^{(1)}$ by bisecting the sides, and select such a subcube R if it satisfies the selection criterion (9.4). Continue this process indefinitely. Also repeat this process for any other cube $Q_j^{(1)}$ of the first generation. We obtain a collection of cubes $\{Q_l^{(2)}\}_l$ of second generation each contained in some $Q_j^{(1)}$ such that versions of (A-1)-(E-1) are satisfied, with the superscript (2) replacing (1) and the superscript (1) replacing (0). We use the superscript (k) to denote the generation of the selected cubes.

For a fixed selected cube $Q_l^{(2)}$ of second generation, introduce the selection criterion

$$\frac{1}{\mathrm{m}(R)}\int_R |f(x)-f_{Q_l^{(2)}}|dx > b. \tag{9.5}$$

and repeat the previous process to obtain a collection of cubes of third generation inside $Q_l^{(2)}$. Repeat this procedure for any other cube $Q_l^{(2)}$ of the second generation. Denote by $\{Q_s^{(3)}\}_s$ the thus obtained collection of all cubes of the third generation.

We iterate this procedure indefinitely to obtain a doubly indexed family of cubes $Q_j^{(k)}$ satisfying the following properties:

(A-k) The interior of every $Q_j^{(k)}$ is contained in $Q_{j'}^{(k-1)}$.

(B-k) $b < \mathrm{m}(Q_j^{(k)})^{-1}\int_{Q_j^{(k)}} |f(x)-f_{Q_{j'}^{(k-1)}}|dx \leqslant 2^n b$.

(C-k) $\left|f_{Q_j^{(k)}} - f_{Q_{j'}^{(k-1)}}\right| \leqslant 2^n b$.

(D-k) $\sum_j \mathrm{m}(Q_j^{(k)}) \leqslant \frac{1}{b}\sum_{j'} \mathrm{m}(Q_{j'}^{(k-1)})$.

(E-k) $|f - f_{Q_{j'}^{(k-1)}}| \leqslant b$ a.e. on the set $Q_{j'}^{(k-1)} \setminus \cup_j Q_j^{(k)}$.

We prove (A-k)-(E-k). Note that (A-k) and the lower inequality in (B-k) are satisfied by construction. The upper inequality in (B-k) is a consequence of the fact that the unique cube $Q_{j_0}^{(k)}$ with double the side length of $Q_j^{(k)}$ that contains it was not selected in the process. Now (C-k) follows from the upper inequality in (B-k). (E-k) is a consequence of the Lebesgue differentiation theorem, since for every point in $Q_j^{(k-1)} \setminus \cup_j Q_j^{(k)}$ there is a sequence of cubes shrinking to it and the averages of

$$|f - f_{Q_{j'}^{(k-1)}}|$$

over all these cubes is at most b. It remains to prove (D-k). We have

$$\begin{aligned}\sum_j \mathrm{m}(Q_j^{(k)}) <& \frac{1}{b}\sum_j \int_{Q_j^{(k)}} |f(x) - f_{Q_{j'}^{(k-1)}}|dx\\ =&\frac{1}{b}\sum_{j'}\sum_{j \text{ corresp. to } j'} \int_{Q_j^{(k)}} |f(x) - f_{Q_{j'}^{(k-1)}}|dx\\ \leqslant&\frac{1}{b}\sum_{j'} \int_{Q_{j'}^{(k-1)}} |f(x) - f_{Q_{j'}^{(k-1)}}|dx\\ \leqslant&\frac{1}{b}\sum_{j'} \mathrm{m}(Q_{j'}^{(k-1)})\|f\|_{\mathrm{BMO}}\\ =&\frac{1}{b}\sum_{j'} \mathrm{m}(Q_{j'}^{(k-1)}).\end{aligned}$$

Having established (A-k)-(E-k) we turn to some consequences. Applying (D-k) successively $k-1$ times, we obtain

$$\sum_j \mathrm{m}(Q_j^{(k)}) \leqslant b^{-k}\mathrm{m}(Q^{(0)}). \tag{9.6}$$

For any fixed j we have that $|f_{Q_j^{(1)}} - f_{Q^{(0)}}| \leqslant 2^n b$ and $|f - f_{Q_j^{(1)}}| \leqslant b$ a.e. on $Q_j^{(1)} \setminus \cup_l Q_l^{(2)}$. This gives

$$|f - f_{Q^{(0)}}| \leqslant 2^n b + b \quad \text{ a.e. on } Q_j^{(1)} \setminus \cup_l Q_l^{(2)},$$

which, combined with (E-1), yields

$$|f - f_{Q^{(0)}}| \leqslant 2^n 2b \quad \text{ a.e. on } Q^{(0)} \setminus \cup_l Q_l^{(2)}. \tag{9.7}$$

For every fixed l, we also have that $|f - f_{Q_l^{(2)}}| \leqslant b$ a.e. on $Q_l^{(2)} \setminus \cup_s Q_s^{(3)}$, which combined with $|f_{Q_l^{(2)}} - f_{Q_{l'}^{(1)}}| \leqslant 2^n b$ and $|f_{Q_{l'}^{(1)}} - f_{Q^{(0)}}| \leqslant 2^n b$ yields

$$|f - f_{Q^{(0)}}| \leqslant 2^n 3b \quad \text{ a.e. on } Q_l^{(2)} \setminus \cup_s Q_s^{(3)}.$$

In view of (9.7), the same estimate is valid on $Q^{(0)} \setminus \cup_s Q_s^{(3)}$. Continuing this reasoning, we obtain by induction that for all $k \geqslant 1$ we have

$$|f - f_{Q^{(0)}}| \leqslant 2^n kb \quad \text{ a.e. on } Q^{(0)} \setminus \cup_s Q_s^{(k)}. \tag{9.8}$$

This proves the almost everywhere inclusion

$$\{x \in Q : |f(x) - f_Q| > 2^n kb\} \subset \cup_j Q_j^{(k)}$$

for all $k = 1, 2, 3, \cdots$. (This also holds when $k = 0$.) We now use (9.6) and (9.8) to prove (9.2). We fix an $\alpha > 0$. If

$$2^n kb < \alpha \leqslant 2^n(k+1)b$$

for some $k \geqslant 0$, then

$$\begin{aligned}\mathbb{m}(\{x\in Q:|f(x)-f_Q|>\alpha\}) \leqslant & \mathbb{m}(\{x\in Q:|f(x)-f_Q|>2^n kb\})\\ \leqslant & \sum_j \mathbb{m}(Q_j^{(k)}) \leqslant \frac{1}{b^k}\mathbb{m}(Q^{(0)})\\ = & \mathbb{m}(Q)e^{-k\ln b}\\ \leqslant & \mathbb{m}(Q)be^{-\alpha\ln b/(2^n b)},\end{aligned}$$

since $-k \leqslant -1 - \frac{\alpha}{2^n b}$. Choosing $b = e > 1$ yields (9.2). □

Having proved the important distribution inequality (9.2), we are now in a position to deduce from it a few corollaries.

Corollary 9.7. *Every BMO function is exponentially integrable over any cube. More precisely, for any $\gamma < 1/(2^n e)$, for all $f \in \mathrm{BMO}(\mathbb{R}^n)$, and all cubes Q we have*

$$\frac{1}{\mathbb{m}(Q)}\int_Q e^{\gamma|f(x)-f_Q|/\|f\|_{\mathrm{BMO}}}dx \leqslant 1 + \frac{2^n e^2\gamma}{1-2^n e\gamma}.$$

Proof. Using identity (2.13) with $\varphi(t) = e^t - 1$, we write

$$\frac{1}{\mathbb{m}(Q)}\int_Q e^h dx = 1 + \frac{1}{\mathbb{m}(Q)}\int_Q (e^h-1)dx = 1 + \frac{1}{\mathbb{m}(Q)}\int_0^\infty e^\alpha \mathbb{m}(\{x\in Q:|h(x)|>\alpha\})d\alpha$$

for a measurable function h. Then we take $h = \gamma|f(x)-f_Q|/\|f\|_{\mathrm{BMO}}$ and we use inequality (9.2) with $\gamma < A = (2^n e)^{-1}$ to obtain

$$\frac{1}{\mathbb{m}(Q)}\int_Q e^{\gamma|f(x)-f_Q|/\|f\|_{\mathrm{BMO}}}dx \leqslant \int_0^\infty e^\alpha e e^{-A(\frac{\alpha}{\gamma}\|f\|_{\mathrm{BMO}})/\|f\|_{\mathrm{BMO}}}d\alpha = C_{n,\gamma},$$

where $C_{n,\gamma}$ is a unit less than the constant in the statement of the inequality. □

Another important corollary is the following.

Corollary 9.8. *For all $0 < p < \infty$, there exits a finite constant $B_{p,n}$ such that*

$$\sup_Q \left(\frac{1}{\mathbb{m}(Q)}\int_Q |f(x)-f_Q|^p dx\right)^{1/p} \leqslant B_{p,n}\|f\|_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{9.9}$$

Proof. This result can be obtained from the one in the preceding corollary or directly in the following way:

$$\begin{aligned}\frac{1}{\mathbb{m}(Q)}\int_Q |f(x)-f_Q|^p dx = & \frac{p}{\mathbb{m}(Q)}\int_0^\infty \alpha^{p-1}\mathbb{m}(\{x\in Q:|f(x)-f_Q|>\alpha\})d\alpha\\ \leqslant & \frac{p}{\mathbb{m}(Q)}e\mathbb{m}(Q)\int_0^\infty \alpha^{p-1}e^{-A\alpha/\|f\|_{\mathrm{BMO}}}d\alpha\\ = p\Gamma(p)\frac{e}{A^p}\|f\|_{\mathrm{BMO}}^p,&\end{aligned}$$

where $A = (2^n e)^{-1}$. Setting $B_{p,n} = (p\Gamma(p)\frac{e}{A^p})^{1/p} = (p\Gamma(p))^{1/p}e^{1+1/p}2^n$, we conclude the proof. □

Since the inequality in Corollary 9.8 can be reversed when $p > 1$ via Hölder's inequality, we obtain the following important L^p characterization of BMO norms.

Corollary 9.9. *For all $1 < p < \infty$, we have*

$$\sup_Q \left(\frac{1}{\mathbb{m}(Q)}\int_Q |f(x)-f_Q|^p dx\right)^{1/p} \sim \|f\|_{\mathrm{BMO}}.$$

Proof. Since the proof is obvious, we omit the details. □

9.3 The sharp maximal function theorem and singular integral of type (L^∞, BMO)

Lemma 9.10. *Let f be a nonnegative integrable function on $\mathbb{R}^n$, $t(\alpha) = \sum_{k=1}^\infty \mathrm{m}(Q_k^\alpha)$ where $\{Q_k^\alpha\}$ is the cubes sequence obtained by the Caldrón-Zygmund decomposition of f at level α. Then for any $A > 0$, we have*

$$t(\alpha) \leqslant (M^\# f)_*(\alpha/A) + \frac{2}{A} t(2^{-n-1}\alpha).$$

Proof. Let $\mu = 2^{-n-1}\alpha$, $\{Q_j^\mu\}$ is the Caldrón-Zygmund decomposition at level μ, then every Q_k^α is contained in a certain Q_j^μ. Thus, there are two possibilities between Q_j^μ and the set $\{x \in \mathbb{R}^n : M^\# f(x) > \alpha/A\}$:

(i) $Q_j^\mu \subset \{x \in \mathbb{R}^n : M^\# f(x) > \alpha/A\}$;

(ii) $Q_j^\mu \not\subset \{x \in \mathbb{R}^n : M^\# f(x) > \alpha/A\}$.

For the case (i), we have

$$\sum_{\{j: Q_j^\mu \text{ in (i)}\}} \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha) \leqslant (M^\# f)_*(\alpha/A). \tag{9.10}$$

For the case (ii), since some points of Q_j^μ don't belong to the set $\{x \in \mathbb{R}^n : M^\# f(x) > \alpha/A\}$, we get

$$\frac{1}{\mathrm{m}(Q_j^\mu)} \int_{Q_j^\mu} |f(x) - f_{Q_j^\mu}| dx \leqslant \frac{\alpha}{A}.$$

Noticing that $f_{Q_j^\mu} \leqslant 2^n \mu = \alpha/2$, we have

$$\begin{aligned}
\alpha \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha) \leqslant & \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \int_{Q_k^\alpha} |f(x)| dx \\
\leqslant & \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \int_{Q_k^\alpha} |f(x) - f_{Q_j^\mu}| dx + |f_{Q_j^\mu}| \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha) \\
\leqslant & \int_{Q_j^\mu} |f(x) - f_{Q_j^\mu}| dx + \frac{\alpha}{2} \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha) \\
\leqslant & \frac{\alpha}{A} \mathrm{m}(Q_j^\mu) + \frac{\alpha}{2} \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha),
\end{aligned}$$

which implies that

$$\sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha) \leqslant \frac{2}{A} \mathrm{m}(Q_j^\mu).$$

Hence, taking the sum for those Q_j^μ in the case (ii) yields

$$\sum_{\{j: Q_j^\mu \text{ in (ii)}\}} \sum_{\{k: Q_k^\alpha \subset Q_j^\mu\}} \mathrm{m}(Q_k^\alpha) \leqslant \frac{2}{A} \sum_{\{j: Q_j^\mu \text{ in (ii)}\}} \mathrm{m}(Q_j^\mu) \leqslant \frac{2}{A} t(\mu) = \frac{2}{A} t(2^{-n-1}\alpha).$$

Combining with (9.10), we obtain the desired result. □

Theorem 9.11 (The sharp maximal theorem)**.** *Let $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, then we have*

$$\|M^\# f\|_p \sim \|f\|_p, \tag{9.11}$$

where the hidden constants depend only on p and n.

Proof. Since

$$M^{\#}f(x) = \sup_{x\in Q}\frac{1}{\mathfrak{m}(Q)}\int_Q |f(x)-f_Q|dx \leqslant \sup_{x\in Q}\frac{1}{\mathfrak{m}(Q)}\int_Q |f(x)|dx + |f_Q| \leqslant 2M'f(x),$$

we have $\|M^{\#}f\|_p \leqslant C\|f\|_p$ by the Hardy-Littlewood maximal function theorem. Since $\|f\|_p \leqslant \|M'f\|_p$, thus we only need to prove that

$$\|M'f\|_p \leqslant C\|M^{\#}f\|_p, \quad 1<p<\infty. \tag{9.12}$$

In fact, from Lemma 3.22 (which can be extended to any integrable functions),changes of variables, we have

$$\begin{aligned}
\|M'f\|_p^p = \int_{\mathbb{R}^n} |M'f(x)|^p dx = & p\int_0^\infty \alpha^{p-1}(M'f)_*(\alpha)d\alpha \\
\leqslant & p\int_0^\infty \alpha^{p-1}2^n\sum_{k=1}^\infty \mathfrak{m}(Q_k^{7^{-n}\alpha})d\alpha \\
= & 2^n7^{np}p\int_0^\infty \beta^{p-1}\sum_{k=1}^\infty \mathfrak{m}(Q_k^\beta)d\beta \\
= & 2^n7^{np}p\int_0^\infty \beta^{p-1}t(\beta)d\beta.
\end{aligned}$$

By Lemma 9.10 and changes of variables, it yields

$$\begin{aligned}
\int_0^\infty \beta^{p-1}t(\beta)d\beta \leqslant & \int_0^\infty \beta^{p-1}(M^{\#}f)_*(\beta/A)d\beta + \frac{2}{A}\int_0^\infty \beta^{p-1}t(2^{-n-1}\beta)d\beta \\
= & A^p\int_0^\infty \alpha^{p-1}(M^{\#}f)_*(\alpha)d\alpha + \frac{2}{A}2^{(n+1)p}\int_0^\infty \alpha^{p-1}t(\alpha)d\alpha,
\end{aligned}$$

which implies, by taking $A = 2^{(n+1)p+2}$, that

$$\int_0^\infty \beta^{p-1}t(\beta)d\beta \leqslant 2^{(n+1)p^2+2p+1}\int_0^\infty \alpha^{p-1}(M^{\#}f)_*(\alpha)d\alpha.$$

Thus, we obtain

$$\|M'f\|_p^p \leqslant 7^{np}2^{(n+1)p^2+2p+1+n}\|M^{\#}f\|_p^p.$$

Therefore, we complete the proof. □

We continue by proving an interpolation result in which the space L^∞ is replaced by BMO. The sharp maximal function plays a key role in the following theorem.

Theorem 9.12. *Let $1 \leqslant p_0 < \infty$. Let T be a linear operator that maps $L^{p_0}(\mathbb{R}^n)$ into $L^{p_0}(\mathbb{R}^n)$ with bound A_0 and $L^\infty(\mathbb{R}^n)$ into* BMO$(\mathbb{R}^n)$ *with bound A_1. Then for all p with $p_0 < p < \infty$ there is a constant C_{n,p,p_0} such that for all $f \in L^p$ we have*

$$\|Tf\|_p \leqslant C_{n,p,p_0}\|f\|_p.$$

Proof. We consider the operator

$$S(f) = M^{\#}(Tf)$$

define for $f \in L^{p_0} + L^\infty$. It is easy to see that S is a sublinear operator. We prove that S maps L^{p_0} into itself and L^∞ into itself:

$$\begin{aligned}
\|S(f)\|_{p_0} = & \|M^{\#}(Tf)\|_{p_0} \leqslant C\|Tf\|_{p_0} \leqslant CA_0\|f\|_{p_0}, \\
\|S(f)\|_\infty = & \|M^{\#}(Tf)\|_\infty = \|Tf\|_{\mathrm{BMO}} \leqslant A_1\|f\|_\infty.
\end{aligned}$$

Interpolating between these estimates using the Marcinkiewicz interpolation theorem, we obtain the estimate

$$\|M^{\#}(Tf)\|_p = \|S(f)\|_p \leqslant C'_{n,p,p_0}A_0^{p_0/p}A_1^{1-p_0/p}\|f\|_p$$

for all $p_0 < p < \infty$. Then by Theorem 9.11, we have

$$\|Tf\|_p \leqslant C'_{n,p,p_0} A_0^{p_0/p} A_1^{1-p_0/p} \|f\|_p$$

for all $f \in L^p(\mathbb{R}^n)$. □

Now, we consider singular integral operators of type (L^∞, BMO).

Theorem 9.13. *Let $K \in L^1(\mathbb{R}^n)$ satisfy*

(i) $|\hat{K}(\xi)| \leqslant A$, $\xi \in \mathbb{R}^n$;

(ii) For all $y \neq 0$, it holds

$$\int_{|x| \geqslant 2|y|} |K(x-y) - K(x)| dx \leqslant A.$$

Then $Tf(x) = K * f(x)$ satisfies

$$\|Tf\|_{\mathrm{BMO}} \leqslant C \|f\|_\infty,$$

where C depends only on n and A.

Remark 9.14. If we apply the Young inequality, the result of the theorem holds for $C = \|K\|_1$. But the key point is that C is independent of $\|K\|_1$.

Proof. Assume that $\|f\|_\infty \leqslant 1$ without loss of generality. Let $Q = Q(0, r)$, $B = B(0, 2\sqrt{n}r)$, and $f(x) = f_1(x) + f_2(x)$ where

$$f_1(x) = f(x)\chi_B(x), \quad f_2(x) = f(x)\chi_{\mathbb{R}^n \setminus B}(x).$$

Denote

$$Tf(x) = K * f_1(x) + K * f_2(x) = u_1(x) + u_2(x).$$

Since $f_1 \in L^2(\mathbb{R}^n)$, it follows that

$$\|u_1\|_2 = \|K * f_1\|_2 = C\|\hat{K}\hat{f}_1\|_2 \leqslant A\|f_1\|_2 \leqslant C_n A \mathrm{m}(Q)^{1/2},$$

and then

$$\int_Q |u_1(x)| dx \leqslant \|u_1\|_2 \mathrm{m}(Q)^{1/2} \leqslant C_n A \mathrm{m}(Q).$$

On the other hand, taking

$$a_Q = \int_{\mathbb{R}^n} K(-y) f_2(y) dy,$$

we have

$$u_2(x) - a_Q = \int_{\mathbb{R}^n} [K(x-y) - K(-y)] f_2(y) dy.$$

Thus, by the condition (ii), it implies

$$\begin{aligned}
\int_Q |u_2(x) - a_Q| dx &\leqslant \int_Q \int_{\mathbb{R}^n \setminus B} |K(x-y) - K(-y)| dy dx \\
&\leqslant \int_Q \int_{|y| > 2|x|} |K(x-y) - K(-y)| dy dx \leqslant A \mathrm{m}(Q).
\end{aligned}$$

Therefore, combining the above estimates, we obtain

$$\begin{aligned}
\frac{1}{\mathrm{m}(Q)} \int_Q |Tf(x) - a_Q| dx \leqslant & \frac{1}{\mathrm{m}(Q)} \int_Q |u_1(x)| dx + \frac{1}{\mathrm{m}(Q)} \int_Q |u_2(x) - a_Q| dx \\
\leqslant & (C_n + 1)A.
\end{aligned}$$

This completes the proof. □

References

Abe11. Helmut Abels. Short lecture notes: Interpolation theory and function spaces. May 2011.

BCD11. Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343 of *GMW*. Springer-Verlag, Berlin Heidelberg, 2011.

Bec75. William Beckner. Inequalities in Fourier analysis on R^n. *Proc. Nat. Acad. Sci. U.S.A.*, 72:638–641, 1975.

BL76. Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

Din07. Guanggui Ding. *New talk of functional analysis*. Science press, Beijing, 2007.

Duo01. Javier Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.

Eva98. Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

GR. I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Seventh edition.

Gra04. Loukas Grafakos. *Classical and modern Fourier analysis*. Pearson Education, Inc., Upper Saddle River, NJ, 2004.

GT01. David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

HLP88. G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.

HMOW10. H. Hajaiej, L. Molinet, T. Ozawa, and Baoxiang Wang. Necessary and sufficient conditions for the fractional Gagliardo-Nirenberg inequalities and applications to the Navier-Stokes and generalized boson equations. *arXiv:1004.4287*, 2010.

HW64. Richard A. Hunt and Guido Weiss. The Marcinkiewicz interpolation theorem. *Proc. Amer. Math. Soc.*, 15:996–998, 1964.

JN61. F. John and L. Nirenberg. On functions of bounded mean oscillation. *Comm. Pure Appl. Math.*, 14:415–426, 1961.

Kri02. Erik Kristiansson. Decreasing rearrangement and Lorentz $L(p, q)$ spaces. Master's thesis, Luleå University of Technology, 2002. http://bioinformatics.zool.gu.se/~erikkr/MasterThesis.pdf.

Mia04. Changxing Miao. *Harmonic analysis and its application in PDEs*, volume 89 of *Fundamental series in mordon Mathematics*. Science Press, Beijing, 2nd edition, 2004.

Rud87. Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

SS03. Elias M. Stein and Rami Shakarchi. *Complex analysis*. Princeton Lectures in Analysis, II. Princeton University Press, Princeton, NJ, 2003.

Ste70. Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

Ste93. Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

SW71. Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

Tri83. Hans Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983.

WHHG11. Baoxiang Wang, Zhaohui Huo, Chengchun Hao, and Zihua Guo. *Harmonic analysis method for nonlinear evolution equations*, volume I. World Scientific Publishing Co. Pte. Ltd., 2011.

Index