GRADUATE LECTURES IN ANALYSIS

LECTURE NOTES ON

Harmonic Analysis

调和分析讲义

HAO Chengchun/郝成春

Lecture Notes on

Harmonic Analysis

HAO Chengchun

Institute of Mathematics, AMSS
Chinese Academy of Sciences

Email: hcc@amss.ac.cn

Preface

These lecture notes are intended primarily as text for a graduate-level analysis course (Harmonic Analysis I, II) taught at the University of Chinese Academy of Sciences (UCAS) with 80 class periods (50 minutes) a semester from 2020 to 2022.

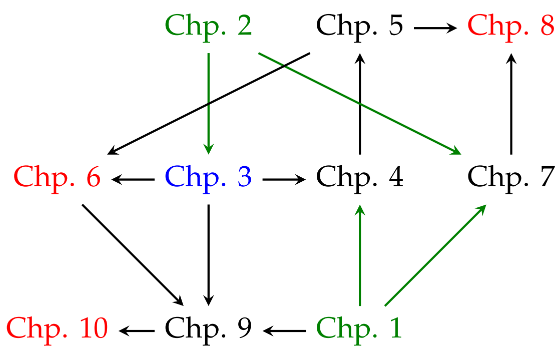
In these notes, our aim is to give some fundamental theory in harmonic analysis, including some basic theories of real analysis, singular integrals of convolution-type, singular integrals of non-convolution types, some function spaces and paraproducts.

Most of the materials in these notes are borrowed from books [Ste70; SW71; Gra14a; Gra14b; BL76a; Fol99; Wan+11; BCD11; MWZ12] and some online lecture notes [AB12; Bro15; Mur19; Tao06] with some necessary modifications and more details.

The prerequisites are some basic knowledge of real analysis (a summary of some relevant facts is provided in chapter 0) and functional analysis and some complex analysis.

There are some exercises at the end of each chapter, some of which have been used in some proofs in the text.

The main dependencies among the chapters are indicated in the following diagram.



Hao Chengchun
Beijing
December 11, 2022

Contents

0 Overview of Real Analysis and Functional Analysis

In this chapter, the purpose is to establish the notation and terminology that will be used throughout the book and to concisely present a few results from analysis that will be needed later. All results can be found in typical real analysis books, e.g., [Fol99; Fol09; SS05].

§ 0.1 Measure theory

The notation for the fundamental number systems is as follows:

$$\begin{aligned}
\mathbb{N} &= \text{the set of positive integers (not including zero)},\\
\mathbb{N}_0 &= \text{the set of nonnegative integers (including zero)},\\
\mathbb{Z} &= \text{the set of integers},\\
\mathbb{Q} &= \text{the set of rational numbers},\\
\mathbb{R} &= \text{the set of real numbers},\\
\mathbb{C} &= \text{the set of complex numbers}.
\end{aligned}$$

The words "family" and "collection" will be used synonymously with "set", usually to avoid phrases such as "set of sets" in set theory. The empty set is denoted by $\varnothing$, and the family of all subsets of a set X is denoted by $\mathcal{P}(X) = \{E : E \subset X\}$. Here and elsewhere, the inclusion sign "$\subset$" is interpreted in the weak sense; that is, the assertion "$E \subset X$" includes the possibility that $E = X$.

Let X be a nonempty set. An ***algebra*** of sets on X is a nonempty collection of subsets of X that is closed under finite unions and complements. A ***σ-algebra*** is an algebra that is closed under countable unions. Moreover, if $\mathcal{A}$ is an algebra, then $\varnothing \in \mathcal{A}$ and $X \in \mathcal{A}$, for if $E \in \mathcal{A}$ we have $\varnothing = E \cap E^c$ and $X = E \cup E^c$ where the complement E^c of a set E (in X) is defined by $E^c = X \setminus E = \{x \in X : x \notin E\}$. If X is a topological space, the σ-algebra generated by the family of open sets in X is called the Borel σ-algebra on X; is denoted by $\mathcal{B}_X$, and its elements are called ***Borel sets***.

Let X be a set equipped with a σ-algebra $\mathcal{M}$. A ***measure*** on $\mathcal{M}$ (or on $(X, \mathcal{M})$) is a function $\mu : \mathcal{M} \to [0, \infty]$ such that

i) $\mu(\varnothing) = 0$,

ii) (Countable additivity) if $\{E_j\}_{j=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{M}$, then

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j).$$

Here are some more basic properties of measures:

(i) *Finite additivity*: If $\{E_j\}_1^n$ is a finite collection of disjoint sets in $\mathcal{M}$, then

$$\mu\left(\bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \mu(E_j).$$

(ii) *Monotonicity*: If $E, F \in \mathcal{M}$ and $E \subset F$, then $\mu(E) \leqslant \mu(F)$.

(iii) *Subadditivity*: If $\{E_j\}_1^{\infty} \subset \mathcal{M}$, then $\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j)$.

(iv) *Continuity from below*: If $\{E_j\}_1^{\infty} \subset \mathcal{M}$ and $E_j \subset E_{j+1}$ for all j, then

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \lim_{j\to\infty} \mu(E_j).$$

(v) *Continuity from above*: If $\{E_j\}_1^{\infty} \subset \mathcal{M}$ and $E_{j+1} \subset E_j$ for all j, and $\mu(E_1) < \infty$, then $\mu\left(\bigcap_{j=1}^{\infty} E_j\right) = \lim_{j\to\infty} \mu(E_j)$.

If X is a set and $\mathcal{M} \subset \mathcal{P}(X)$ is a σ-algebra, $(X, \mathcal{M})$ is called a ***measurable space*** and the sets in $\mathcal{M}$ are called ***measurable sets***. If μ is a measure on $(X, \mathcal{M})$, then $(X, \mathcal{M}, \mu)$ is called a ***measure space***, where a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ is called a ***null set***. A measure whose domain includes all subsets of null sets is called ***complete***. Completeness can sometimes obviate annoying technical points, and it can always be achieved by enlarging the domain of μ; indeed, there is a unique extension of μ to a complete measure on the completion of $\mathcal{M}$ with respect to μ.

There is a unique measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that the measure of each interval is its length, and for $n > 1$, there is a unique measure on $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ such that the measure of the Cartesian product of n intervals is the product of their lengths. The completions of these measures are called the ***Lebesgue measure*** on $\mathbb{R}$ and $\mathbb{R}^n$, respectively. Its domain is called the class of ***Lebesgue measurable*** sets, and we denote it by $\mathcal{L}$.

A measure space $(X, \mathcal{M}, \mu)$ is called ***finite*** if $\mu(X) < \infty$, and is called ***σ-finite*** if $X = \bigcup_{j=1}^{\infty} E_j$ where $E_j \in \mathcal{M}$ and $\mu(E_j) < \infty$ for all j. If for each $E \in \mathcal{M}$ with $\mu(E) = \infty$ there exists $F \in \mathcal{M}$ with $F \subset E$ and $0 < \mu(F) < \infty$, μ is called ***semifinite***. Every σ-finite measure is semifinite, but not conversely.

If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable spaces, a mapping $f : X \to Y$ is called ***$(\mathcal{M}, \mathcal{N})$-measurable***, or just ***measurable*** when $\mathcal{M}$ and $\mathcal{N}$ are understood, if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.

If $(X, \mathcal{M})$ is a measurable space, a real- or complex-valued function f on X is called $\mathcal{M}$-measurable, or just measurable, if it is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$ or $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ measurable. $\mathcal{B}_{\mathbb{R}}$ or $\mathcal{B}_{\mathbb{C}}$ is always understood as the σ-algebra on the range space. In particular, $f : \mathbb{R} \to \mathbb{C}$ is ***Lebesgue*** (resp. ***Borel***) ***measurable*** if it is $(\mathcal{L}, \mathcal{B}_{\mathbb{C}})$ (resp. $(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{C}})$) measurable; likewise for $f : \mathbb{R} \to \mathbb{R}$.

We now recall the basic building blocks of the theory of integration, that is, the so-called "simple functions". Here are the definitions. Suppose that $(X, \mathcal{M})$ is a measurable space. If $E \subset X$, the ***characteristic function*** $\chi_E : X \to \{0, 1\}$ of E is defined by

$$\chi_E(x) = \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

It is easy to check that χ_E is measurable ***iff*** (the abbreviation of "if and only if") $E \in \mathcal{M}$. A ***simple function*** on X is a finite linear combination, with complex coefficients, of characteristic functions of sets in $\mathcal{M}$. (We do not allow simple functions to assume the values $\pm\infty$.) Equivalently, $f : X \to \mathbb{C}$ is simple iff f is measurable and the range of f is a finite subset of $\mathbb{C}$. Indeed, we have

$$f = \sum_{j=1}^{N} a_j \chi_{E_j}, \text{ where } E_j = f^{-1}(\{a_j\}) \text{ and } \mathrm{rangle}(f) = \{a_1, \cdots, a_N\}.$$

We call this the ***standard representation*** of f. It exhibits f as a linear combination, with distinct coefficients, of characteristic functions of disjoint sets whose union is X. Now, we recall that arbitrary measurable functions can be approximated in a nice way by simple functions.

Theorem 0.1 ([Fol99, Theorem 2.10]). *Let $(X, \mathcal{M})$ be a measurable space.*

(i) *If $f : X \to [0, \infty]$ is measurable, there is a sequence $\{\phi_n\}$ of simple functions such that $0 \leqslant \phi_1 \leqslant \phi_2 \leqslant \cdots \leqslant f$, $\phi_n \to f$ pointwise, and $\phi_n \to f$ uniformly on any set on which f is bounded.*

(ii) *If $f : X \to \mathbb{C}$ is measurable, there is a sequence $\{\phi_n\}$ of simple functions such that $0 \leqslant |\phi_1| \leqslant |\phi_2| \leqslant \cdots \leqslant |f|$, $\phi_n \to f$ pointwise, and $\phi_n \to f$ uniformly on any set on which f is bounded.*

Next, we recall the notion of convergence in measure.

Definition 0.2. Let f, f_n, $n = 1, 2, \cdots$, be measurable functions on the measure space $(X, \mathcal{M}, \mu)$. The sequence $\{f_n\}$ is said to ***converge in measure*** to f, denoted by $f_n \xrightarrow{\mu} f$, if for all $\varepsilon > 0$, there exists an $n_0 \in \mathbb{Z}^+$ such that

$$n > n_0 \Longrightarrow \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) < \varepsilon. \tag{0.1}$$

Remark 0.3. The above definition is equivalent to the following statement:

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0, \quad \forall \varepsilon > 0. \tag{0.2}$$

Clearly, (0.2) implies (0.1). To see the converse, given $\varepsilon > 0$, pick $0 < \delta < \varepsilon$ and apply (0.1) for this δ. There exists an $n_0 \in \mathbb{Z}^+$ such that

$$\mu(\{x \in X : |f_n(x) - f(x)| > \delta\}) < \delta$$

holds for $n > n_0$. Since

$$\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \leqslant \mu(\{x \in X : |f_n(x) - f(x)| > \delta\}),$$

we conclude that

$$\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) < \delta$$

for all $n > n_0$. Let $n \to \infty$, and we deduce that

$$\limsup_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \leqslant \delta. \tag{0.3}$$

Since (0.3) holds for all $\delta \in (0, \varepsilon)$, (0.2) follows by letting $\delta \to 0$.

Theorem 0.4 (Riesz Theorem, cf. [Fol99, Theorem 2.30]). *Let $\{f_n\}$ ad f be complex-valued measurable functions on a measure space $(X, \mathcal{M}, \mu)$ and suppose that f_n converges to f in measure. Then, some subsequence of $\{f_n\}$ converges to f μ-a.e.*

Proof. Since f_n converges to f in measure, we have by definition that for any $k \in \mathbb{N}$, there exists n_k such that

$$\mu(A_k) < 2^{-k} \tag{0.4}$$

and such that $n_1 < n_2 < \cdots < n_k < \cdots$, where

$$A_k = \left\{x \in X : |f_{n_k}(x) - f(x)| > 2^{-k}\right\}.$$

It follows from (0.4) that

$$\mu\left(\bigcup_{k=m}^{\infty} A_k\right) \leqslant \sum_{k=m}^{\infty} \mu(A_k) \leqslant \sum_{k=m}^{\infty} 2^{-k} = 2^{1-m}, \quad \forall m \in \mathbb{N}, \tag{0.5}$$

which implies that

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leqslant 1 < \infty. \tag{0.6}$$

Using (0.5) and (0.6), we conclude that the sequence of the measures of the sets $\left\{\bigcup_{k=m}^{\infty} A_k\right\}_{m=1}^{\infty}$ converges as $m \to \infty$ to

$$\mu\left(\bigcap_{m=1}^{\infty}\bigcup_{k=m}^{\infty} A_k\right) = 0. \tag{0.7}$$

It is clear that the null set in (0.7) contains the set of all $x \in X$ for which

$f_{n_k}(x)$ does not converge to $f(x)$. ❑

Theorem 0.5 (Inner regularity of Lebesgue measure). *If A is Lebesgue measurable, then $\mu(A) = \sup\{\mu(K) : K \subset A, K \text{ compact}\}$.*

§ 0.2 Integration

In this section, we develop the theory of integration of real- or complex-valued functions on a measure space.

If $(X, \mathcal{M})$ is a measurable space, a ***simple function*** on X is a finite linear combination of characteristic functions of measurable sets with complex coefficients. Every simple function can be written uniquely as the canonical form $\sum_{j=1}^{N} a_j \chi_{E_j}$ where $N < \infty$, the a_j's are distinct complex numbers (one of which may be 0), and the E_j's are disjoint measurable sets whose union is X.

Now suppose $(X, \mathcal{M}, \mu)$ is a measurable space. If $\phi = \sum_{j=1}^{N} a_j \chi_{E_j}$ is a nonnegative simple function, its integral with respect to μ, $\int \phi d\mu$, is defined in the obvious way:

$$\int \phi d\mu = \sum a_j \mu(E_j),$$

with the understanding that if $a_j = 0$ and $\mu(E_j) = \infty$, then $a_j \mu(E_j) = 0$. Note that $\int \phi d\mu$ may be $+\infty$ if some of the sets E_j have infinite measures. To extend this notion of an integral to more general functions, one can approximate such functions by simple functions, cf. [SS05]. Given a measure space $(X, \mathcal{M}, \mu)$, we set

$$L^+(X) = \{f : X \to [0, \infty] : \ f \text{ is measurable}\},$$

and for $f \in L^+(X)$ we define the integral of f with respect to μ by

$$\int f d\mu = \sup \left\{ \int \phi d\mu : \phi \text{ is simple and } 0 \leqslant \phi \leqslant f \right\}.$$

Thus, $\int f d\mu$ is an element of $[0, \infty]$. We say that a measurable $f : X \to \mathbb{C}$ is integrable if $\int |f| d\mu < \infty$, and we denote the set of integrable functions by $L^1(X, \mu)$:

$$L^1(X, \mu) = \left\{ f : X \to \mathbb{C} : \ f \text{ is measurable and } \int |f| d\mu < \infty \right\}.$$

We now recall three basic convergence theorems that address the question of when "the integral of the limit is the limit of the integrals".

Theorem 0.6 (Monotone convergence theorem, cf. [Fol99, Theorem 2.14]). *Let $\{f_n\}$ be a sequence in L^+ such that $f_n(x) \leqslant f_{n+1}(x)$ for all n and x and let*

$$f(x) = \lim_{n\to\infty} f_n(x) = \sup_n f_n(x)$$

(which always exists since we allow the value ∞). Then,

$$\int f = \lim_{n\to\infty} \int f_n.$$

Lemma 0.7 (Fatou's lemma, cf. [Fol99, Theorem 2.18]). *Let $\{f_n\}$ be a sequence in L^+, then*

$$\int \liminf_{n\to\infty} f_n \leqslant \liminf_{n\to\infty} \int f_n.$$

In particular, if $f_n \to f$ a.e., then

$$\int f \leqslant \liminf_{n\to\infty} \int f_n.$$

If we impose a bound on the functions f_n that forbids the areas under their graphs from escaping to infinity, we obtain another positive result.

Theorem 0.8 (Dominated convergence theorem, cf. [Fol99, Theorem 2.24]). *Let $\{f_n\}$ be a sequence in L^1 such that $f_n \to f$ a.e., and there exists a nonnegative $g \in L^1$ such that $|f_n| \leqslant g$ a.e. for all n, then $f \in L^1$ and*

$$\int f = \lim_{n\to\infty} \int f_n.$$

The next theorem gives a criterion, less restrictive than those found in most advanced calculus books, for the validity of interchanging a limit or a derivative with an integral.

Theorem 0.9 ([Fol99, Theorem 2.27]). *Suppose that $f : X \times [a, b] \to \mathbb{C}$ $(-\infty < a < b < \infty)$ and that $f(\cdot, t) : X \to \mathbb{C}$ is integrable for each $t \in [a, b]$. Let $F(t) = \int_X f(x, t) d\mu(x)$.*

(i) *Suppose that there exists $g \in L^1(\mu)$ such that $|f(x, t)| \leqslant g(x)$ for all x, t. If $\lim_{t\to t_0} f(x, t) = f(x, t_0)$ for every x, then $\lim_{t\to t_0} F(t) = F(t_0)$; in particular, if $f(x, \cdot)$ is continuous for each x, then F is continuous.*

(ii) *Suppose that $\partial f/\partial t$ exists and there is a $g \in L^1(\mu)$ such that $|(\partial f/\partial t)(x, t)| \leqslant g(x)$ for all x, t. Then, F is differentiable, and $F'(t) = \int (\partial f/\partial t)(x, t) d\mu(x)$.*

The Fubini-Tonelli theorem is an essential tool in analysis. It is most commonly used to justify interchanging the order of integration in an iterated integral. Let us first recall some notations.

Suppose that $(X_j, \mathcal{M}_j, \mu_j)$ is a σ-finite measure space for $j = 1, \cdots, n$, and let $X = \prod_1^n X_j$ and $\mathcal{M} = \bigotimes_1^n \mathcal{M}_j$, there is a unique measure π on $(X, \mathcal{M})$ such that

$$\pi(E_1 \times E_2 \times \cdots \times E_n) = \mu_1(E_1)\mu_2(E_2)\cdots\mu_n(E_n) \quad \text{for all } E_j \in \mathcal{M}_j,$$

with the understanding that any numerical product containing 0 as a factor has the value 0, even if one or more of the other factors is ∞. This measure is called the ***product*** of $\mu_1, \cdots, \mu_n$ and is denoted by $\mu_1 \times \cdots \times \mu_n$. In what follows we restrict the discussion to the case of two factors to keep the notation more manageable, but the generalization to n factors is straightforward.

Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. If f is a function on $X \times Y$, we can consider not only the integral of f with respect to the product measure but also the iterated integrals of f with respect to μ and ν or with respect to ν and μ. It will be convenient to employ the following notation for the functions on X and Y obtained from f by fixing one of its arguments:

$$f^y(x) = f(x, y) = f_x(y).$$

Here is the main result. Parts ii) and iii) are due to Tonelli and Fubini, respectively, in the case where $X = Y = \mathbb{R}$ and $\mu = \nu =$ Lebesgue measure. Fubini came first, and the whole theorem is often simply called ***Fubini's theorem***.

Theorem 0.10 (Fubini-Tonelli theorem, [Fol99, Theorem 2.37]). *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces.*

(i) *If f is an $\mathcal{M} \otimes \mathcal{N}$-measurable function on $X \times Y$, then f^y is $\mathcal{M}$-measurable for all $y \in Y$ and f_x is $\mathcal{N}$-measurable for all $x \in X$.*

(ii) *(Tonelli) If $f \in L^+(X \times Y)$, the functions $g(x) = \int f_x d\nu$ and $h(y) = \int f^y d\mu$ are in $L^+(X)$ and $L^+(Y)$, respectively, and*

$$\begin{aligned} \int_{X\times Y} f d(\mu \times \nu) &= \int_X \left[\int_Y f(x,y) d\nu(y) \right] d\mu(x) \\ &= \int_Y \left[\int_X f(x,y) d\mu(x) \right] d\nu(y). \end{aligned} \tag{0.8}$$

(iii) *(Fubini) If $f \in L^1(X \times Y)$, then $f_x \in L^1(\nu)$ for a.e. $x \in X$ and $f^y \in L^1(\mu)$ for a.e. $y \in Y$; the a.e.-defined functions $g(x) = \int f_x d\nu$ and $h(y) = \int f^y d\mu$ are in $L^1(\mu)$ and $L^1(\nu)$, respectively, and* (0.8) *holds.*

§ 0.3 L^p spaces

L^p spaces are a class of Banach spaces (for $p \in [1, \infty]$) of functions whose norms are defined in terms of integrals and that generalize the L^1 spaces and play a central role in modern analysis.

Let $(X, \mathcal{M}, \mu)$ be a measure space. If f is a measurable function on X and $0 < p < \infty$, we define

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

(allowing the possibility that $\|f\|_p = \infty$), and

$$L^p(X, \mathcal{M}, \mu) = \left\{ f : X \to \mathbb{C} : f \text{ is measurable and } \|f\|_p < \infty \right\}.$$

We abbreviate $L^p(X, \mathcal{M}, \mu)$ by $L^p(\mu)$, $L^p(X)$, or simply L^p when this will cause no confusion. We consider two functions to define the same element of L^p when they are equal almost everywhere.

For $p = \infty$, L^∞ consists of all μ-measurable and bounded functions. Then, we write

$$\|f\|_\infty = \operatorname*{ess\,sup}_X |f(x)| = \inf\{a > 0 : \mu(\{x : |f(x)| > a\}) = 0\},$$

with the convention that $\inf \varnothing = \infty$.

For $p \in [1, \infty]$, let $p' = p/(p-1)$ be the conjugate exponent of p, i.e., $1/p + 1/p' = 1$ (with the notations $1' = \infty$ and $\infty' = 1$). Then with this notation, we summarize some results about L^p.

Theorem 0.11. (i) *(Hölder's inequality, [Fol99, Theorem 6.2]) Suppose $p \in [1, \infty]$. If f and g are measurable functions on X, then*

$$\|fg\|_1 \leqslant \|f\|_p \|f\|_{p'}. \tag{0.9}$$

For $p = p' = 2$, this is the Cauchy-Schwartz inequality. In particular, if $f \in L^p$ and $g \in L^{p'}$, then $fg \in L^1$, and in this case equality holds in (0.9) iff $\alpha|f|^p = \beta|g|^{p'}$ a.e. for $p > 1$ and some constants α, β with $(\alpha, \beta) \neq (0, 0)$, or $|g(x)| = \|g\|_\infty$ a.e. for $p = 1$ on the set where $f(x) \neq 0$.

(ii) *(Minkowski's inequality, [Fol99, Theorem 6.5]) Suppose $p \in [1, \infty]$ and $f, g \in L^p$, then*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p.$$

(iii) *(Completeness: Riesz-Fisher theorem, [Fol99, Theorems 6.6, 6.8]) For $p \in [1, \infty]$, L^p is a Banach space.*

(iv) *(Riesz representation theorem, [DiB16, Theorem 11.1]) Suppose $1 < p < \infty$ and $E \in \mathcal{M}$. For every bounded linear functional F in $L^p(E)$,* *there exists a unique function $g \in L^{p'}(E)$ such that F is represented by*

$$F(f) = \int_E fg d\mu \quad \text{for all } f \in L^p(E).$$

Moreover, $\|F\| = \|g\|_{p'}$, hence $L^{p'}$ is isometrically isomorphic to $(L^p)'$ which denotes the dual space of L^p. The same conclusion holds for $p = 1$ provided $(X, \mathcal{M}, \mu)$ is σ-finite.

(v) *([Fol99, Proposition 6.13]) Suppose $p \in [1, \infty)$. If $f \in L^p$, then*

$$\|f\|_p = \sup \left\{ \left| \int fg \right| : \|g\|_{p'} = 1 \right\}.$$

If μ is semifinite, this result holds also for $p = \infty$.

(vi) *(Density, [Fol99, Proposition 6.7, Theorem 6.8]) For $p \in [1, \infty)$, the set of (finitely) simple functions $f = \sum_1^N a_j \chi_{E_j}$, where $\mu(E_j) < \infty$ for all j, is dense in L^p. In addition, the simple functions (not necessarily with finite measure support) are dense in L^∞.*

Proposition 0.12 (cf. [Fol99, Exercise 6.9]). *Suppose that $(X, \mathcal{M}, \mu)$ is a measure space and $p \in [1, \infty)$. If a sequence $\{f_k\} \subset L^p$ converges in L^p to f, then there is a subsequence $\{f_{k_j}\}$ that converges to f μ-a.e.*

Theorem 0.13 (Vitali convergence theorem, cf. [Fol99, Exercise 6.15]). *Suppose $1 \leqslant p < \infty$ and $\{f_n\}_1^\infty \subset L^p$. $\{f_n\}$ is Cauchy in the L^p norm iff the following three conditions hold:*

(i) *$\{f_n\}$ is Cauchy in measure;*

(ii) *the sequence $\{|f_n|^p\}$ is uniformly integrable;*

(iii) *for every $\varepsilon > 0$ there exists $E \subset X$ such that $\mu(E) < \infty$ and $\int_{E^c} |f_n|^p < \varepsilon$ for all n.*

The next result is a rather general theorem about the boundedness of integral operators on L^p spaces.

Theorem 0.14 (cf. [Fol99, Theorem 6.18]). *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces, and let K be an $(\mathcal{M} \otimes \mathcal{N})$-measurable function on $X \times Y$. Suppose that there exists $C > 0$ such that $\int |K(x, y)| d\mu(x) \leqslant C$ for a.e. $y \in Y$ and $\int |K(x, y)| d\nu(y) \leqslant C$ for a.e. $x \in X$ and that $1 \leqslant p \leqslant \infty$. If $f \in L^p(Y)$, then the integral*

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

converges absolutely for a.e. $x \in X$, the function Tf thus defined is in $L^p(X)$, *and* $\|Tf\|_p \leqslant C\|f\|_p$.

Proof. Suppose that $1 < p < \infty$. Let p' be the conjugate exponent to p. By applying Hölder's inequality to the product

$$|K(x,y)f(y)| = |K(x,y)|^{1/p'}(|K(x,y)|^{1/p}|f(y)|),$$

we have

$$\begin{aligned}\int |K(x,y)f(y)|d\nu(y) \leqslant &\left[\int |K(x,y)|d\nu(y)\right]^{1/p'}\left[\int |K(x,y)||f(y)|^p d\nu(y)\right]^{1/p}\\ \leqslant &C^{1/p'}\left[\int |K(x,y)||f(y)|^p d\nu(y)\right]^{1/p}\end{aligned}$$

for a.e. $x \in X$. Hence, by Tonelli's theorem,

$$\begin{aligned}\int\left[\int |K(x,y)f(y)|d\nu(y)\right]^p d\mu(x) \leqslant &C^{p/p'}\iint |K(x,y)||f(y)|^p d\nu(y)d\mu(x)\\ \leqslant &C^{1+p/p'}\int |f(y)|^p d\nu(y).\end{aligned}$$

Since the last integral is finite, Fubini's theorem implies that $K(x,\cdot)f \in L^1(Y)$ for a.e. x, so that Tf is well defined a.e., and

$$\int |Tf(x)|^p d\mu(x) \leqslant CC^{1+p/p'}\|f\|_p^p.$$

Taking pth roots, we are done.

For $p = 1$ the proof is similar but easier and requires only the hypothesis

$$\int |K(x,y)|d\mu(x) \leqslant C;$$

for $p = \infty$ the proof is trivial and requires only the hypothesis

$$\int |K(x,y)|d\nu(y) \leqslant C.$$

Details are left to the reader (Exercise 0.1). ❑

Minkowski's inequality states that the L^p norm of a sum is at most the sum of the L^p norms. There is a generalization of this result in which sums are replaced by integrals:

Theorem 0.15 (Minkowski's integral inequality, [Fol99, Exercise 6.19]). *Suppose that* $(X,\mathcal{M},\mu)$ *and* $(Y,\mathcal{N},\nu)$ *are* σ*-finite measure spaces, and let* f *be an* $(\mathcal{M}\otimes\mathcal{N})$*-measurable function on* $X \times Y$.

(i) *If* $f \geqslant 0$ *and* $p \in [1,\infty)$, *then*

$$\left[\int\left(\int f(x,y)d\nu(y)\right)^p d\mu(x)\right]^{1/p} \leqslant \int\left[\int f^p(x,y)d\mu(x)\right]^{1/p} d\nu(y).$$

(ii) *If* $p \in [1,\infty]$, $f(\cdot,y) \in L^p(\mu)$ *for a.e.* y, *and the function* $y \mapsto$

$\|f(\cdot,y)\|_p$ is in $L^1(\nu)$, then $f(x,\cdot)\in L^1(\nu)$ for a.e. x, the function $x\mapsto \int f(x,y)d\nu(y)$ is in $L^p(\mu)$, and

$$\left\|\int f(\cdot,y)d\nu(y)\right\|_p \leqslant \int \|f(\cdot,y)\|_p d\nu(y).$$

Proof. If $p=1$, (i) is merely Tonelli's theorem. If $1<p<\infty$, let p' be the conjugate exponent to p and suppose $g\in L^{p'}(X)$. Then by Tonelli's theorem and Hölder's inequality,

$$\begin{aligned}\int\left[\int f(x,y)d\nu(y)\right]|g(x)|d\mu(x)=&\iint f(x,y)|g(x)|d\mu(x)d\nu(y)\\ \leqslant&\|g\|_{p'}\int\left[\int f^p(x,y)d\mu(x)\right]d\nu(y).\end{aligned}$$

Therefore, (i) follows from (v) in Theorem 0.11. For $p<\infty$, (ii) follows from (i) (with f replaced by $|f|$) and Fubini's theorem; for $p=\infty$, it is a simple consequence of the monotonicity of the integral. ❑

As an application, the next result is a theorem concerning integral operators on $(0,\infty)$ with the Lebesgue measure.

Theorem 0.16 ([Fol99, Theorem 6.20]). *Let K be a Lebesgue measurable function on $(0,\infty)\times(0,\infty)$ such that $K(\lambda x,\lambda y)=\lambda^{-1}K(x,y)$ for all $\lambda>0$ and $\int_0^\infty |K(x,1)|x^{-1/p}dx=C<\infty$ for some $p\in[1,\infty]$, and let p' be the conjugate exponent to p. For $f\in L^p$ and $g\in L^{p'}$, let*

$$Tf(y)=\int_0^\infty K(x,y)f(x)dx,\quad Sg(x)=\int_0^\infty K(x,y)g(y)dy.$$

Then Tf and Sg are defined a.e., and $\|Tf\|_p\leqslant C\|f\|_p$ and $\|Sg\|_{p'}\leqslant C\|g\|_{p'}$.

Proof. Setting $z=x/y$, we have

$$\int_0^\infty |K(x,y)f(x)|dx=\int_0^\infty |K(yz,y)f(yz)|ydz=\int_0^\infty |K(z,1)f_z(y)|dz$$

where $f_z(y)=f(yz)$; moreover,

$$\|f_z\|_p=\left[\int_0^\infty |f(yz)|^p dy\right]^{1/p}=\left[\int_0^\infty |f(x)|^p z^{-1}dx\right]^{1/p}=z^{-1/p}\|f\|_p.$$

Therefore, by Minkowski's inequality for integrals, Tf exists a.e. and

$$\|Tf\|_p\leqslant\int_0^\infty |K(z,1)|\|f_z\|_p dz=\|f\|_p\int_0^\infty |K(z,1)|z^{-1/p}dz=C\|f\|_p.$$

Finally, setting $u=y^{-1}$, we have

$$\int_0^\infty |K(1,y)|y^{-1/p'}dy=\int_0^\infty |K(y^{-1},1)|y^{-1-1/p'}dy$$

$$= \int_0^\infty |K(u,1)| u^{-1/p} du = C,$$

so the same reasoning shows that Sg is defined a.e. and that $\|Sg\|_{p'} \leqslant C\|g\|_{p'}$. ❑

Corollary 0.17 ([Fol99, Corollary 6.21]). *Let*

$$Tf(y) = y^{-1}\int_0^y f(x)dx, \quad Sg(x) = \int_x^\infty y^{-1}g(y)dy.$$

Then for $1 < p \leqslant \infty$ *and* $1 \leqslant q < \infty$,

$$\|Tf\|_p \leqslant \frac{p}{p-1}\|f\|_p, \quad \|Sg\|_q \leqslant q\|g\|_q.$$

Proof. Let $K(x,y) = y^{-1}\chi_E(x,y)$ where $E = \{(x,y) : x < y\}$. Then

$$\int_0^\infty |K(x,1)| x^{-1/p} dx = \int_0^1 x^{-1/p} dx = p/(p-1),$$

and

$$\int_0^\infty |K(x,1)| x^{-1/q'} dx = \int_0^1 x^{-1/q'} dx = q'/(q'-1) = q,$$

where q' is the conjugate exponent to q, so Theorem 0.16 yields the result. ❑

Corollary 0.17 is a special case of Hardy's inequalities; the general result is in Exercise 0.4.

Definition 0.18. We define $L^{p_1}(X) + L^{p_2}(X)$ to be the space of all functions f, such that $f = f_1 + f_2$, with $f_1 \in L^{p_1}(X)$ and $f_2 \in L^{p_2}(X)$.

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

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

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

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

since $p_1 - p \leqslant 0$. Similarly, due to $p_2 \geqslant p$,

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

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

Proposition 0.19 (Interpolation of L^p spaces)**.** *If $0 < p < q < r \leqslant \infty$, then $L^p \cap L^r \subset L^q \subset L^p + L^r$ and $\|f\|_q \leqslant \|f\|_p^{1-\theta}\|f\|_r^{\theta}$, where $\theta \in (0,1)$ is defined by*

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

The three most obviously important L^p spaces are L^1, L^2 and L^∞. With L^1 we are very familiar, L^2 is special because it is a Hilbert space, and the topology on L^∞ is closely related to the topology of uniform convergence. Unfortunately, L^1 and L^∞ are pathological in many respects, and it is more fruitful to deal with the intermediate L^p spaces, e.g., the duality theory. Many operators of interest in Fourier analysis and differential equations are bounded on L^p for $1 < p < \infty$ but not on L^1 or L^∞.

§ 0.4 Weak* topology

Let X be a Banach space. The weak topology on X is the weakest topology such that every bounded linear functional on X is continuous. The dual space of X is the space X' of all bounded linear functionals on X. The norm of $x^* \in X'$ is defined by

$$\|x^*\| = \sup\{|x^*(x)| : \|x\| \leqslant 1\}.$$

The weak* topology on X' is the weakest topology such that for all $x \in X$ the functional $x^* \mapsto x^*(x)$ is continuous. The weak* closure of a set $Z \subseteq X'$ is denoted $\mathrm{cl}^* Z$. Note that about any point $x_0^* \in X'$ there is a weak* neighborhood basis consisting of all sets of the form

$$\{x^* \in X' : |x^*(x_1) - x_0^*(x_1)| < 1, \ldots, |x^*(x_n) - x_0^*(x_n)| < 1\}$$

for some finite set $x_1, \ldots, x_n \in X$.

A key theorem (cf. [DS88, p.424]) concerning the weak* topology is:

Theorem 0.20 (Banach-Alaoglu theorem)**.** *Let X' be the dual to some Banach space X. Then, for any $r > 0$, the closed ball*

$$B_r(X') = \{x^* \in X' : \|x^*\| \leqslant r\}$$

is weak closed and weak* compact. Furthermore, if X is separable then $B_r(X')$ is weak* metrizable.*

Exercises

Exercise 0.1. [Fol99, Exercise 6.26] Complete the proof of Theorem 0.14 for the cases $p = 1$ and $p = \infty$.

Exercise 0.2 (Young's conjugate functions [Pey18, Exercise 2.7])**.** Let φ be a continuous increasing function from $[0, \infty)$ onto $[0, \infty)$. One sets

$$\Phi(x) = \int_0^x \varphi(t)dt \quad \text{and} \quad \Psi(y) = \int_0^y \varphi^{-1}(t)dt.$$

(i) Prove that, for all $a \geqslant 0$ and $b \geqslant 0$, $ab \leqslant \Phi(a) + \Psi(b)$ and that the equality holds iff $b = \varphi(a)$.

(ii) Prove the inequality, valid for nonnegative a and b,

$$ab \leqslant (a+1)\ln(a+1) + e^b.$$

Hint (i) Draw a picture. (ii) Use the function $\varphi(t) = \ln(1+t)$.

Exercise 0.3. [Pey18, Exercise 2.8] Let Φ and Ψ be as in Exercise 0.2, and let f and g be two measurable functions.

(i) Prove the inequality

$$\int |f(x)g(x)|dx \leqslant \int \Phi(|f(x)|)dx + \int \Psi(|g(x)|)dx.$$

(ii) Define

$$\|f\|_\Phi = \inf\left\{t > 0 : \int \Phi(|f(x)|/t)dx \leqslant 1\right\}.$$

Then

$$\int |f(x)g(x)|dx \leqslant \|f\|_\Phi \left(1 + \int \Psi(|g(x)|)dx\right)$$

and

$$\int |f(x)g(x)|dx \leqslant 2\|f\|_\Phi \|g\|_\Psi.$$

(iii) Prove that the set of f such that $\|f\|_\Phi < \infty$ is a normed vector space. Such a space is called an Orlicz space.

(iv) Let f_j be a sequence of functions. Then $\lim\limits_{j\to\infty} \int \Phi(t|f_j(x)|)dx = 0$ for all $t > 0$, iff $\lim\limits_{j\to\infty} \|f_j\|_\Phi = 0$.

Exercise 0.4 (Hardy's inequalities [Fol99, Exercise 6.29])**.** Suppose that $1 \leqslant p < \infty$, $r > 0$, and h is a nonnegative measurable function on $(0, \infty)$. Prove:

(i) $$\int_0^\infty \left[\int_0^t h(y)dy\right]^p t^{-r-1}dt \leqslant \left(\frac{p}{r}\right)^p \int_0^\infty (yh(y))^p y^{-r-1}dy,$$

(ii) $$\int_0^\infty \left[\int_t^\infty h(y)dy\right]^p t^{r-1}dt \leqslant \left(\frac{p}{r}\right)^p \int_0^\infty (yh(y))^p y^{r-1}dy.$$

Hint (i) Apply Theorem 0.16 with $K(t,y) = t^{\beta-1}y^{-\beta}\chi_{(0,\infty)}(y-t)$ and

$f(t) = t^\gamma h(t)$ for suitable β, γ; or use Jensen's inequality.

(ii) Let $h(x) = f(1/x)/x^2$ and use (i).

1

Interpolation of Operators

In this chapter, we introduce several interpolation theorems. There are generally two ways to derive them: "complex interpolation", such as the Riesz-Thorin interpolation theorem, and "real interpolation", such as the Marcinkiewicz interpolation theorem. The former gives sharper results and is more elegant because it is based on analytic-function theory in the complex plane. However, this also carries some disadvantages: The L^p-spaces need to be defined over $\mathbb{C}$, and the method only works for linear operators. The "real interpolation" method is less elegant in terms of results and proofs, but it works in larger generality and includes even nonlinear operators such as maximal functions (given in §2.1).

These two interpolation theorems, Riesz-Thorin and Marcinkiewicz, have been developed into full theories of interpolation of operators between function spaces in advanced functional analysis, with the goal of studying which families of function spaces and operators between them can be interpolated.

An important observation of E. Stein is that the proof of the Riesz-Thorin interpolation theorem can be generalized to the case where the operator T itself varies analytically, so we will introduce the Stein interpolation theorem. We also introduce the notion of weak-L^p spaces that are natural objects for interpolation theory and used to prove the Marcinkiewicz interpolation theorem and weak-type estimates (cf. Definition 1.25).

§ 1.1 Riesz-Thorin interpolation theorem

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 $L^q = L^q(Y, d\nu)$. 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.,

$$A = \sup_{\{f:\|f\|_p \neq 0\}} \frac{\|Tf\|_q}{\|f\|_p} = \sup_{\|f\|_p=1} \|Tf\|_q < \infty.$$

The number A is called the norm of mapping T.

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

If $1 \leqslant p < q < r \leqslant \infty$, then $(L^p \cap L^r) \subset L^q \subset (L^p + L^r)$, and it is natural to ask whether a linear operator T on $L^p + L^r$ that is bounded on both L^p and L^r is also bounded on L^q. The answer is affirmative, and this result can be generalized in various ways. A fundamental theorem on this question is the Riesz-Thorin interpolation theorem of "complex interpolation" as follows:

Theorem 1.1 (Riesz-Thorin interpolation theorem). *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be a pair of measure spaces and $p_0, p_1, q_0, q_1 \in [1, \infty]$. If $q_0 = q_1 = \infty$, suppose also that the measure ν on Y is semifinite. For $0 < \theta < 1$, define p and q by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{1.1}$$

If T is a linear operator from $(L^{p_0} + L^{p_1})(X, d\mu)$ into $(L^{q_0} + L^{q_1})(Y, d\nu)$ such that $\|Tf\|_{q_0} \leqslant A_0 \|f\|_{p_0}$, for $f \in L^{p_0}(X, d\mu)$ and $\|Tf\|_{q_1} \leqslant A_1 \|f\|_{p_1}$, for $f \in L^{p_1}(X, d\mu)$, then

$$\|Tf\|_q \leqslant A_\theta \|f\|_p, \quad \text{for } f \in L^p(X, d\mu),\ 0 < \theta < 1,$$

with

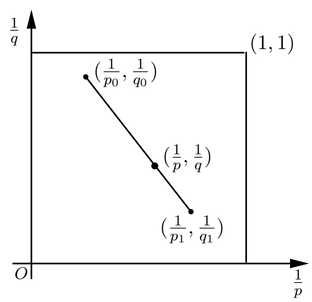
$$A_\theta \leqslant A_0^{1-\theta} A_1^{\theta}. \tag{1.2}$$

Remark 1.2. 1) (1.2) means that A_θ is logarithmically convex, i.e., $\ln A_\theta$ is convex. Because of this, the above theorem is sometimes known as the Riesz(-Thorin) convexity theorem.

2) The geometrical meaning of (1.1) is that the points $(1/p, 1/q)$ are the points on the line segment between $(1/p_0, 1/q_0)$ and $(1/p_1, 1/q_1)$; see the figure.

3) One can only assume the boundedness of T for all finitely simple functions f on X and obtain the boundedness for all finitely simple functions. When $p < \infty$, by density, T has a unique bounded extension from $L^p(X, \mu)$ to $L^q(Y, \nu)$ when p and q are as in (1.1).

The Riesz-Thorin interpolation theorem is due to M. Riesz, who, in 1926, in the process of proving the convergence of Fourier series in the L^p norm, developed an interpolation method in the framework of bilinear forms in L^p with a long and difficult calculation subject to conditions $p_0 \leqslant q_0$, $p_1 \leqslant q_1$ and L^p spaces over $\mathbb{R}$. Later, one of his students, O. Thorin, removed the restriction on



the indices by introducing complex analysis methods in 1939. This is what is now called the Riesz-Thorin interpolation theorem, whose current widely presented proof is actually due to a further simplification by A. Zygmund. The proof of Theorem 1.1 is conceptually fairly simple but wonderful in a number of specific aspects. The key idea is a result of complex analysis proved by J. Hadamard in 1896.

Lemma 1.3 (Hadamard three lines lemma). *Let* $S = \{z \in \mathbb{C} : 0 \leqslant \operatorname{Re} z \leqslant 1\}$ *and* $f : S \to \mathbb{C}$ *be bounded and continuous on* S *and analytic on the interior* $\mathring{S}$ *of* S. *Let* $A_\theta = \sup_{t\in\mathbb{R}} |f(\theta + it)|$. *Then we have* $A_\theta \leqslant A_0^{1-\theta} A_1^{\theta}$ *for all* $\theta \in [0, 1]$.

Proof. Without loss of generality, we can assume that $A_0 = A_1 = 1$. Otherwise, we replace f by the function $g : S \to \mathbb{C}$ defined by $g(z) = f(z)/(A_0^{1-z}A_1^z)$. By the assumption on f, it follows that g is continuous and bounded on S (because $|g(z)| \leqslant |f(z)|/(A_0^{1-\operatorname{Re} z}A_1^{\operatorname{Re} z}) \leqslant |f(z)|/(\min(1, A_0, A_1))$ and A_0, A_1 cannot be zero) and analytic on $\mathring{S}$, with $\sup_{\operatorname{Re} z=0} |g(z)| = \sup_{\operatorname{Re} z=1} |g(z)| = 1$. Hence, we can assume that

$$\sup_{\operatorname{Re} z=0,1} |f(z)| = 1$$

and, under this assumption, we need to show that

$$\sup_{z\in S} |f(z)| \leqslant 1.$$

To this end, we define the sequence $f_n(z) = f(z)e^{(z^2-1)/n}$ and we observe that $|f_n(z)| \leqslant |f(z)|$ for all $z \in S$, in particular, $\sup_{\operatorname{Re} z=0,1} |f_n(z)| \leqslant 1$. Moreover, $f_n(z)$ is analytic in $\mathring{S}$ for all $n \geqslant 1$, and $|f_n(x + iy)| \to 0$, as $|y| \to \infty$, for every fixed n, uniformly in x. Hence, we obtain for every $n \geqslant 1$,

$$\sup_{z\in S} |f_n(z)| \leqslant 1,$$

because analytic functions attain their maximum and minimum on the boundary of any compact set (cf. [SS03, p.92], consider the compact domain $K = \{z : |\operatorname{Im} z| \leqslant \kappa, 0 \leqslant \operatorname{Re} z \leqslant 1\}$, where κ is so large that $|f_n(x + iy)| \leqslant 1$ for all $|y| \geqslant \kappa$, and $x \in [0, 1]$). Since $|f_n(z)| \to |f(z)|$ as $n \to \infty$, it follows that $|f(z)| \leqslant 1$ for all $z \in S$. ❑

The proof of the Riesz-Thorin interpolation theorem then follows by building a function, using duality, that depends holomorphically on z corresponding to a complex parameter such that $\operatorname{Re}(z) = 1/p$ and then using the Hadamard three lines lemma to obtain the intermediate bounds. Duality again yields the final estimates.

Proof of Theorem 1.1. We observe that the case $p_0 = p_1$ follows from Proposition 0.19: If $p = p_0 = p_1$, then

$$\|Tf\|_q \leqslant \|Tf\|_{q_0}^{1-\theta}\|Tf\|_{q_1}^{\theta} \leqslant A_0^{1-\theta}A_1^{\theta}\|f\|_p.$$

Thus, we may assume that $p_0 \neq p_1$ and, in particular, that $p < \infty$ for $0 < \theta < 1$.

Denote

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

and $1/q' = 1 - 1/q$. Then, we have, by the dual (i.e., (v) in Theorem 0.11, where we need that ν is semifinite if $q = \infty$, i.e., $q_0 = q_1 = \infty$),

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

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, $q' < \infty$. (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 $q' = \infty$.) 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 does g.

For $0 \leqslant \operatorname{Re} z \leqslant 1$, we set

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

and

$$\eta(z) = \eta(x,z) = \begin{cases} |f(x)|^{\frac{p}{p(z)}}\frac{f(x)}{|f(x)|}, & x \in \{x \in X : f(x) \neq 0\}; \\ 0, & \text{otherwise,} \end{cases}$$

$$\zeta(z) = \zeta(y,z) = \begin{cases} |g(y)|^{\frac{q'}{q'(z)}}\frac{g(y)}{|g(y)|}, & y \in \{y \in Y : g(y) \neq 0\}; \\ \zeta(z) = 0, & \text{otherwise.} \end{cases}$$

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

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

Thus,

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

We have

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

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

On one hand, we have $\lim_{|f(x)|\to 0_+}|f(x)|^\alpha \ln|f(x)| = 0$ for any $\alpha > 0$, i.e., $\forall \varepsilon > 0$, $\exists\delta > 0$ s.t. $||f(x)|^\alpha \ln|f(x)|| < \varepsilon$ if $|f(x)| < \delta$. On the other hand, if $|f(x)| \geqslant \delta$, 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,

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

which yields

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

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

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

Since

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

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

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's three lines lemma, reproduced as Lemma 1.3, we obtain 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 a rather simple application of the Riesz-Thorin interpolation theorem.

Theorem 1.4 (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. Fix $f \in L^p$, $p \in [1, \infty]$, then we 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 $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$. Thus, we obtain the stated inequality for precisely the exponents p, q and r in the hypothesis. ❑

Remark 1.5. 1) The sharp form of Young's inequality for convolutions can be found in [Bec75b, Theorem 3] or [Bec75a; BL76b], and we state it as follows. Under the assumption of Theorem 1.4, 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 Hölder conjugate numbers, i.e., $1/m + 1/m' = 1$.

2) The Riesz-Thorin interpolation theorem is valid for a sublinear operator. One can see [CZ56] for details.

§1.2 Stein interpolation theorem

An important observation of E. Stein is that the above proof of the Riesz-Thorin interpolation theorem can be generalized to the case where the operator T itself varies analytically. For the motivation and importance, one can read Terence Tao's blog (`https://wp.me/p3qzP-1g5`). 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

$$\int_Y |T_z(\chi_A)\chi_B| d\nu < \infty \tag{1.3}$$

whenever A and B are subsets of finite measure of X and Y, respectively. The family $\{T_z\}_z$ is said to be ***analytic*** if the function

$$z \to \int_Y T_z(f) g d\nu \tag{1.4}$$

is analytic in the open strip $\mathring{S}$ 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

$$e^{-a|\operatorname{Im} z|} \ln \left| \int_Y T_z(f) g d\nu \right| \leqslant C_{f,g} < \infty \tag{1.5}$$

for all $z \in S$.

Note that if there is $a \in (0, \pi)$ such that for all measurable subsets A of X and B of Y of finite measure there is a constant $c(A, B)$ such that

$$e^{-a|\operatorname{Im} z|} \ln \left| \int_B T_z(\chi_A) d\nu \right| \leqslant c(A, B), \tag{1.6}$$

then (1.5) holds for $f = \sum\limits_{k=1}^{M} a_k \chi_{A_k}$ and $g = \sum\limits_{j=1}^{N} b_j \chi_{B_j}$ with

$$C_{f,g} = \ln(MN) + \sum_{k=1}^{M} \sum_{j=1}^{N} \left(c(A_k, B_j) + \big|\ln |a_k b_j|\big| \right).$$

In fact, by the linearity of T_z, the increase in ln and (1.6), we obtain

$$\begin{aligned}
\ln \left| \int_Y T_z(f) g d\nu \right| &= \ln \left| \int_Y T_z \left(\sum_{k=1}^{M} a_k \chi_{A_k} \right) \sum_{j=1}^{N} b_j \chi_{B_j} d\nu \right| \\
&= \ln \left| \sum_{k=1}^{M} \sum_{j=1}^{N} a_k b_j \int_{B_j} T_z(\chi_{A_k}) d\nu \right| \\
&\leqslant \ln \sum_{k=1}^{M} \sum_{j=1}^{N} |a_k b_j| \left| \int_{B_j} T_z(\chi_{A_k}) d\nu \right| \\
&\leqslant \ln \left[MN \max_{k,j} \left(|a_k b_j| \exp\left(c(A_k, B_j) e^{a|\operatorname{Im} z|} \right) \right) \right] \\
&\leqslant \ln(MN) + \max_{k,j} \left| \ln \left[\left(|a_k b_j| \exp\left(c(A_k, B_j) e^{a|\operatorname{Im} z|} \right) \right) \right] \right| \\
&\leqslant \ln(MN) + \max_{k,j} \left[\big|\ln |a_k b_j|\big| + c(A_k, B_j) e^{a|\operatorname{Im} z|} \right]
\end{aligned}$$

$$\leqslant \ln(MN) + \sum_{k=1}^{M}\sum_{j=1}^{N}\left[\left|\ln|a_k b_j|\right| + c(A_k, B_j)e^{a|\operatorname{Im} z|}\right]$$
$$\leqslant \left[\ln(MN) + \sum_{k=1}^{M}\sum_{j=1}^{N}\left(\left|\ln|a_k b_j|\right| + c(A_k, B_j)\right)\right]e^{a|\operatorname{Im} z|}.$$

Then, we have an extension of the three lines theorem due to I. Hirschman [Hir53].

Lemma 1.6 (Hirschman lemma). *Let F be analytic on the open strip $\mathring{S} = \{z \in \mathbb{C} : 0 < \operatorname{Re} z < 1\}$ and continuous on its closure S such that for some $A < \infty$ and $0 \leqslant a < \pi$, we have*

$$\ln|F(z)| \leqslant Ae^{a|\operatorname{Im} z|} \tag{1.7}$$

for all $z \in S$. Then

$$|F(x+iy)| \leqslant \exp\left\{\frac{\sin \pi x}{2}\int_{-\infty}^{\infty}\left[\frac{\ln|F(it+iy)|}{\cosh \pi t - \cos \pi x} + \frac{\ln|F(1+it+iy)|}{\cosh \pi t + \cos \pi x}\right]dt\right\},$$

whenever $0 < x < 1$, and y is real.

Before we give the proof of Lemma 1.6, we first recall the Poisson-Jensen formula[1] from [Rub96, p.21].

Theorem 1.7 (Poisson-Jensen formula). *Suppose that f is meromorphic[a] in the disk $D_R = \{z \in \mathbb{C} : |z| < R\}$, $r < R$. Then for any $z = re^{i\theta}$ in D_R, we have*

$$\ln|f(re^{i\theta})| = \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|f(Re^{i\varphi})|\frac{R^2 - r^2}{|Re^{i\varphi} - re^{i\theta}|^2}d\varphi + \sum_{|z_\nu|<R}\ln|B_R(z : z_\nu)|$$
$$- \sum_{|w_\nu|<R}\ln|B_R(z : w_\nu)| - k\ln\frac{R}{r},$$

where B is the Blaschke factor defined by

$$B_R(z : a) = \frac{R(z-a)}{R^2 - \bar{a}z}$$

and the z_ν are the zeros of f, the w_ν are the poles of f, and k is the order of the zero or the pole at the origin.

[a] In complex analysis, a meromorphic function on an open subset D of the complex plane is a function that is holomorphic on all D except a set of isolated points, which are poles for the function.

[1] This is a generalization of Jensen's Theorem (i.e., the case of $r = 0$, [SS03, p.135]) by using the Gauss mean value theorem.

Now, we consider the modulus of the Blaschke factor since

$$B_R(z:a)=\frac{R(z-a)}{R^2-\bar{a}z}=\frac{(z/R-a/R)}{1-\bar{a}z/R^2},$$

it suffices to consider $B_1(z:a)$ with $|z|<1$ and $|a|<1$, thus, we consider for $z=re^{i\theta}$

$$\left|\frac{z-a}{1-\bar{a}z}\right|=\left|\frac{re^{i\theta}-a}{1-\bar{a}re^{i\theta}}\right|=\left|\frac{r-ae^{-i\theta}}{1-\overline{ae^{-i\theta}}r}\right|.$$

Letting $a^*=ae^{-i\theta}$, this becomes

$$\left|\frac{r-a^*}{1-\overline{a^*}r}\right|.$$

Therefore, it is sufficient to consider the case in which $z=r$ is a real number. Note further that replacing a^* by $\overline{a^*}$ is equivalent to taking the complex conjugate of the entire fraction. Therefore, it is sufficient to consider

$$\left(\frac{r-a}{1-\bar{a}r}\right)\left(\frac{r-\bar{a}}{1-ar}\right)=\frac{r^2-2r\operatorname{Re}a+|a|^2}{1-2r\operatorname{Re}a+r^2|a|^2}\leqslant 1,$$

since $r^2+|a|^2-(1+r^2|a|^2)=(r^2-1)(1-|a|^2)<0$ for $r<1$ and $|a|<1$. Thus, we obtain $|B_1(z:a)|<1$, and then

$$\ln|B_R(z:a)|<0. \tag{1.8}$$

Proof of Lemma 1.6. It is not difficult to verify that

$$h(\zeta)=\frac{1}{\pi i}\ln\left(i\frac{1+\zeta}{1-\zeta}\right)$$

is a conformal map from $D=\{z:|z|<1\}$ onto the strip $\mathring{S}=(0,1)\times\mathbb{R}$. Indeed, $i(1+\zeta)/(1-\zeta)$ lies in the upper half-plane, and the preceding complex logarithm is a well-defined holomorphic function that takes the upper half-plane onto the strip $\mathbb{R}\times(0,\pi)$. Since $F\circ h$ is a holomorphic function in D, which implies that there are no poles in D, by the Poisson-Jensen formula with (1.8), we have

$$\ln|F(h(z))|\leqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|F(h(Re^{i\varphi}))|\frac{R^2-\rho^2}{|Re^{i\varphi}-\rho e^{i\theta}|^2}d\varphi \tag{1.9}$$

when $z=\rho e^{i\theta}$ and $|z|=\rho<R$. Let $\zeta=e^{i\varphi}$, $h(R\zeta)=\frac{1}{\pi i}\ln\left(i\frac{1+R\zeta}{1-R\zeta}\right)$, we observe that for $R<|\zeta|=1$ the hypothesis on F implies that

$$\ln|F(h(Re^{i\varphi}))|\leqslant Ae^{a\left|\operatorname{Im}\frac{1}{\pi i}\ln\left(i\frac{1+R\zeta}{1-R\zeta}\right)\right|}\quad\text{(by (1.7))}$$

(If $z=|z|e^{i\beta}$, then $\ln z=\ln|z|+i\beta$, thus $\operatorname{Im}i\ln z=\ln|z|$.)

$$=Ae^{\frac{a}{\pi}\left|\ln\left|\frac{1+R\zeta}{1-R\zeta}\right|\right|}$$

$$=Ae^{\frac{a}{\pi}\left|\ln\frac{|1+R\zeta|}{|1-R\zeta|}\right|}\quad\text{(due to }|z_1/z_2|=|z_1|/|z_2|)$$

$$=Ae^{\frac{a}{\pi}\left|\ln\sqrt{\frac{(1+R\cos\varphi)^2+(R\sin\varphi)^2}{(1-R\cos\varphi)^2+(R\sin\varphi)^2}}\right|}$$

(the square root is $\geqslant 1$ if $\cos\varphi \geqslant 0$ and < 1 otherwise)

$$=A\left(\frac{1+R^2+2R|\cos\varphi|}{1+R^2-2R|\cos\varphi|}\right)^{\frac{a}{2\pi}}.$$

Since

$$\begin{aligned}1+R^2-2R|\cos\varphi| &=(R-|\cos\varphi|)^2+\sin^2\varphi \geqslant \sin^2\varphi,\\ 1+R^2+2R|\cos\varphi| &\leqslant(1+R)^2\leqslant 4,\end{aligned}$$

we obtain

$$\ln|F(h(Re^{i\varphi}))| \leqslant A\left(\frac{4}{\sin^2\varphi}\right)^{\frac{a}{2\pi}} \leqslant A2^{\frac{a}{\pi}}|\sin\varphi|^{-\frac{a}{\pi}}.$$

Now,

$$\begin{aligned}\int_{-\pi}^{\pi}|\sin\varphi|^{-\frac{a}{\pi}}d\varphi &=4\int_0^{\frac{\pi}{2}}\sin^{-\frac{a}{\pi}}\varphi d\varphi\\ &=4\int_0^{\frac{\pi}{2}}\sin^{2(\frac{1}{2}-\frac{a}{2\pi})-1}\varphi\cos^{2\cdot\frac{1}{2}-1}\varphi d\varphi\\ &=2B\left(\frac{1}{2},\frac{1}{2}-\frac{a}{2\pi}\right)<\infty,\end{aligned}$$

since $a < \pi$ and the fact that the Beta function

$$B(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = 2\int_0^{\frac{\pi}{2}}\sin^{2\beta-1}\varphi\cos^{2\alpha-1}\varphi d\varphi$$

converges for $\operatorname{Re}\alpha, \operatorname{Re}\beta > 0$. Moreover, for $1 > R > \frac{1}{2}(\rho+1)$, it holds that

$$\begin{aligned}\frac{R^2-\rho^2}{|Re^{i\varphi}-\rho e^{i\theta}|^2} &=\frac{R^2-\rho^2}{R^2-2R\rho\cos(\theta-\varphi)+\rho^2}\leqslant\frac{R^2-\rho^2}{R^2-2R\rho+\rho^2}\\ &=\frac{(R-\rho)(R+\rho)}{(R-\rho)^2}=\frac{R+\rho}{R-\rho}\leqslant\frac{2}{\frac{1}{2}(\rho+1)-\rho}\leqslant\frac{4}{1-\rho}.\end{aligned}$$

Thus, (1.9) is uniformly bounded w.r.t. $R\in(\frac{1}{2}(\rho+1),1)$.

We will now use the following consequence of Fatou's lemma: suppose that $F_R \leqslant G$, where $G \geqslant 0$ is integrable, then

$$\limsup_R \int F_R d\varphi \leqslant \int \limsup_R F_R d\varphi.$$

Letting $R \nearrow 1$ in (1.9) and using this convergence result, we obtain

$$\ln|F(h(\rho e^{i\theta}))| \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|F(h(e^{i\varphi}))|\frac{1-\rho^2}{1-2\rho\cos(\theta-\varphi)+\rho^2}d\varphi. \quad (1.10)$$

Setting $x = h(\rho e^{i\theta})$, we obtain that

$$\rho e^{i\theta} =h^{-1}(x)=\frac{e^{\pi i x}-i}{e^{\pi i x}+i}=\frac{\cos\pi x+i\sin\pi x-i}{\cos\pi x+i\sin\pi x+i}$$

$$
\begin{aligned}
&= \frac{(\cos \pi x + i(\sin \pi x - 1))(\cos \pi x - i(\sin \pi x + 1))}{|\cos \pi x + i(\sin \pi x + 1)|^2} \\
&= -i \frac{\cos \pi x}{1 + \sin \pi x} = \left(\frac{\cos \pi x}{1 + \sin \pi x} \right) e^{-\frac{\pi}{2} i},
\end{aligned}
$$

from which it follows that $\rho = (\cos \pi x)/(1 + \sin \pi x)$ and $\theta = -\pi/2$ when $x \in (0, \frac{1}{2}]$, while $\rho = -(\cos \pi x)/(1 + \sin \pi x)$ and $\theta = \pi/2$ when $x \in [\frac{1}{2}, 1)$. In either case, we have $\rho = (\operatorname{sgn}(\frac{1}{2} - x))(\cos \pi x)/(1 + \sin \pi x)$ and $\theta = -(\operatorname{sgn}(\frac{1}{2} - x))\pi/2$ for $x \in (0, 1)$. We easily deduce that

$$
\begin{aligned}
&\frac{1 - \rho^2}{1 - 2\rho \cos(\theta - \varphi) + \rho^2} \\
=& \frac{1 - \frac{\cos^2 \pi x}{(1+\sin \pi x)^2}}{1 - 2(\operatorname{sgn}(\frac{1}{2} - x)) \frac{\cos \pi x}{1+\sin \pi x} \cos((\operatorname{sgn}(\frac{1}{2} - x))\frac{\pi}{2} + \varphi) + \frac{\cos^2 \pi x}{(1+\sin \pi x)^2}} \\
=& \frac{(1 + \sin \pi x)^2 - \cos^2 \pi x}{(1 + \sin \pi x)^2 + 2(1 + \sin \pi x) \cos \pi x \sin \varphi + \cos^2 \pi x} \\
=& \frac{2 \sin \pi x + 2 \sin^2 \pi x}{2(1 + \sin \pi x)(1 + \cos \pi x \sin \varphi)} \\
=& \frac{\sin \pi x}{1 + \cos \pi x \sin \varphi},
\end{aligned}
$$

since

$$
\begin{aligned}
\cos((\operatorname{sgn}(\frac{1}{2} - x))\frac{\pi}{2} + \varphi) &= \cos(\frac{\pi}{2} + (\operatorname{sgn}(\frac{1}{2} - x))\varphi) \\
&= -\sin((\operatorname{sgn}(\frac{1}{2} - x))\varphi) = -\operatorname{sgn}(\frac{1}{2} - x) \sin \varphi.
\end{aligned}
$$

Using this we write (1.10) as

$$
\ln |F(x)| \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sin \pi x}{1 + \cos \pi x \sin \varphi} \ln |F(h(e^{i\varphi}))| d\varphi. \tag{1.11}
$$

We now change the variables. On the interval $[-\pi, 0)$, we use the change of variables $it = h(e^{i\varphi})$ or, equivalently,

$$
\begin{aligned}
e^{i\varphi} =& h^{-1}(it) = \frac{e^{-\pi t} - i}{e^{-\pi t} + i} = \frac{(e^{-\pi t} - i)^2}{e^{-2\pi t} + 1} = \frac{e^{-2\pi t} - 1 - 2ie^{-\pi t}}{e^{-2\pi t} + 1} \\
=& \frac{e^{-\pi t} - e^{\pi t} - 2i}{e^{-\pi t} + e^{\pi t}} = -\tanh \pi t - i \operatorname{sech} \pi t.
\end{aligned}
$$

Observe that as φ ranges from $-\pi$ to 0, t ranges from $+\infty$ to $-\infty$. Furthermore, $d\varphi = -\pi \operatorname{sech} \pi t \, dt$. We have

$$
\begin{aligned}
&\frac{1}{2\pi} \int_{-\pi}^{0} \frac{\sin \pi x}{1 + \cos \pi x \sin \varphi} \ln |F(h(e^{i\varphi}))| d\varphi \\
=& \frac{1}{2} \int_{-\infty}^{\infty} \frac{\sin \pi x}{\cosh \pi t - \cos \pi x} \ln |F(it)| dt.
\end{aligned} \tag{1.12}
$$

On the interval $(0, \pi]$, we use the change of variables $1 + it = h(e^{i\varphi})$ or, equivalently,

$$\begin{aligned}
e^{i\varphi} =& h^{-1}(1+it) = \frac{e^{\pi i(1+it)}-i}{e^{\pi i(1+it)}+i} = \frac{e^{\pi i}e^{-\pi t}-i}{e^{\pi i}e^{-\pi t}+i} \\
=& \frac{(e^{\pi i}e^{-\pi t}-i)(e^{-\pi i}e^{-\pi t}-i)}{e^{-2\pi t}+1} \\
=& \frac{e^{-2\pi t}-1-ie^{-\pi t}(e^{-\pi i}+e^{\pi i})}{1+e^{-2\pi t}} = \frac{e^{-\pi t}-e^{\pi t}+2i}{e^{\pi t}+e^{-\pi t}} \\
=& -\tanh \pi t + i \operatorname{sech} \pi t.
\end{aligned}$$

Observe that as φ ranges from 0 to π, t ranges from $-\infty$ to $+\infty$. Furthermore, $d\varphi = \pi \operatorname{sech} \pi t\, dt$. Similarly, we obtain

$$\begin{aligned}
&\frac{1}{2\pi}\int_0^\pi \frac{\sin \pi x}{1+\cos \pi x \sin \varphi} \ln |F(h(e^{i\varphi}))| d\varphi \\
=& \frac{1}{2}\int_{-\infty}^{\infty} \frac{\sin \pi x}{\cosh \pi t + \cos \pi x} \ln |F(1+it)| dt. \qquad (1.13)
\end{aligned}$$

Adding (1.12) and (1.13) and using (1.11), we conclude the proof when $y = 0$.

We now consider the case when $y \neq 0$. Fix $y \neq 0$ and define the function $G(z) = F(z+iy)$. Then, G is analytic on the open strip $\mathring{S} = \{z \in \mathbb{C} : 0 < \operatorname{Re} z < 1\}$ and continuous on its closure S. Moreover, for some $A < \infty$ and $a \in [0, \pi)$, we have

$$\ln |G(z)| = \ln |F(z+iy)| \leqslant Ae^{a|\operatorname{Im} z + y|} \leqslant Ae^{a|y|}e^{a|\operatorname{Im} z|}$$

for all $z \in S$. Then, the case $y = 0$ for G (with A replaced by $Ae^{a|y|}$) yields

$$|G(x)| \leqslant \exp\left\{\frac{\sin \pi x}{2}\int_{-\infty}^{\infty}\left[\frac{\ln |G(it)|}{\cosh \pi t - \cos \pi x} + \frac{\ln |G(1+it)|}{\cosh \pi t + \cos \pi x}\right] dt\right\},$$

which yields the required conclusion for any real y, since $G(x) = F(x+iy)$, $G(it) = F(it+iy)$, and $G(1+it) = F(1+it+iy)$. ❑

The extension of the Riesz-Thorin interpolation theorem is now stated.

Theorem 1.8 (Stein interpolation theorem)**.** *Let (X, μ) and (Y, ν) be a pair of σ-finite measure spaces. 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*

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

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

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \text{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \qquad (1.14)$$

Suppose that

$$\|T_{it}(f)\|_{q_0} \leqslant M_0(t)\|f\|_{p_0}, \quad \|T_{1+it}(f)\|_{q_1} \leqslant M_1(t)\|f\|_{p_1} \tag{1.15}$$

for all finitely simple functions f on X. Then,

$$\|T_\theta(f)\|_q \leqslant M(\theta)\|f\|_p, \quad \text{when } 0<\theta<1 \tag{1.16}$$

for all simple finitely functions f on X, where

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

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 (1.14).

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 $F(z)=\langle T\eta(z),\zeta(z)\rangle$, and 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. We leave it as an exercise.

§1.3 Distribution functions and weak L^p

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

Definition 1.9. Let f be a measurable function on $(X,\mathcal{M},\mu)$; we define its ***distribution function*** $f_*:[0,\infty)\mapsto[0,\infty]$ by

$$f_*(\alpha)=\mu(\{x\in X:|f(x)|>\alpha\}).$$

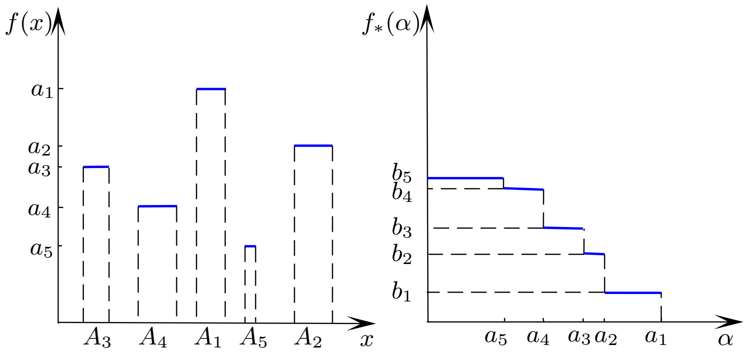
The distribution function f_* provides information about the size of f but not about the behavior of f itself near any given point. For instance, a function on $\mathbb{R}^n$ and each of its translations have the same distribution function.

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 in $f_*(\alpha)$ 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.

Example 1.10 (Distribution function 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 set A (see Figure (a)). Then,



(a) (b)

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

where $b_j = \sum\limits_{i=1}^{j} |A_i|$, $B_j = [a_{j+1}, a_j)$ 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)).

Example 1.11. 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

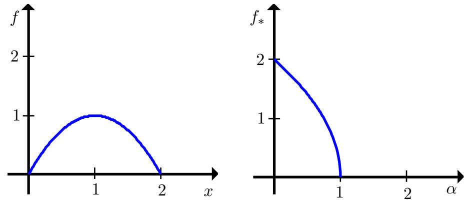
$$\begin{aligned} f_*(\alpha) =& |\{x \in [0, \infty) : 1 - (x-1)^2 > \alpha\}| \\ =& |\{x \in [0, \infty) : 1 - \sqrt{1-\alpha} < x < 1 + \sqrt{1-\alpha}\}| = 2\sqrt{1-\alpha}. \end{aligned}$$

That is,

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

Observe that the integrals of f and f_* are the same, i.e.,

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



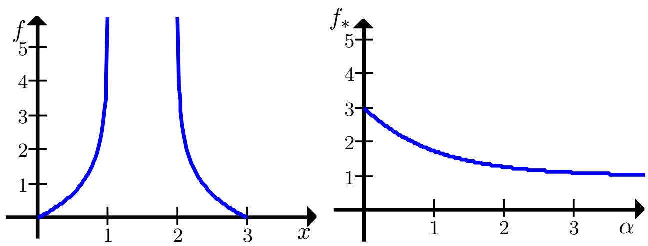
(a) (b)

Example 1.12. 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 is still well-defined and can be calculated for any $\alpha \geqslant 0$

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



(a) (b)

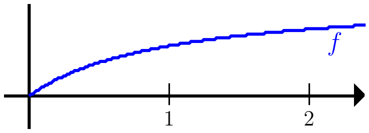
Example 1.13. Consider the function $f(x) = x$ for all $x \in [0, \infty)$. Then, $f_*(\alpha) = |\{x \in [0, \infty) : x > \alpha\}| = \infty$ for all $\alpha \geqslant 0$.

Example 1.14. 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

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

That is,

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



Now, we give some basic properties of distribution functions.

Proposition 1.15. *Let f and g be measurable functions on $(X, \mathcal{M}, \mu)$. Then for all $\alpha, \beta > 0$, we have*

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

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

(iii) *$(cf)_*(\alpha) = f_*(\alpha/|c|)$, for all $c \in \mathbb{C} \setminus \{0\}$.*

(iv) *If $|f(x)| \leqslant |g(x)| + |h(x)|$, then $f_*(\alpha+\beta) \leqslant g_*(\alpha) + h_*(\beta)$.*

(v) *$(fg)_*(\alpha\beta) \leqslant f_*(\alpha) + g_*(\beta)$.*

(vi) *(Chebyshev's inequality) For any $p \in (0, \infty)$ and $\alpha > 0$, it holds*

$$f_*(\alpha) \leqslant \left(\frac{\|f\|_p}{\alpha}\right)^p.$$

(vii) *If $f \in L^p$, $p \in (0, \infty)$, then*

$$\lim_{\alpha \to +\infty} \alpha^p f_*(\alpha) = 0 = \lim_{\alpha \to 0} \alpha^p f_*(\alpha).$$

(viii) *If $\int_0^\infty \alpha^{p-1} f_*(\alpha) d\alpha < \infty$, $p \in (0, \infty)$, then*

$$\lim_{\alpha \to +\infty} \alpha^p f_*(\alpha) = 0 = \lim_{\alpha \to 0} \alpha^p f_*(\alpha).$$

(ix) *If $|f(x)| \leqslant \liminf_{k \to \infty} |f_k(x)|$ for a.e. x, then*

$$f_*(\alpha) \leqslant \liminf_{k \to \infty} (f_k)_*(\alpha).$$

(x) *If $|f_k|$ increases to $|f|$, then $(f_k)_*$ increases to f_*.*

Proof. (i) For simplicity, denote $E_f(\alpha) = \{x \in X : |f(x)| > \alpha\}$ for $\alpha > 0$. Let $\{\alpha_k\}$ be a decreasing positive sequence that tends to α; then, we have $E_f(\alpha) = \bigcup_{k=1}^\infty E_f(\alpha_k)$. Since $\{E_f(\alpha_k)\}$ is a increasing sequence of sets, it follows that $\lim_{k \to \infty} f_*(\alpha_k) = f_*(\alpha)$. This implies the continuity of $f_*(\alpha)$ on the right.

(v) Noticing that

$$\{x \in X : |f(x)g(x)| > \alpha\beta\} \subset \{x \in X : |f(x)| > \alpha\} \cup \{x \in X : |g(x)| > \beta\},$$

we have the desired result.

(vi) We have

$$\begin{aligned} f_*(\alpha) =& \mu(\{x : |f(x)| > \alpha\}) \\ =& \int_{\{x \in X : |f(x)| > \alpha\}} d\mu(x) \\ \leqslant& \int_{\{x \in X : |f(x)| > \alpha\}} \left(\frac{|f(x)|}{\alpha}\right)^p d\mu(x) \end{aligned}$$

$$\leqslant \left(\frac{\|f\|_p}{\alpha}\right)^p.$$

(vii) From (vi), it follows that

$$\alpha^p f_*(\alpha) \leqslant \int_{\{x\in X:|f(x)|>\alpha\}} |f(x)|^p d\mu(x) \leqslant \int |f(x)|^p d\mu(x).$$

Thus, $\mu(\{x \in X : |f(x)| > \alpha\}) \to 0$ as $\alpha \to +\infty$ and

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

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

$$\begin{aligned}\lim_{\alpha\to0} \alpha^p f_*(\alpha) &= \lim_{\alpha\to0} \alpha^p (f_*(\alpha) - f_*(\beta)) \\ &= \lim_{\alpha\to0} \alpha^p \mu(\{x \in X : \alpha < |f(x)| \leqslant \beta\}) \\ &\leqslant \int_{\{x\in X:|f(x)|\leqslant\beta\}} |f(x)|^p d\mu(x).\end{aligned}$$

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

(ix) Let $E = \{x \in X : |f(x)| > \alpha\}$ and $E_k = \{x \in X : |f_k(x)| > \alpha\}$, $k \in \mathbb{N}$. By the assumption and the definition of the 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 $\ell \geqslant 1$,

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

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

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

(x) If $|f_k|$ increases to $|f|$, then $E_f(\alpha)$ is the increasing union of $\{E_{f_k}(\alpha)\}$, so $(f_k)_*$ increases to f_*.

For others, they are easy to verify. ❑

In view of (i) in Proposition 1.15, f_* defines a negative Borel measure ν on $(0, \infty)$ such that $\nu((a, b]) = f_*(b) - f_*(a)$ whenever $0 < a < b$. We can therefore consider the Lebesgue-Stieltjes integrals $\int \varphi df_* = \int \varphi d\nu$ of functions φ on $(0, \infty)$. The following result shows that the integrals of the functions of $|f|$ on X can be reduced to Lebesgue-Stieltjes integrals.

Theorem 1.16. *If $f_*(\alpha) < \infty$ for all $\alpha > 0$ and φ is a nonnegative Borel measurable function on $[0, \infty)$ such that $\varphi(0) = 0$, then*

$$\int_X \varphi(|f|)d\mu = -\int_0^\infty \varphi(\alpha)df_*(\alpha). \tag{1.17}$$

Proof. If ν is the negative measure determined by f_*, we have

$$\nu((a,b]) = f_*(b) - f_*(a) = -\mu(\{x : a < |f(x)| \leqslant b\}) = -\mu(|f|^{-1}((a,b]).$$

It follows that $\nu(E) = -\mu(|f|^{-1}(E))$ for all Borel sets $E \subset (0, \infty)$, by the uniqueness of extensions (cf. [Fol99, Theorem 1.14]). However, this means (1.17) when φ is the characteristic function of a Borel set and hence when φ is simple. The general case then follows by virtue of Theorem 0.1 and the monotone convergence theorem. ❑

The case of this result in which we are most interested is $\varphi(\alpha) = \alpha^p$, which gives

$$\int |f|^p d\mu = -\int_0^\infty \alpha^p df_*(\alpha). \tag{1.18}$$

A more useful form of this equation is obtained by integrating the right side by parts as follows.

Theorem 1.17 (The equivalent norm of L^p). *If $0 < p < \infty$, then*

$$\int |f|^p d\mu = p\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha.$$

Proof. If $f_*(\alpha) = \infty$ for some $\alpha > 0$, then the values of both sides are infinite, and this is clearly true. If not, and f is simple, then for either $f \in L^p(X)$ or $\int_0^\infty \alpha^{p-1} f_*(\alpha)d\alpha < \infty$, we always have $\alpha^p f_*(\alpha) \to 0$ as $\alpha \to +\infty$ and $\alpha \to 0$ by the property (vii) and (viii) in Proposition 1.15, so the integration by parts described above works in r.h.s. of (1.18). Therefore, we have

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

For the general case, let $\{g_n\}$ be a sequence of simple functions that increases to $|f|$; then, the desired result is true for g_n, and it follows for f by (x) in Proposition 1.15 and the monotone convergence theorem. ❑

A variant of the L^p spaces that turns up rather often is the following. Using the distribution function f_*, we now introduce the weak L^p-spaces denoted by $L^{p,\infty}$.

Definition 1.18 (*Weak L^p-space*). For $0 < p < \infty$, the space $L^{p,\infty}(X, \mu)$

consists of all μ-measurable functions f such that

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

In the limiting case $p=\infty$, we put $L^{\infty,\infty}=L^\infty$.

Two functions in $L^{p,\infty}(X,\mu)$ are considered equal if they are equal μ-a.e. We can show that $L^{p,\infty}$ is a quasinormed linear space.

1° If $\|f\|_{L^{p,\infty}}=0$, then for any $\alpha>0$, it holds that $\mu(\{x\in X:|f(x)|>\alpha\})=0$; thus, $f=0$, μ-a.e.

2° From (iii) in Proposition 1.15, we can show that for any $k\in\mathbb{C}\setminus\{0\}$

$$\begin{aligned}\|kf\|_{L^{p,\infty}} =& \sup_{\alpha>0} \alpha (kf)_*^{1/p}(\alpha) = \sup_{\alpha>0} \alpha f_*^{1/p}(\alpha/|k|)\\ =&|k| \sup_{\alpha>0} \alpha f_*^{1/p}(\alpha) = |k|\|f\|_{L^{p,\infty}},\end{aligned}$$

and it is clear that $\|kf\|_{L^{p,\infty}}=|k|\|f\|_{L^{p,\infty}}$ also holds for $k=0$.

3° By part (iv) in Proposition 1.15, we have

$$\begin{aligned}\|f+g\|_{L^{p,\infty}} =& \sup_{\alpha>0} \alpha (f+g)_*^{\frac{1}{p}}(\alpha)\\ \leqslant& \sup_{\alpha>0} \alpha \left(f_*\left(\frac{\alpha}{2}\right)+g_*\left(\frac{\alpha}{2}\right)\right)^{\frac{1}{p}}\\ \leqslant& \max(2^{\frac{1}{p}},2) \sup_{\alpha>0} \frac{\alpha}{2}\left(f_*^{\frac{1}{p}}\left(\frac{\alpha}{2}\right)+g_*^{\frac{1}{p}}\left(\frac{\alpha}{2}\right)\right)\\ \leqslant& \max(2^{\frac{1}{p}},2)\left(\sup_{\alpha>0}\alpha f_*^{\frac{1}{p}}(\alpha)+\sup_{\alpha>0}\alpha g_*^{\frac{1}{p}}(\alpha)\right)\\ \leqslant& \max(2^{\frac{1}{p}},2)(\|f\|_{L^{p,\infty}}+\|g\|_{L^{p,\infty}}).\end{aligned}$$

Thus, $L^{p,\infty}$ is a quasinormed linear space.

The weak L^p spaces are larger than the usual L^p spaces. We have the following:

Theorem 1.19. *For any $0<p\leqslant\infty$ and any $f\in L^p(X,\mu)$, we have*

$$\|f\|_{L^{p,\infty}} \leqslant \|f\|_p.$$

Hence, $L^p(X,\mu)\hookrightarrow L^{p,\infty}(X,\mu)$.

Proof. It is clear for $p=\infty$. For $p\in(0,\infty)$, from the part (vi) in Proposition 1.15, we have

$$\alpha f_*^{1/p}(\alpha) \leqslant \left(\int_{\{x\in X:|f(x)|>\alpha\}} |f(x)|^p d\mu(x)\right)^{1/p} \leqslant \|f\|_p,$$

which yields the desired result. ❑

The inclusion $L^p \hookrightarrow L^{p,\infty}$ is strict for $0 < p < \infty$. For example, on $\mathbb{R}^n$ with the usual Lebesgue measure, let $h(x) = |x|^{-n/p}$. Obviously, h is not in $L^p(\mathbb{R}^n)$ due to

$$\int |x|^{-n} dx = \omega_{n-1} \int_0^\infty r^{-n} r^{n-1} dr = \infty,$$

where $\omega_{n-1} = 2\pi^{n/2}/\Gamma(n/2)$ is the surface area of the unit sphere $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$, but h is in $L^{p,\infty}(\mathbb{R}^n)$ and we may check easily that

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

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 $\operatorname{Re} z > 0$.

Convergence in measure is a weaker notion than convergence in either L^p or $L^{p,\infty}$, $0 < p \leqslant \infty$, as the following proposition indicates:

Proposition 1.20. *Let $p \in (0, \infty]$ and $f_n, f \in L^{p,\infty}(X, \mu)$.*
(i) *If $f_n, f \in L^p$ and $f_n \to f$ in L^p, then $f_n \to f$ in $L^{p,\infty}$.*
(ii) *If $f_n \to f$ in $L^{p,\infty}$, then $f_n \xrightarrow{\mu} f$.*

Proof. For $p \in (0, \infty)$, Theorem 1.19 gives that

$$\|f_n - f\|_{L^{p,\infty}} \leqslant \|f_n - f\|_p,$$

which implies (i) for the case $p \in [1, \infty)$. The case $p = \infty$ is trivial due to $L^{\infty,\infty} = L^\infty$.

For (ii), given $\varepsilon > 0$, there exists an n_0 such that for $n > n_0$,

$$\|f_n - f\|_{L^{p,\infty}} = \sup_{\alpha>0} \alpha\mu(\{x \in X : |f_n(x) - f(x)| > \alpha\})^{\frac{1}{p}} < \varepsilon^{1+\frac{1}{p}}.$$

Taking $\alpha = \varepsilon$, we obtain the desired result. ❑

Remark 1.21. Note that there is no general converse of statement (ii) in the above proposition. Fix $p \in [1, \infty)$ and on $[0, 1]$, we define the functions

$$f_{k,j} = k^{1/p} \chi_{(\frac{j-1}{k}, \frac{j}{k})}, \quad 1 \leqslant j \leqslant k.$$

Consider the sequence $\{f_{1,1}, f_{2,1}, f_{2,2}, f_{3,1} f_{3,2}, f_{3,3}, \cdots\}$. Observe that

$$|\{x : f_{k,j}(x) > 0\}| = 1/k \to 0, \text{ as } k, j \to \infty.$$

Therefore, $f_{k,j} \xrightarrow{\mu} 0$. Similarly, we have

$$\begin{aligned}\|f_{k,j}\|_{L^{p,\infty}} =& \sup_{\alpha>0} \alpha|\{x : f_{k,j}(x) > \alpha\}|^{1/p}\\ =& \sup_{\alpha>0} \alpha|\{x : k^{1/p} \chi_{(\frac{j-1}{k}, \frac{j}{k})}(x) > \alpha\}|^{1/p}\end{aligned}$$

$$= \sup_{\alpha>0} \alpha \left| \left\{ x \in \left(\frac{j-1}{k}, \frac{j}{k} \right) : k^{1/p} > \alpha \right\} \right|^{1/p}$$
$$= \sup_{0<\alpha<k^{1/p}} \alpha(1/k)^{1/p}$$
$$\geqslant \sup_{k\geqslant 1} \left(1 - \frac{1}{k^2}\right)^{1/p} \quad (\text{taking } \alpha = (k-1/k)^{1/p})$$
$$=1,$$

which implies that $f_{k,j}$ does not converge to 0 in $L^{p,\infty}$.

It turns out that every sequence convergent in $L^p(X,\mu)$ or in $L^{p,\infty}(X,\mu)$ has a subsequence that converges a.e. to a limit in view of Theorem 0.4.

A sequence $\{f_k\}_{k=1}^{\infty} \subset L^{p,\infty}$ is Cauchy if $\|f_k - f_m\|_{L^{p,\infty}} \to 0$ as $k, m \to \infty$. We now have

Theorem 1.22. *For each $p \in (0,\infty]$, the space $L^{p,\infty}$ is complete.*

Proof. Since $L^{\infty,\infty} = L^{\infty}$, we will focus on $p \in (0,\infty)$. Let $\{f_k\}_{k=1}^{\infty}$ be a Cauchy sequence in $L^{p,\infty}$; then, $\{f_k\}$ is Cauchy in measure by Proposition 1.20. Thus, it has a subsequence $\{f_{k_j}\}$ that converges a.e. to some f by Theorem 0.4. Fixed j_0 and apply (ix) in Proposition 1.15. Since $|f - f_{k_{j_0}}| = \lim_{j\to\infty} |f_{k_j} - f_{k_{j_0}}|$, it follows that

$$(f - f_{k_{j_0}})_*(\alpha) \leqslant \liminf_{j\to\infty} (f_{k_j} - f_{k_{j_0}})_*(\alpha).$$

Thus,

$$\|f - f_{k_{j_0}}\|_{L^{p,\infty}} \leqslant \liminf_{j\to\infty} \|f_{k_j} - f_{k_{j_0}}\|_{L^{p,\infty}}.$$

Let $j_0 \to \infty$ and use the fact that $\{f_k\}$ is Cauchy to conclude that f_{k_j} converges to f in $L^{p,\infty}$. It follows that f_k converges to f in $L^{p,\infty}$ by the triangle inequality for the quasinorm. ❑

It is a useful fact that a function $f \in L^p(X,\mu) \cap L^q(X,\mu)$ with $p < q$ implies $f \in L^r(X,\mu)$ for all $r \in (p,q)$. The usefulness of the spaces $L^{p,\infty}$ can be seen from the following sharpening of this statement:

Proposition 1.23 (Interpolation of $L^{p,\infty}$ spaces)**.** *Let $1 \leqslant p < q \leqslant \infty$ and $f \in L^{p,\infty}(X,\mu) \cap L^{q,\infty}(X,\mu)$, where X is a σ-finite measure space. Then, $f \in L^r(X,\mu)$ for all $r \in (p,q)$ (i.e., $\theta \in (0,1)$) and*

$$\|f\|_r \leqslant \left(\frac{r}{r-p} + \frac{r}{q-r} \right)^{1/r} \|f\|_{L^{p,\infty}}^{1-\theta} \|f\|_{L^{q,\infty}}^{\theta}, \tag{1.19}$$

with the interpretation that $1/\infty = 0$, *where*

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

Proof. We first consider the case $q < \infty$. From Theorem 1.17 and the definition of the distribution function, it follows that

$$\begin{aligned} \|f\|_r^r =& r \int_0^\infty \alpha^{r-1} f_*(\alpha) d\alpha \\ \leqslant& r \int_0^\infty \alpha^{r-1} \min\left(\frac{\|f\|_{L^{p,\infty}}^p}{\alpha^p}, \frac{\|f\|_{L^{q,\infty}}^q}{\alpha^q}\right) d\alpha. \end{aligned} \tag{1.20}$$

We take a suitable α such that $\frac{\|f\|_{L^{p,\infty}}^p}{\alpha^p} \leqslant \frac{\|f\|_{L^{q,\infty}}^q}{\alpha^q}$, i.e., $\alpha \leqslant \left(\frac{\|f\|_{L^{q,\infty}}^q}{\|f\|_{L^{p,\infty}}^p}\right)^{\frac{1}{q-p}} =: B$. Then, we obtain

$$\begin{aligned} \|f\|_r^r \leqslant& r \int_0^B \alpha^{r-1-p} \|f\|_{L^{p,\infty}}^p d\alpha + r \int_B^\infty \alpha^{r-1-q} \|f\|_{L^{q,\infty}}^q d\alpha \\ =& \frac{r}{r-p} \|f\|_{L^{p,\infty}}^p B^{r-p} + \frac{r}{q-r} \|f\|_{L^{q,\infty}}^q B^{r-q} \quad (\text{due to } p < r < q) \\ =& \left(\frac{r}{r-p} + \frac{r}{q-r}\right) \|f\|_{L^{p,\infty}}^{r(1-\theta)} \|f\|_{L^{q,\infty}}^{r\theta}. \end{aligned}$$

For the case $q = \infty$, because $f_*(\alpha) = 0$ for $\alpha > \|f\|_\infty$, we only use the inequality $f_*(\alpha) \leqslant \alpha^{-p} \|f\|_{L^{p,\infty}}^p$ for $\alpha \leqslant \|f\|_\infty$ for the integral in (1.20) to obtain

$$\begin{aligned} \|f\|_r^r \leqslant& r \int_0^{\|f\|_\infty} \alpha^{r-1-p} \|f\|_{L^{p,\infty}}^p d\alpha \\ =& \frac{r}{r-p} \|f\|_{L^{p,\infty}}^p \|f\|_\infty^{r-p}, \end{aligned}$$

which implies the result since $p = r(1-\theta)$ and $L^{\infty,\infty} = L^\infty$. ❑

Frequently, it is convenient to express a function as the sum of a "small" part and a "big" part. The following is a way of doing this that gives a simple formula for the distribution functions.

Proposition 1.24. *If* f *is a measurable function and* $N > 0$, *let* $E(N) = \{x : |f(x)| > N\}$, *and set*

$$\begin{aligned} h_N =& f\chi_{X\setminus E(N)} + N(\operatorname{sgn} f)\chi_{E(N)}, \\ g_N =& f - h_N = (\operatorname{sgn} f)(|f| - N)\chi_{E(N)}. \end{aligned}$$

Then,

$$(g_N)_*(\alpha) = f_*(\alpha + N), \quad (h_N)_*(\alpha) = \begin{cases} f_*(\alpha) & \text{if } \alpha < N, \\ 0, & \text{if } \alpha \geqslant N. \end{cases}$$

We leave the proof as an exercise.

§1.4 Marcinkiewicz interpolation theorem

We now turn to the Marcinkiewicz interpolation theorem, for which we need some more terminology. Let T be an operator from some vector space $\mathcal{D}$ of measurable functions on $(X, \mathcal{M}, \mu)$ to the space of all measurable functions on $(Y, \mathcal{N}, \nu)$.

Definition 1.25. (i) T is called ***quasilinear*** if $|T(f+g)| \leqslant K(|Tf| + |Tg|)$ and $|T(\lambda f)| = |\lambda||Tf|$ for all $f, g \in \mathcal{D}$, where $K \geqslant 1$ is a positive constant independent of f and g. If $K = 1$, then T is called ***sublinear***.

(ii) A quasilinear operator T is of ***strong type*** (p, q) $(1 \leqslant p, q \leqslant \infty)$ if $L^p(X, d\mu) \subset \mathcal{D}$, T maps $L^p(X, d\mu)$ into $L^q(Y, d\nu)$, and there exists $C > 0$ such that $\|Tf\|_q \leqslant C\|f\|_p$ for all $f \in L^p(X, d\mu)$.

(iii) A quasilinear operator T is of ***weak type*** (p, q) $(1 \leqslant p \leqslant \infty$, $1 \leqslant q < \infty)$ if $L^p(X, d\mu) \subset \mathcal{D}$, T maps $L^p(X, d\mu)$ into $L^{q,\infty}(Y, d\nu)$, and there exists $C > 0$ such that $\|Tf\|_{L^{q,\infty}} \leqslant C\|f\|_p$ for all $f \in L^p(X, d\mu)$. Additionally, we shall say that T is of weak type (p, ∞) iff T is strong type (p, ∞).

Now, we give the Marcinkiewicz interpolation theorem.[2]

Theorem 1.26 (Marcinkiewicz interpolation theorem). *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are measure spaces. Assume that $1 \leqslant p_j \leqslant q_j \leqslant \infty$ for $j = 0, 1$, $q_0 \neq q_1$ and*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \text{ and } \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}, \quad \text{where } 0 < \theta < 1.$$

If T is a quasi-linear map from $L^{p_0}(X) + L^{p_1}(X)$ to the space of measurable functions on Y that is of weak types (p_0, q_0) and (p_1, q_1), then T is of strong type (p, q). More precisely, if $\|Tf\|_{L^{q_j,\infty}(Y)} \leqslant A_j\|f\|_{p_j}$ for $j = 0, 1$, then $\|Tf\|_q \leqslant A_p\|f\|_p$ where A_p depends only on A_j, p_j, q_j, θ, and for $j = 0, 1$, $A_p|p - p_j|$ (resp. A_p) remains bounded as $p \to p_j$ if $p_j < \infty$ (resp. $p_j = \infty$).

[2] J. Marcinkiewicz (1910–1940) was a Polish mathematician and a student of A. Zygmund. The theorem was first announced by Marcinkiewicz (1939), who showed this result to A. 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.

Proof. The case $p_0 = p_1$ is easy and is left to the reader (Exercise 1.6). Without loss of generality, we may therefore assume that $p_0 < p_1$, and for the time being, we also assume that $q_0 < \infty$ and $q_1 < \infty$ (where also $p_0 < p_1 < \infty$). Given $f \in L^p(X)$ and $N > 0$, let g_N and h_N be as in Proposition 1.24. Then, by Theorem 1.17 and Proposition 1.24,

$$\begin{aligned}\int |g_N|^{p_0} d\mu =& p_0 \int_0^\infty \beta^{p_0-1}(g_N)_*(\beta)d\beta = p_0 \int_0^\infty \beta^{p_0-1} f_*(\beta+N)d\beta \\ =& p_0 \int_N^\infty (\beta-N)^{p_0-1} f_*(\beta)d\beta \leqslant p_0 \int_N^\infty \beta^{p_0-1} f_*(\beta)d\beta, \quad (1.21)\\ \int |h_N|^{p_1} d\mu =& p_1 \int_0^\infty \beta^{p_1-1}(h_N)_*(\beta)d\beta = p_1 \int_0^N \beta^{p_1-1} f_*(\beta)d\beta, \quad (1.22)\end{aligned}$$

and

$$\int |Tf|^q d\nu = q \int_0^\infty \alpha^{q-1}(Tf)_*(\alpha)d\alpha = (2K)^q q \int_0^\infty \alpha^{q-1}(Tf)_*(2K\alpha)d\alpha. \quad (1.23)$$

Since T is quasilinear, by (iv) and (iii) in Proposition 1.15, we have

$$(Tf)_*(2K\alpha) \leqslant (KTg_N)_*(K\alpha) + (KTh_N)_*(K\alpha) = (Tg_N)_*(\alpha) + (Th_N)_*(\alpha). \quad (1.24)$$

Then, by (1.21)-(1.24), and the weak type estimates of T, we obtain

$$\begin{aligned}\|Tf\|_q^q \leqslant& (2K)^q q \int_0^\infty \alpha^{q-1} \left[(A_0\|g_N\|_{p_0}/\alpha)^{q_0} + (A_1\|h_N\|_{p_1}/\alpha)^{q_1}\right] d\alpha\\ \leqslant& (2K)^q q A_0^{q_0} p_0^{q_0/p_0} \int_0^\infty \alpha^{q-q_0-1} \left[\int_N^\infty \beta^{p_0-1} f_*(\beta)d\beta\right]^{q_0/p_0} d\alpha\\ &+ (2K)^q q A_1^{q_1} p_1^{q_1/p_1} \int_0^\infty \alpha^{q-q_1-1} \left[\int_0^N \beta^{p_1-1} f_*(\beta)d\beta\right]^{q_1/p_1} d\alpha\\ =& \sum_{j=0}^{1} (2K)^q q A_j^{q_j} p_j^{q_j/p_j} \int_0^\infty \left[\int_0^\infty \phi_j(N,\beta)d\beta\right]^{q_j/p_j} d\alpha, \quad (1.25)\end{aligned}$$

where χ_0 and χ_1 denote the characteristic functions of $\{(N,\beta) : \beta > N\}$ and $\{(N,\beta) : \beta < N\}$,

$$\phi_j(N,\beta) = \chi_j(N,\beta)\alpha^{(q-q_j-1)p_j/q_j}\beta^{p_j-1} f_*(\beta).$$

Since $q_j/p_j \geqslant 1$, we may apply Minkowski's inequality for integrals to obtain

$$\int_0^\infty \left[\int_0^\infty \phi_j(N,\beta)d\beta\right]^{q_j/p_j} d\alpha \leqslant \left[\int_0^\infty \left[\int_0^\infty \phi_j(N,\beta)^{q_j/p_j} d\alpha\right]^{p_j/q_j} d\beta\right]^{q_j/p_j}. \quad (1.26)$$

Since (1.24) is true for all $\alpha > 0$ and $N > 0$, we may take N to depend on α, say $N = \gamma(\alpha)$ or $\alpha = \gamma^{-1}(N)$ for some bijective map γ. In addition, we have

$$\int_0^\infty \left[\int_0^\infty \phi_0(N,\beta)^{q_0/p_0} d\alpha\right]^{p_0/q_0} d\beta$$
$$= \int_0^\infty \left[\int_0^\infty \chi_0(N,\beta)\alpha^{(q-q_0-1)} d\alpha\right]^{p_0/q_0} \beta^{p_0-1} f_*(\beta) d\beta. \quad (1.27)$$

If $q_1 > q_0$, then $q - q_0$ is positive. Therefore, we need to assume that $\gamma(\alpha)$ is increasing with respect to α in order to make the inner integral of the above converge. Then, the inequality $\beta > \gamma(\alpha)$ is equivalent to $\alpha < \gamma^{-1}(\beta)$. Thus, we obtain

$$(1.27) = \int_0^\infty \left[\int_0^{\gamma^{-1}(\beta)} \alpha^{(q-q_0-1)} d\alpha\right]^{p_0/q_0} \beta^{p_0-1} f_*(\beta) d\beta$$
$$=(q-q_0)^{-p_0/q_0} \int_0^\infty (\gamma^{-1}(\beta))^{(q-q_0)p_0/q_0} \beta^{p_0-1} f_*(\beta) d\beta. \quad (1.28)$$

Since we expect to control it by $\|f\|_p$, in view of Theorem 1.17, we may take the inverse map γ^{-1} such that

$$(\gamma^{-1}(\beta))^{(q-q_0)p_0/q_0} \beta^{p_0-1} = \beta^{p-1},$$

i.e.,

$$\gamma^{-1}(\beta) = \beta^{\frac{q_0(p-p_0)}{p_0(q-q_0)}}.$$

It follows from the equations defining p and q that

$$\sigma = \frac{p_0(q-q_0)}{q_0(p-p_0)} = \frac{p^{-1}(q^{-1}-q_0^{-1})}{q^{-1}(p^{-1}-p_0^{-1})} = \frac{p^{-1}(q^{-1}-q_1^{-1})}{q^{-1}(p^{-1}-p_1^{-1})} = \frac{p_1(q_1-q)}{q_1(p_1-p)}, \quad (1.29)$$

where σ is positive for the case $q_1 > q_0$, and let $\tau = 1/\sigma$, then $\gamma^{-1}(\beta) = \beta^\tau$, namely, we may choose $N = \alpha^\sigma$. Thus, it follows

$$(1.28) =(q-q_0)^{-p_0/q_0} \int_0^\infty \beta^{p-1} f_*(\beta) d\beta = |q-q_0|^{-p_0/q_0} p^{-1} \|f\|_p^p.$$

On the other hand, if $q_1 < q_0$, then $q - q_0$ and σ are negative, and the inequality $\beta > \alpha^\sigma$ is equivalent to $\alpha > \beta^\tau$, so as above, we obtain

$$(1.27) = \int_0^\infty \left[\int_{\beta^\tau}^\infty \alpha^{q-q_0-1} d\alpha\right]^{p_0/q_0} \beta^{p_0-1} f_*(\beta) d\beta$$
$$=(q-q_0)^{-p_0/q_0} \int_0^\infty \beta^{p-1} f_*(\beta) d\beta$$
$$=|q-q_0|^{-p_0/q_0} p^{-1} \|f\|_p^p.$$

A similar calculation shows that

$$\int_0^\infty \left[\int_0^\infty \phi_1(N,\beta)^{q_1/p_1} d\alpha\right]^{p_1/q_1} d\beta = |q-q_1|^{-p_1/q_1} p^{-1} \|f\|_p^p.$$

Combining these results with (1.25) and (1.26), we see that

$$\sup\{\|Tf\|_q : \|f\|_p = 1\} \leqslant A_p = 2Kq^{1/q}\left[\sum_{j=0}^{1} A_j^{q_j}(p_j/p)^{q_j/p_j}|q-q_j|^{-1}\right]^{1/q}.$$

However, since $|T(\lambda f)| = |\lambda||Tf|$, this implies that $\|Tf\|_q \leqslant A_p\|f\|_p$ for all $f \in L^p(X)$, and we are done.

It remains to be shown how to modify this argument to address the exceptional cases $q_0 = \infty$ or $q_1 = \infty$. We distinguish three cases by noticing the condition $p_j \leqslant q_j$.

Case I: $p_1 = q_1 = \infty$ (so $p_0 \leqslant q_0 < \infty$). Instead of taking $N = \alpha^\sigma$ in the decomposition of f, we take $N = \alpha/A_1$. Then, $\|Th_N\|_\infty \leqslant A_1\|h_N\|_\infty \leqslant \alpha$, so $(Th_N)_*(\alpha) = 0$, and we obtain (1.25) with $\phi_1 = 0$. The same argument as above then gives

$$\|Tf\|_q \leqslant 2K\left[qA_0^{q_0}A_1^{q-q_0}(p_0/p)^{q_0/p_0}|q-q_0|^{-1}\right]^{1/q}\|f\|_p.$$

Case II: $p_0 < p_1 < \infty$, $q_0 < q_1 = \infty$. Again, the idea is to choose N so that $(Th_N)_*(\alpha) = 0$, and the proper choice is $N = (\alpha/d)^\sigma$ where $d = A_1[p_1\|f\|_p^p/p]^{1/p_1}$ and $\sigma = p_1/(p_1-p)$ (the limiting value of the σ defined by (1.29) as $q_1 \to \infty$). Indeed, since $p_1 > p$, we have

$$\begin{aligned}\|Th_N\|_\infty^{p_1} \leqslant& A_1^{p_1}\|h_N\|_{p_1}^{p_1} = A_1^{p_1}p_1\int_0^N \alpha^{p_1-1}f_*(\alpha)d\alpha\\ \leqslant& A_1^{p_1}p_1N^{p_1-p}\int_0^N \alpha^{p-1}f_*(\alpha)d\alpha = A_1^{p_1}\frac{p_1}{p}\left[\frac{\alpha}{d}\right]^{p_1}\|f\|_p^p = \alpha^{p_1}.\end{aligned}$$

As in Case I, then, we find that $\phi_1 = 0$ in (1.25) and the integral involving ϕ_0 is majorized by a constant A_p when $\|f\|_p = 1$, which yields the desired result.

Case III: $p_0 < p_1 < \infty$, $q_1 < q_0 = \infty$. The argument is essentially the same as in Case II, except that we take $N = (\alpha/d)^\sigma$ with d chosen so that $(Tg_N)_*(\alpha) = 0$. ❑

A less superficial generalization of the theorem can be given in terms of the notation of Lorentz spaces, which unifies and generalizes 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 [BL76a, Chapter 5].

Exercises

Exercise 1.1. Prove Theorem 1.8.

Exercise 1.2 (Hölder's inequality for weak spaces[Gra14a, Exercise 1.1.15]). Let f_j be in $L^{p_j,\infty}$ of a measure space X where $p_j \in [1,\infty)$ and $1 \leqslant j \leqslant k$.

Let

$$\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}.$$

Prove that

$$\|f_1 \cdots f_k\|_{L^{p,\infty}} \leqslant p^{-1/p} \left(\prod_{j=1}^{k} p_j^{1/p_j} \right) \left(\prod_{j=1}^{k} \|f_j\|_{L^{p_j,\infty}} \right).$$

Hint Take $\|f_j\|_{L^{p_j,\infty}} = 1$ for all j. Control $(f_1 \cdots f_k)_*(\alpha)$ by

$$\begin{aligned} &\mu(\{|f_1| > \alpha/s_1\}) + \cdots + \mu(\{|f_{k-1}| > s_{k-2}/s_{k-1}\}) + \mu(\{|f_k| > s_{k-1}\}) \\ \leqslant &(s_1/\alpha)^{p_1} + (s_2/s_1)^{p_2} + \cdots + (s_{k-1}/s_{k-2})^{p_{k-1}} + (1/s_{k-1})^{p_k}. \end{aligned}$$

Set $x_1 = s_1/\alpha$, $x_2 = s_2/s_1$, $\cdots$, $x_k = 1/s_{k-1}$. Minimize $x_1^{p_1} + \cdots + x_k^{p_k}$ subject to the constraint $x_1 \cdots x_k = 1/\alpha$.

Exercise 1.3 (Normability of $L^{p,\infty}$ for $p > 1$[Gra14a, Exercises 1.1.11(a), 1.1.12]**).** Let (X, μ) be a σ-finite measure space and let $1 \leqslant p < \infty$. Pick $0 < r < p$ and define

$$|||f|||_{L^{p,\infty}} = \sup_{0<\mu(E)<\infty} \mu(E)^{-\frac{1}{r}+\frac{1}{p}} \left(\int_E |f|^r d\mu \right)^{\frac{1}{r}},$$

where the supremum is takes over all measurable subsets E of X of finite measure.

(i) Show that for $E \subset X$ with $\mu(E) < \infty$,

$$\int_E |f(x)|^r d\mu(x) \leqslant \frac{p}{p-r} \mu(E)^{1-\frac{r}{p}} \|f\|_{L^{p,\infty}}^r,$$

and then

$$|||f|||_{L^{p,\infty}} \leqslant \left(\frac{p}{p-r} \right)^{\frac{1}{r}} \|f\|_{L^{p,\infty}}$$

for all f in $L^{p,\infty}(X, \mu)$. (It is not needed that X is σ-finite here.)

(ii) Prove that for every measurable function f on (X, μ),

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

(iii) Show that $L^{p,\infty}(X, \mu)$ is normable when $p > 1$, i.e., there is a norm on the space equivalent to $\| \cdot \|_{L^{p,\infty}}$.

(iv) Use the characterization of the $L^{p,\infty}$ quasinorm obtained in parts (i) and (ii) to prove Fatou's lemma for this space: For all measurable functions g_n on X, we have

$$\left\| \liminf_{n\to\infty} |g_n| \right\|_{L^{p,\infty}} \leqslant C_p \liminf_{n\to\infty} \|g_n\|_{L^{p,\infty}}$$

for some constant C_p that depends only on $p \in [1, \infty)$.

Hint Part (i): Use $\mu(E \cap \{|f| > \alpha\}) \leqslant \min(\mu(E), \alpha^{-p}\|f\|_{L^{p,\infty}}^p)$. Part (ii): Write $X = \bigcup_{k=1}^\infty X_k$ with $\mu(X_k) < \infty$ and take $E = \{|f| > \alpha\} \cap X_k$.

Exercise 1.4. Prove Proposition 1.24.

Exercise 1.5. Let $p \in (0, \infty]$. If $f \in L^{p,\infty}$ and $\mu(f \neq 0) < \infty$, then $f \in L^q$ for all $q \in (0, p)$. In particular, for measure space $(X, \mathscr{M}, \mu)$,

$$\mu(X) < \infty \Longrightarrow \forall q < p,\ L^{p,\infty} \subset L^q.$$

On the other hand, if $f \in L^{p,\infty} \cap L^\infty$, then $f \in L^q$ for all $q > p$.

Exercise 1.6. [Fol99, Exercise 6.42] Prove the Marcinkiewicz interpolation theorem in the case $p_0 = p_1$.

Hint Setting $p = p_0 = p_1$, we have $(Tf)_*(\alpha) \leqslant (A_0 \|f\|_p / \alpha)^{q_0}$ and $(Tf)_*(\alpha) \leqslant (A_1 \|f\|_p / \alpha)^{q_1}$. Use whichever estimate is better, depending on α, to majorize

$$q \int_0^\infty \alpha^{q-1} (Tf)_*(\alpha) d\alpha.$$

Exercise 1.7. Write out the proof of the Marcinkiewicz interpolation theorem for two special cases: (i) $p_0 = q_0 = 1$, $p_1 = q_1 = 2$, and (ii) $p_0 = q_0 = 1$, $p_1 = q_1 = \infty$.

Exercise 1.8. [Zho99, Example 3, on p.89] Let (X, μ) and (Y, ν) be two σ-finite measure spaces and $q \in (1, \infty)$. Assume that $K(x, y)$ is a measurable function on $X \times Y$ satisfying

$$\|K(x, \cdot)\|_{L^{q,\infty}} \leqslant C, \quad \text{a.e. } x \in X,$$

$$\|K(\cdot, y)\|_{L^{q,\infty}} \leqslant C, \quad \text{a.e. } y \in Y.$$

To show that if $f \in L^p(Y)$ for $p \in [1, \infty)$, then the integral

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

converges for a.e. $x \in X$, and for $1 < p < r < \infty$ and $1/p + 1/q = 1/r + 1$, and T is of weak type $(1, q)$ and of type (p, r).

Exercise 1.9. [Gra14a, Exercise 1.3.2] Let (X, μ) and (Y, ν) be two σ-finite measure spaces. Let $1 < p < r \leqslant \infty$ and suppose that T is a linear operator defined on the space $L^1(X) + L^\infty(X)$ and taking values in the space of measurable functions on Y. Assume that T maps $L^1(X)$ to $L^{1,\infty}(Y)$ with norm A_0 and $L^r(X)$ to $L^r(Y)$ with norm A_1. Prove that T maps $L^p(X)$ to $L^p(Y)$ with norm at most

$$C(p-1)^{-\frac{1}{p}} A_0^{\frac{\frac{1}{p} - \frac{1}{r}}{1 - \frac{1}{r}}} A_1^{\frac{1 - \frac{1}{p}}{1 - \frac{1}{r}}}.$$

Hint First interpolate between L^1 and L^r using the Marcinkiewicz interpolation theorem and then interpolate between $L^{\frac{p+1}{2}}$ and L^r using the Riesz-Thorin interpolation theorem.

2 Maximal Functions and Calderón-Zygmund Decomposition

Maximal functions appear in many forms in harmonic analysis, such as the Hardy-Littlewood maximal function, dyadic maximal function, and nontangential maximal function [CM85; CM86]. One of the most important of these is the Hardy-Littlewood maximal function. It plays an important role in understanding, for example, the differentiability properties of functions, singular integrals and partial differential equations. It often provides a deeper and more simplified approach to understanding problems in these areas than other methods. We also introduce the Calderón-Zygmund decomposition as an application of maximal functions.

§ 2.1 Hardy-Littlewood maximal function

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 address $F'(x)$, 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 (e.g., 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 open ball $B(x,r)$ of radius r, centered at x, and denote its measure by $|B(x,r)|$. It follows

$$\lim_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy = f(x), \quad \text{for a.e. } x? \tag{2.1}$$

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

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

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

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

as desired.

In general, for this "averaging problem" (2.1), we shall have an affirmative answer. To study the limit (2.1), we consider its quantitative analog, where "$\lim\limits_{r\to 0}$" is replaced by "$\sup\limits_{r>0}$", which is the ***(centered) maximal function***. Since the properties of this maximal function are expressed in terms of relative size and do not involve any cancellation of positive and negative values, we replace f by $|f|$.

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 $L^1_{\text{loc}}(\mathbb{R}^n)$ the space of all locally integrable functions. Loosely speaking, the behavior at infinity does not affect the local integrability of a function. For example, the functions $e^{|x|}$ and $|x|^{-1/2}$ are both locally integrable but not integrable on $\mathbb{R}^n$.

Definition 2.1. If f is locally integrable on $\mathbb{R}^n$, we define its ***maximal function*** $Mf : \mathbb{R}^n \to [0,\infty]$ by

$$Mf(x) = \sup_{r>0}\frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y)|dy, \quad x \in \mathbb{R}^n, \tag{2.2}$$

where the supremum takes over all open balls $B(x,r)$ centered at x. Moreover, M is also called the ***centered Hardy-Littlewood maximal operator***.

It is immediate from the definition that

Theorem 2.2. *If $f \in L^\infty(\mathbb{R}^n)$, then $Mf \in L^\infty(\mathbb{R}^n)$ and*

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

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

$$M'f(x) = \sup_{r>0} \frac{1}{(2r)^n} \int_{Q(x,r)} |f(y)|dy, \quad x \in \mathbb{R}^n. \tag{2.3}$$

When $n = 1$, M and M' coincide. If $n > 1$, then

$$V_n 2^{-n} Mf(x) \leqslant M'f(x) \leqslant V_n 2^{-n} n^{n/2} Mf(x). \tag{2.4}$$

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

$$M''f(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)|dy, \tag{2.5}$$

where the supremum is taken over all cubes containing x. Again, M'' is pointwise equivalent to M; indeed, $V_n 2^{-n} Mf(x) \leqslant M''f(x) \leqslant V_n n^{n/2} Mf(x)$. One sometimes distinguishes between M' and M'' by referring to the former as the centered operator and the latter as the noncentered maximal operator.

Alternatively, we could define the noncentered maximal function with balls instead of cubes:

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

at each $x \in \mathbb{R}^n$. Here, the supremum is taken over all open balls B in $\mathbb{R}^n$ that contain the point x.

Clearly, $Mf \leqslant \widetilde{M}f \leqslant 2^n Mf$ and the boundedness properties of $\widetilde{M}$ are identical to those of M.

Remark 2.3. (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 that are often equivalent.

Next, we state some immediate properties of the maximal function. The proofs are left to interested readers.

Proposition 2.4. *Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then*

(i) *Positivity: $Mf(x) \geqslant 0$ for all $x \in \mathbb{R}^n$.*

(ii) *Sublinearity:* $M(f+g)(x) \leqslant Mf(x) + Mg(x)$.
(iii) *Homogeneity:* $M(\alpha f)(x) = |\alpha| Mf(x)$, $\alpha \in \mathbb{R}$.
(iv) *Translation invariance:* $M(\tau_y f) = (\tau_y Mf)(x) = Mf(x-y)$.

We are now ready to obtain some basic properties of maximal functions. We need the following simple covering lemma.

Lemma 2.5 (Wiener's Vitali-type covering lemma)**.** *Suppose* $\mathcal{B} = \{B_1, \cdots, B_N\}$ *is a finite collection of open balls in* $\mathbb{R}^n$*. Then, there exists a disjoint subcollection* $B_{j_1}, \cdots, B_{j_k}$ *of* $\mathcal{B}$ *such that*

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

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 2.1.)

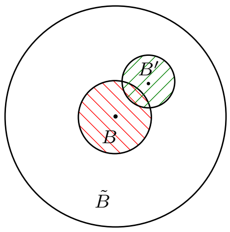


Figure 2.1: The balls B and $\tilde{B}$

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.

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 a radius equal to or smaller than B_{j_i}, we must have $\bigcup_{B \cap B_{j_i} \neq \varnothing} B \subset \tilde{B}_{j_i}$, thus,

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

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. ❑

With Wiener's Vitali-type covering lemma, we can state and prove the main results for the maximal function.

Theorem 2.6 (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 a.e.*

(ii) *If $f \in L^1(\mathbb{R}^n)$, then for every $\alpha > 0$, M is of weak type $(1,1)$, i.e.,*

$$|\{x : Mf(x) > \alpha\}| \leqslant \frac{3^n}{\alpha}\|f\|_1.$$

(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 $A_p = 2(3^n/(p-1))^{1/p}$ for $p \in (1, \infty)$ and $A_\infty = 1$.

Proof. We first prove the second one, i.e., (ii). Since $Mf \leqslant \widetilde{M}f \leqslant 2^n Mf$, we only need to prove it for $\widetilde{M}$. Denote for $\alpha > 0$

$$E_\alpha = \left\{x : \widetilde{M}f(x) > \alpha\right\},$$

we claim that the set E_α is open. Indeed, from the definitions of $\widetilde{M}f$ and the supremum, for each $x \in E_\alpha$ and $0 < \varepsilon < \widetilde{M}f(x) - \alpha$, there exists an open ball $B_x \ni x$ such that

$$\frac{1}{|B_x|}\int_{B_x} |f(y)|dy > \widetilde{M}f(x) - \varepsilon > \alpha.$$

Then for any $z \in B_x$, we have $\widetilde{M}f(z) > \alpha$, and thus, $B_x \subset E_\alpha$. This implies that E_α is open.

Therefore, for any $x \in E_\alpha$ and the above open balls, we have

$$|B_x| < \frac{1}{\alpha}\int_{B_x} |f(y)|dy. \tag{2.6}$$

Fix a compact subset K of E_α. Since K is covered by $\cup_{x\in E_\alpha} B_x$, by the Heine-Borel theorem, we may select a finite subcover of K, say $K \subset \bigcup_{\ell=1}^N B_\ell$. Lemma 2.5 guarantees the existence of a subcollection $B_{j_1}, \cdots, B_{j_k}$ of disjoint balls with

$$\left|\bigcup_{\ell=1}^N B_\ell\right| \leqslant 3^n \sum_{i=1}^k |B_{j_i}|. \tag{2.7}$$

Since the balls $B_{j_1}, \cdots, B_{j_k}$ are disjoint and satisfy (2.6) as well as (2.7), we find that

$$|K| \leqslant \left|\bigcup_{\ell=1}^N B_\ell\right| \leqslant 3^n \sum_{i=1}^k |B_{j_i}| \leqslant \frac{3^n}{\alpha}\sum_{i=1}^k \int_{B_{j_i}} |f(y)|dy$$

$$=\frac{3^n}{\alpha}\int_{\bigcup_{i=1}^k B_{j_i}}|f(y)|dy \leqslant \frac{3^n}{\alpha}\int_{E_\alpha}|f(y)|dy.$$

Since this inequality is true for all compact subsets K of E_α, taking the supremum over all compact $K \subset E_\alpha$ and using the inner regularity of the Lebesgue measure (i.e., Theorem 0.5), we deduce the weak type inequality (ii) for the maximal operator $\widetilde{M}$. It follows from $Mf \leqslant \widetilde{M}f$ that

$$|\{x : Mf(x) > \alpha\}| \leqslant |\{x : \widetilde{M}f(x) > \alpha\}| \leqslant \frac{3^n}{\alpha}\int_{E_\alpha}|f(y)|dy.$$

The above proof also gives the proof of (i) for the case when $p = 1$. For the case $p = \infty$, by Theorem 2.2, (i) and (iii) are true with $A_\infty = 1$.

Now, by using the Marcinkiewicz interpolation theorem (according to the case $p_1 = q_1 = \infty$ and $p_0 = q_0 = 1$ in its proof) between $L^1 \to L^{1,\infty}$ and $L^\infty \to L^\infty$, we can simultaneously obtain (i) and (iii) for the case $p \in (1, \infty)$. ❑

Now, we make some clarifying comments.

Remark 2.7. (1) It is useful for certain applications to observe that

$$A_p = O\left(\left(\frac{1}{p-1}\right)^{1/p}\right), \quad \text{as } p \to 1.$$

(2) It is easier to use $\widetilde{M}$ in proving (ii) than M, and one can see the proof that E_α is open.

§ 2.2 Differentiation theorems

We introduce some notation that will be used frequently hereafter: If φ is any function on $\mathbb{R}^n$ and $\varepsilon > 0$, we set

$$\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon). \tag{2.8}$$

We observe that if $\varphi \in L^1(\mathbb{R}^n)$, then $\int \varphi_\varepsilon$ is independent of ε by a change in variables,

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

Moreover, the "mass" of φ_ε becomes concentrated at the origin as $\varepsilon \to 0$.

Theorem 2.8. *Suppose $\varphi \in L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n}\varphi(x)dx = a$. Let $f \in \mathcal{C}_c(\mathbb{R}^n)$. Then,*

$$\lim_{\varepsilon \to 0^+}\varphi_\varepsilon * f(x) = af(x).$$

Proof. Since $\varphi_\varepsilon * f(x) - af(x) = \int_{\mathbb{R}^n}(f(x-y) - f(x))\varphi_\varepsilon(y)dy$. Since f is continuous at x, for any $\sigma > 0$, there exists a $\delta > 0$ such that

$$|f(x-y) - f(x)| < \frac{\sigma}{\|\varphi\|_1},$$

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

$$\begin{aligned}
|\varphi_\varepsilon * f(x) - af(x)| \leqslant & \frac{\sigma}{\|\varphi\|_1} \int_{|x|<\delta} |\varphi_\varepsilon(x)|dx + 2\|f\|_\infty \int_{|x|\geqslant\delta} |\varphi_\varepsilon(x)|dx \\
\leqslant & \frac{\sigma}{\|\varphi\|_1} \|\varphi\|_1 + 2\|f\|_\infty \int_{|y|\geqslant\delta/\varepsilon} |\varphi(y)|dy \\
= & \sigma + 2\|f\|_\infty I_\varepsilon.
\end{aligned}$$

However, $I_\varepsilon \to 0$ as $\varepsilon \to 0$. This proves the result. ❑

By the density (cf. [Fol99, Propositions 7.9 and 4.35]) of $\mathcal{C}_c$ in L^p ($1 \leqslant p < \infty$) and $\mathcal{C}_0$ (with L^∞ norm), we immediately have the following result.

Theorem 2.9. *Suppose $\varphi \in L^1(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} \varphi(x)dx = a$. 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 $1 \leqslant p \leqslant \infty$*

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

In most applications of the preceding theorem, one has $a = 1$, although the case $a = 0$ is also useful. If $a = 1$, $\{\varphi_\varepsilon\}_{\varepsilon>0}$ is called an ***approximate identity***, as it furnishes an approximation to the identity operator on L^p by convolution operators.

The Hardy-Littlewood maximal function is an important tool that can be used to study the identity operator. At first, what could be easier to understand than the identity? We will illustrate that the identity operator can be interesting by using the Hardy-Littlewood maximal function to prove the Lebesgue differentiation theorem, i.e., the identity operator is a pointwise limit of averages on balls.

The Hardy-Littlewood maximal function Mf is obtained as the supremum of the averages of a function f with respect to the dilates of the kernel $k = V_n^{-1}\chi_{B(0,1)}$ in $\mathbb{R}^n$. Indeed, we have

$$\begin{aligned}
Mf(x) = & \sup_{\varepsilon>0} \frac{1}{V_n\varepsilon^n} \int_{\mathbb{R}^n} |f(x-y)|\chi_{B(0,1)}(y/\varepsilon)dy \\
= & \sup_{\varepsilon>0} (|f| * k_\varepsilon)(x).
\end{aligned}$$

Note that the function $k = V_n^{-1}\chi_{B(0,1)}$ has integral 1, and the convolution with k_ε is an averaging operation where we have used the notation in (2.8). However, it is not hard to see that many radially symmetric averaging processes can be estimated by using M. Before stating the results, given a function φ on $\mathbb{R}^n$, we define the ***least decreasing radial majorant*** of φ by

$$\varphi^*(x) = \sup_{|y|\geqslant|x|} |\varphi(y)|. \tag{2.9}$$

It is obvious that $\varphi \in L^1$ if $\varphi^* \in L^1$.

Theorem 2.10. *If φ has the least decreasing radial majorant $\varphi^* \in L^1$. then for $f \in L^1_{\text{loc}}(\mathbb{R}^n)$,*

$$\sup_{\varepsilon>0} |(f * \varphi_\varepsilon)(x)| \leqslant \|\varphi^*\|_1 Mf(x).$$

Proof. With a slight abuse of notation, let us write $\varphi^*(r) = \varphi^*(x)$, if $|x| = r$; it should cause no confusion since $\varphi^*(x)$ is anyway radial. Now, observe that $\varphi^*(r)$ is decreasing and then $\int_{r/2\leqslant|x|\leqslant r} \varphi^*(x)dx \geqslant \varphi^*(r)\int_{r/2\leqslant|x|\leqslant r} dx = c\varphi^*(r)r^n$. Therefore, the assumption $\varphi^* \in L^1$ proves that $r^n\varphi^*(r) \to 0$ as $r \to 0$ or $r \to \infty$. We need to show that

$$(f * \varphi^*_\varepsilon)(x) \leqslant AMf(x), \tag{2.10}$$

where $f \geqslant 0$, $\varepsilon > 0$ and $A = \int_{\mathbb{R}^n} \varphi^*(x)dx$.

Since (2.10) is clearly translation invariant w.r.t f and also dilation invariant w.r.t. φ^* and the maximal function, it suffices to show that

$$(f * \varphi^*)(0) \leqslant AMf(0). \tag{2.11}$$

In proving (2.11), we may clearly assume that $Mf(0) < \infty$. Let us write $\lambda(r) = \int_{\mathbb{S}^{n-1}} f(rx')d\sigma(x')$, and $\Lambda(r) = \int_{|x|\leqslant r} f(x)dx$, so

$$\Lambda(r) = \int_0^r \int_{\mathbb{S}^{n-1}} f(tx')d\sigma(x')t^{n-1}dt = \int_0^r \lambda(t)t^{n-1}dt, \text{ i.e., } \Lambda'(r) = \lambda(r)r^{n-1}.$$

We have

$$\begin{aligned}
(f * \varphi^*)(0) &= \int_{\mathbb{R}^n} f(x)\varphi^*(x)dx = \int_0^\infty r^{n-1}\int_{\mathbb{S}^{n-1}} f(rx')\varphi^*(r)d\sigma(x')dr \\
&= \int_0^\infty r^{n-1}\lambda(r)\varphi^*(r)dr = \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N \lambda(r)\varphi^*(r)r^{n-1}dr \\
&= \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N \Lambda'(r)\varphi^*(r)dr \\
&= \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \left\{ [\Lambda(r)\varphi^*(r)]_\varepsilon^N - \int_\varepsilon^N \Lambda(r)d\varphi^*(r) \right\}.
\end{aligned}$$

Since $\Lambda(r) \leqslant V_n r^n Mf(0)$, and the fact $r^n\varphi^*(r) \to 0$ as $r \to 0$ or $r \to \infty$, we have

$$0 \leqslant \lim_{N\to\infty} \Lambda(N)\varphi^*(N) \leqslant V_n Mf(0) \lim_{N\to\infty} N^n\varphi^*(N) = 0,$$

which implies $\lim_{N\to\infty} \Lambda(N)\varphi^*(N) = 0$ and similarly $\lim_{\varepsilon\to 0} \Lambda(\varepsilon)\varphi^*(\varepsilon) = 0$. Thus, by integration by parts, we have

$$\begin{aligned}
(f * \varphi^*)(0) &= \int_0^\infty \Lambda(r)d(-\varphi^*(r)) \leqslant V_n Mf(0)\int_0^\infty r^n d(-\varphi^*(r)) \\
&= nV_n Mf(0)\int_0^\infty \varphi^*(r)r^{n-1}dr = Mf(0)\int_{\mathbb{R}^n} \varphi^*(x)dx,
\end{aligned}$$

where two of the integrals are of the Lebesgue-Stieltjes type since $\varphi^*(r)$ is decreasing, which implies $\partial_r \varphi^*(r) \leqslant 0$, and $nV_n = \omega_{n-1}$. This proves (2.11) and then (2.10). ❑

Theorem 2.11. *If φ has the least decreasing radial majorant $\varphi^* \in L^1$, and $f \in L^p$ for some $p \in [1, \infty]$, then for a.e. $x \in \mathbb{R}^n$,*

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

Proof. The proofs for $p = 1$, $1 < p < \infty$ and $p = \infty$ are slightly different from each other.

Let $a = \int \varphi dx$ and

$$\theta(f)(x) = \limsup_{\varepsilon \to 0^+} |\varphi_\varepsilon * f(x) - af(x)|.$$

Our goal is to show that $\theta(f) = 0$ a.e. Observe that by Theorem 2.8, we have for $g \in \mathcal{C}_c$

$$\theta(f) = \theta(f - g).$$

Additionally, according to Theorem 2.10, there is a constant C such that

$$\theta(f - g)(x) \leqslant |a||f(x) - g(x)| + CM(f - g)(x).$$

Thus, if $f \in L^1$ and $\alpha > 0$, we have by Chebyshev's inequality and Theorem 2.6 (ii) that for any $g \in \mathcal{C}_c$

$$\begin{aligned} |\{x : \theta(f)(x) > \alpha\}| \leqslant & |\{x : |a(f - g)(x)| > \alpha/2\}| \\ & + |\{x : CM(f - g)(x) > \alpha/2\}| \\ \leqslant & \frac{C}{\alpha} \|f - g\|_1. \end{aligned}$$

Since $\mathcal{C}_c$ is dense in L^1, we can approximate f in the L^1 norm by functions $g \in \mathcal{C}_c$ and conclude that $|\{x : \theta(f)(x) > \alpha\}| = 0$. Since it holds for each $\alpha > 0$, then we obtain $|\{x : \theta(f)(x) > 0\}| = 0$.

If $f \in L^p$, $1 < p < \infty$, we can argue as above and use the strong type (p, p) estimates of the maximal operator, i.e., Theorem 2.6 (iii) to conclude that for any $g \in \mathcal{C}_c$,

$$|\{x : \theta(f)(x) > \alpha\}| \leqslant \frac{C}{\alpha^p} \|f - g\|_p.$$

Again, $\mathcal{C}_c$ is dense in L^p if $p < \infty$; thus, we can obtain $\theta(f) = 0$ a.e.

Finally, if $p = \infty$, we claim that for each $N \in \mathbb{N}$, the set $\{x : \theta(f)(x) > 0 \text{ and } |x| < N\}$ has measure zero. This implies the theorem. To establish the claim, we write $f = \chi_{B(0,2N)} f + (1 - \chi_{B(0,2N)}) f =: f_1 + f_2$. Since $f_1 \in L^p$ for $p < \infty$, we have $\theta(f_1) = 0$ a.e. It is easy to see that $\theta(f_2)(x) = 0$ if $|x| < 2N$. Since $\theta(f)(x) \leqslant \theta(f_1)(x) + \theta(f_2)(x)$, the claim follows. ❑

The standard Lebesgue differentiation theorem is a special case of the result proved above.

Theorem 2.12 (Lebesgue differentiation theorem). *If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, then*

$$\lim_{r\to 0^+} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)dy = f(x), \quad \text{for a.e. } x.$$

§ 2.3 Calderón-Zygmund decomposition

Applying the Lebesgue differentiation theorem, we give a decomposition of $\mathbb{R}^n$, called the Calderón-Zygmund decomposition, which is extremely useful in harmonic analysis.

Theorem 2.13 (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, such that for each* Q_k

$$\alpha < \frac{1}{|Q_k|} \int_{Q_k} |f(x)|dx \leqslant 2^n \alpha. \tag{2.12}$$

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

$$\frac{1}{|Q_k^{(0)}|} \int_{Q_k^{(0)}} |f(x)|dx \leqslant \alpha, \tag{2.13}$$

since $f \in L^1$.

Split each $Q_k^{(0)}$ into 2^n congruent cubes which we denote by $Q_k^{(1)}$, $k = 1, 2, \cdots$. There are two possibilities:

$$\text{either } \frac{1}{|Q_k^{(1)}|} \int_{Q_k^{(1)}} |f(x)|dx \leqslant \alpha, \text{ or } \frac{1}{|Q_k^{(1)}|} \int_{Q_k^{(1)}} |f(x)|dx > \alpha.$$

In the first case, we split $Q_k^{(1)}$ again into 2^n congruent cubes to obtain $Q_k^{(2)}$ $(k = 1, 2, \cdots)$. In the second case, we have

$$\alpha < \frac{1}{|Q_k^{(1)}|} \int_{Q_k^{(1)}} |f(x)|dx \leqslant \frac{1}{2^{-n}|Q_{\tilde{k}}^{(0)}|} \int_{Q_{\tilde{k}}^{(0)}} |f(x)|dx \leqslant 2^n \alpha$$

in view of (2.13), 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

$$|Q_{k_j}^{(j)}| \to 0 \text{ as } j \to \infty, \quad \text{and } \frac{1}{|Q_{k_j}^{(j)}|} \int_{Q_{k_j}^{(j)}} |f(y)| dy \leqslant \alpha \quad (j = 0, 1, \cdots).$$

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 2.14. *Suppose f, α, F, Ω and Q_k have the same meaning as in Theorem 2.13. Then, there exist two constants A and B (depending only on the dimension n), such that (i) and (ii) of Theorem 2.13 hold and*

(a) $|\Omega| \leqslant \dfrac{A}{\alpha} \|f\|_1$,

(b) $\dfrac{1}{|Q_k|} \displaystyle\int_{Q_k} |f| dx \leqslant B\alpha$.

Proof. From (2.12), it follows that

$$|\Omega| = \sum_k |Q_k| < \frac{1}{\alpha} \int_\Omega |f(x)| dx \leqslant \frac{1}{\alpha} \|f\|_1.$$

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 2.13 from which it was deduced, but by using the maximal function theorem (Theorem 2.6) and the theorem about the decomposition of an arbitrary open set as a union of disjoint cubes as follows. 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***.

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.

Lemma 2.15 (Whitney decomposition). *Let F be a nonempty closed set in $\mathbb{R}^n$ and Ω be its complement. Then, there exists a countable collection of cubes $\mathcal{F} = \{Q_k\}_{k=1}^{\infty}$ whose sides are parallel to the axes, such that*

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

(ii) $\mathring{Q}_j \cap \mathring{Q}_k = \varnothing$ *if $j \neq k$, where $\mathring{Q}$ 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 integers. 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 $\mathscr{M}_k = 2^{-k}\mathscr{M}_0$.

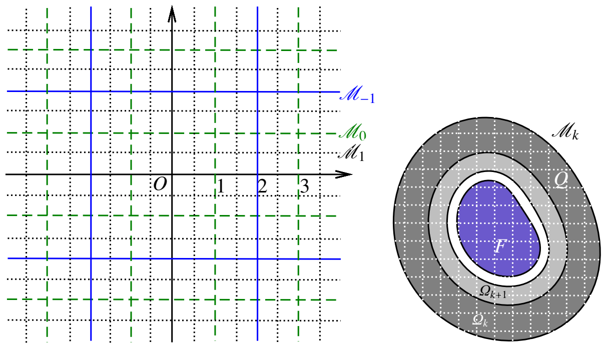


Figure 2.2: Meshes and layers: $\mathscr{M}_0$ with dashed lines; $\mathscr{M}_1$ with dotted lines; $\mathscr{M}_{-1}$ with solid lines

Thus, each cube in mesh $\mathscr{M}_k$ gives rise to 2^n cubes in mesh $\mathscr{M}_{k+1}$ by bisecting the sides. The cubes in 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 that 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 $\mathcal{F}_0$. Our choice is made as follows. We consider the cubes of the mesh $\mathscr{M}_k$ (each cube is of size approximately 2^{-k}) and include a cube of this mesh in $\mathcal{F}_0$ if it intersects Ω_k (the points of the latter are all approximately at a distance of 2^{-k} from F). Namely,

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

For an appropriate choice of c, we claim that

$$\operatorname{diam}(Q) \leqslant \operatorname{dist}(Q, F) \leqslant 4\operatorname{diam}(Q), \quad Q \in \mathcal{F}_0. \tag{2.14}$$

Let us prove (2.14) first. Suppose $Q \in \mathscr{M}_k$; then $\operatorname{diam}(Q) = \sqrt{n}2^{-k}$. Since $Q \in \mathcal{F}_0$, 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 obtain (2.14). Then, by (2.14), the cubes $Q \in \mathcal{F}_0$ are disjoint from F and clearly cover Ω. Therefore, (i) is also proven.

Note that the collection $\mathcal{F}_0$ 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 $\mathcal{F}_0$, eliminating those cubes that were truly unnecessary.

We require the following simple observation. Suppose that Q_1 and Q_2 are two cubes (taken from meshes $\mathscr{M}_{k_1}$ and $\mathscr{M}_{k_2}$, respectively). Then, if $Q_1 \cap Q_2 = \varnothing$, 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 $Q \in \mathcal{F}_0$, and consider the maximal cube in $\mathcal{F}_0$ that contains it. In view of the inequality (2.14), for any cube $Q' \in \mathcal{F}_0$ that contains $Q \in \mathcal{F}_0$, 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'' that contain Q have obviously a nontrivial intersection. Thus, by the observation made above each cube, $Q \in \mathcal{F}_0$ has a unique maximal cube in $\mathcal{F}_0$ that contains it. By the same taken, these maximal cubes are also disjoint. We let $\mathcal{F}$ denote the collection of maximal cubes of $\mathcal{F}_0$. Then, obviously,

(i) $\bigcup_{Q \in \mathcal{F}} Q = \Omega$,
(ii) The cubes of $\mathcal{F}$ are almost disjoint,
(iii) $\operatorname{diam}(Q) \leqslant \operatorname{dist}(Q, F) \leqslant 4 \operatorname{diam}(Q)$, $Q \in \mathcal{F}$.

Therefore, we complete the proof. ❑

Another proof of Corollary 2.14. We know that in F, $|f(x)| \leqslant \alpha$, but this fact does not determine F. The set F is, in effect, determined by the fact that the (uncentered) maximal function satisfies $\widetilde{M}f(x) \leqslant \alpha$ on it. Therefore, we choose $F = \left\{x : \widetilde{M}f(x) \leqslant \alpha\right\}$ and $\Omega = E_\alpha = \left\{x : \widetilde{M}f(x) > \alpha\right\}$. Then, by Theorem 2.6 (ii), we know that $|\Omega| \leqslant \frac{3^n}{\alpha}\|f\|_1$. Thus, we can take $A = 3^n$.

From the proof of Theorem 2.6, we know that Ω is open, and then $F = \Omega^c$ is closed. Then, we can choose cubes Q_k according to Lemma 2.15, 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 $p_k \in F$ 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

$$\gamma_k = \frac{|B_k|}{|Q_k|}.$$

We have, because $p_k \in \left\{x : \widetilde{M}f(x) \leqslant \alpha\right\}$, that

$$\alpha \geqslant \widetilde{M}f(p_k) \geqslant \frac{1}{|B_k|}\int_{B_k} |f(x)|dx \geqslant \frac{1}{\gamma_k |Q_k|}\int_{Q_k} |f(x)|dx.$$

Thus, we can take an upper bound of γ_k as the value of B.

The elementary geometry and inequality (iii) of Lemma 2.15 then show that

$$\operatorname{radius}(B_k) \leqslant \operatorname{dist}(p_k, Q_k) + \operatorname{diam}(Q_k) = \operatorname{dist}(F, Q_k) + \operatorname{diam}(Q_k)$$

$$\leqslant (c_2+1)\,\mathrm{diam}\,(Q_k),$$

and so

$$\begin{aligned}|B_k| =& V_n(\mathrm{radius}(B_k))^n \leqslant V_n(c_2+1)^n(\,\mathrm{diam}\,(Q_k))^n\\ =& V_n(c_2+1)^n n^{n/2}|Q_k|,\end{aligned}$$

since $|Q_k| = (\,\mathrm{diam}\,(Q_k)/\sqrt{n})^n$. 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 2.16. Theorem 2.13 may be used to give another proof of the fundamental inequality for the maximal function in part (ii) of Theorem 2.6. (See [Ste70, §5.1, p.22–23] for more details.)

The Calderón-Zygmund decomposition is a key step in the real 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. 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 a mean value of 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 for functions. The latter is a very important tool in harmonic analysis.

Theorem 2.17 (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) $b = \sum\limits_j b_j$, *where each b_j is supported in a dyadic cube Q_j satisfying $\int_{Q_j} b_j(x)dx = 0$ and $\|b_j\|_1 \leqslant 2^{n+1}\alpha|Q_j|$. Furthermore, cubes Q_j and Q_k have disjoint interiors when $j \neq k$.*

(iii) $\sum\limits_j |Q_j| \leqslant \alpha^{-1}\|f\|_1$.

Proof. Applying Corollary 2.14 (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 Q_j mutually disjoint;

4) $|\Omega| \leqslant \alpha^{-1} \int_{\mathbb{R}^n} |f(x)|dx$, and $\alpha < \frac{1}{|Q_j|} \int_{Q_j} |f(x)|dx \leqslant 2^n \alpha$.

From 3) and 4), it is easy to obtain (iii).

Now define

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

$b = \sum_j b_j$ and $g = f - b$. It is clear that $\int_{Q_j} b_j(x)dx = 0$. Consequently,

$$\begin{aligned}\int_{Q_j} |b_j|dx \leqslant & \int_{Q_j} |f(x)|dx + |Q_j| \left| \frac{1}{|Q_j|} \int_{Q_j} f(x)dx \right| \\ \leqslant & 2 \int_{Q_j} |f(x)|dx \leqslant 2^{n+1}\alpha|Q_j|,\end{aligned}$$

which proves that $\|b_j\|_1 \leqslant 2^{n+1}\alpha|Q_j|$. Thus, (ii) is proved with the help of 3).

Next, we need to obtain the estimates on g. Write $\mathbb{R}^n = \bigcup_j Q_j \cup F$, where F is the set obtained by Corollary 2.14. Since $b = 0$ on F and $f - b_j = \frac{1}{|Q_j|} \int_{Q_j} f(x)dx$ on Q_j, we have

$$g = \begin{cases} f, & \text{on } F, \\ \dfrac{1}{|Q_j|} \displaystyle\int_{Q_j} f(x)dx, & \text{on } Q_j. \end{cases} \tag{2.15}$$

On the cube Q_j, g is equal to the constant $\frac{1}{|Q_j|} \int_{Q_j} f(x)dx$, and this is bounded by $2^n\alpha$ by 4). Then, by 2), we can obtain $\|g\|_\infty \leqslant 2^n\alpha$. Finally, it follows from (2.15) that $\|g\|_1 \leqslant \|f\|_1$. This completes the proof of (i) and then of the theorem. ❑

As an application of the Marcinkiewicz interpolation theorem and the Calderón-Zygmund decomposition, we now prove the weighted estimates for the Hardy-Littlewood maximal function (cf. [FS71, p.111, Lemma 1]).

Theorem 2.18 (Weighted inequality for Hardy-Littlewood maximal function)**.** *For $p \in (1, \infty)$, there exists a constant $C = C_{n,p}$ such that, for any nonnegative real-valued locally integrable function $\varphi(x)$ on $\mathbb{R}^n$, we have, for $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, the inequality*

$$\int_{\mathbb{R}^n} (Mf(x))^p \varphi(x)dx \leqslant C \int_{\mathbb{R}^n} |f(x)|^p M\varphi(x)dx. \tag{2.16}$$

We first prove the following lemma.

Lemma 2.19. *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 $\mathbb{R}^n$ (i.e., Theorem 2.13)* *for f and $\alpha > 0$, then*

$$\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha\} \subset \bigcup_k Q_k^*,$$

where $Q_k^* = 2Q_k$. It follows

$$|\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha\}| \leqslant 2^n \sum_k |Q_k|.$$

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

$$\frac{1}{|Q|}\int_Q |f(x)|dx \leqslant \alpha.$$

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

$$\begin{aligned}
\int_Q |f(x)|dx \leqslant & \int_{Q\cap F} |f(x)|dx + \sum_{Q_k \cap Q \neq \varnothing} \int_{Q_k} |f(x)|dx \\
\leqslant & \alpha|Q| + \sum_{Q_k\cap Q\neq\varnothing} 2^n\alpha|Q_k| \\
\leqslant & \alpha|Q| + 2^n\alpha|3Q| \\
\leqslant & 7^n\alpha|Q|.
\end{aligned}$$

Thus, we know that $M'f(x) \leqslant 7^n\alpha$ for any $x \notin \bigcup_k Q_k^*$, which yields that

$$|\{x \in \mathbb{R}^n : M'f(x) > 7^n\alpha\}| \leqslant \left|\bigcup_k Q_k^*\right| \leqslant 2^n \sum_k |Q_k|.$$

We complete the proof of the lemma. ❑

Proof of Theorem 2.18. Except when $M\varphi(x) = \infty$ a.e., in which case (2.16) holds trivially, $M\varphi$ is the density of a positive measure σ. Thus, we may assume that $M\varphi(x) < \infty$ a.e. $x \in \mathbb{R}^n$ and $M\varphi(x) > 0$. If we denote

$$d\sigma(x) = M\varphi(x)dx \quad \text{and} \quad d\nu(x) = \varphi(x)dx.$$

Then, by the Marcinkiewicz interpolation theorem, to obtain (2.16), it suffices to prove that M is both of weak type $(L^\infty(\sigma), L^\infty(\nu))$ and of weak type $(L^1(\sigma), L^1(\nu))$.

Let us first show that M is of weak type $(L^\infty(\sigma), L^\infty(\nu))$. If $\|f\|_{L^\infty(\sigma)} = \alpha$, then

$$\int_{\{x\in\mathbb{R}^n:|f(x)|>\alpha\}} M\varphi(x)dx = \sigma(\{x \in \mathbb{R}^n : |f(x)| > \alpha\}) = 0.$$

Since $M\varphi(x) > 0$ for any $x \in \mathbb{R}^n$, we have $|\{x \in \mathbb{R}^n : |f(x)| > \alpha\}| = 0$, 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 then $|\{x : Mf(x) > \alpha\}| = 0$ which implies that $\nu(\{Mf(x) > \alpha\}) = \int_{\{x:Mf(x)>\alpha\}} \varphi(x)dx = 0$ and thus $\|Mf\|_{L^\infty(\nu)} \leqslant \alpha$. Therefore, $\|Mf\|_{L^\infty(\nu)} \leqslant \|f\|_{L^\infty(\sigma)}$.

Let us turn to the proof of the weak type $(L^1(\sigma), L^1(\nu))$. We need to prove that there exists a constant C such that for any $\alpha > 0$ and $f \in L^1(\sigma)$

$$\begin{aligned}\int_{\{x\in\mathbb{R}^n:Mf(x)>\alpha\}} \varphi(x)dx =&\nu(\{x \in \mathbb{R}^n : Mf(x) > \alpha\}) \\ \leqslant&\frac{C}{\alpha}\int_{\mathbb{R}^n} |f(x)|M\varphi(x)dx.\end{aligned} \tag{2.17}$$

We may assume that $f \in L^1(\mathbb{R}^n)$. In fact, if we take $f_\ell = |f|\chi_{B(0,\ell)}$, then $f_\ell \in L^1(\mathbb{R}^n)$, $0 \leqslant f_\ell(x) \leqslant f_{\ell+1}(x)$ for $x \in \mathbb{R}^n$ and $\ell = 1, 2, \cdots$. By $\lim\limits_{\ell\to\infty} f_\ell(x) = |f(x)|$ and Exercise 2.9, we have

$$\{x \in \mathbb{R}^n : Mf(x) > \alpha\} = \bigcup_\ell \{x \in \mathbb{R}^n : Mf_\ell(x) > \alpha\}.$$

Due to $Mf(x) \leqslant c_n M'f(x)$ with $c_n = 2^n/V_n$ 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 obtain a sequence $\{Q_k\}$ of cubes satisfying

$$\alpha' < \frac{1}{|Q_k|}\int_{Q_k} |f(x)|dx \leqslant 2^n\alpha'.$$

By Lemma 2.19 and $M''\varphi \leqslant V_n n^{n/2} M\varphi$, we have

$$\begin{aligned}&\int_{\{x\in\mathbb{R}^n:Mf(x)>\alpha\}} \varphi(x)dx \\ \leqslant&\int_{\{x\in\mathbb{R}^n:M'f(x)>7^n\alpha'\}} \varphi(x)dx \\ \leqslant&\int_{\bigcup_k Q_k^*} \varphi(x)dx \leqslant \sum_k \int_{Q_k^*} \varphi(x)dx \\ \leqslant&\sum_k \left(\frac{1}{|Q_k|}\int_{Q_k^*} \varphi(x)dx\right)\left(\frac{1}{\alpha'}\int_{Q_k} |f(y)|dy\right) \\ =&\frac{c_n 7^n}{\alpha}\sum_k \int_{Q_k} |f(y)|\left(\frac{2^n}{|Q_k^*|}\int_{Q_k^*} \varphi(x)dx\right)dy \\ \leqslant&\frac{c_n 14^n}{\alpha}\sum_k \int_{Q_k} |f(y)|M''\varphi(y)dy \\ \leqslant&\frac{28^n n^{n/2}}{\alpha}\int_{\mathbb{R}^n} |f(y)|M\varphi(y)dy.\end{aligned}$$

Thus, M is of weak type $(L^1(\sigma), L^1(\nu))$, and the inequality can be obtained by applying the Marcinkiewicz interpolation theorem. ❑

Exercises

Exercise 2.1. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \chi_{(0,1)}(x)$. To calculate Mf, $M'f$, $M''f$ and $\widetilde{M}f$.

Exercise 2.2. [Gra14b, Exercise 2.1.13] Observe that the proof of Theorem 2.6 yields the estimate

$$\alpha|\{\widetilde{M}f > \alpha\}|^{1/p} \leqslant 3^n|\{\widetilde{M}f > \alpha\}|^{-1+1/p} \int_{\{\widetilde{M}f>\alpha\}} |f(y)|dy$$

for $\alpha > 0$ and f locally integrable. Use the result of Exercise 1.3 (i) to prove that the Hardy-Littlewood maximal operator M maps the space $L^{p,\infty}(\mathbb{R}^n)$ to itself for $1 < p < \infty$.

Exercise 2.3. [Zho99, Exercise 10, on p.75] Assume that $f(x) \geqslant 0$ is a locally integrable function. To show that for any $x > 0$, we have

$$x \int_x^\infty \frac{f(t)}{t^2} dt \leqslant CMf(\xi), \quad \xi \in (0, x].$$

Exercise 2.4. [Gra14a, Exercise 2.1.8] Prove that for any fixed $1 < p < \infty$, the operator norm of $\widetilde{M}$ on $L^p(\mathbb{R}^n)$ tends to infinity as $n \to \infty$.
Hint Let f_0 be the characteristic function of the unit ball in $\mathbb{R}^n$. Consider the averages $|B_x|^{-1} \int_{B_x} f_0 dy$, where $B_x = B(\frac{1}{2}(|x| - |x|^{-1})\frac{x}{|x|}, \frac{1}{2}(|x| + |x|^{-1}))$ for $|x| > 1$.

Exercise 2.5. [Pey18, Exercise 1.3] Let $f = \chi_{B(0,1)}$ be the characteristic function of the unit ball in $\mathbb{R}^n$. Show that, for $|x| > 1$, $Mf(x) \leqslant C/(|x| - 1)^n$, where $C > 0$ is a constant. Conclude that, for $p > 1$, $Mf \in L^p(\mathbb{R}^n)$.

Exercise 2.6. [Pey18, Exercise 1.4] If $f \in L^1(\mathbb{R}^n)$ and $f \neq 0$, then $Mf \notin L^1(\mathbb{R}^n)$.
Hint Prove that $Mf(x) \geqslant C/|x|^n$ for $|x|$ large enough, where $C > 0$ is a constant.

Exercise 2.7. [Pey18, Exercise 1.12] 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

$$\int_E Mf(x)dx \leqslant 2|E| + C \int_E |f(x)| \ln^+ |f(x)|dx,$$

where $\ln^+ t = \max(\ln t, 0)$.

Exercise 2.8 ((Gagliardo-Nirenberg-) Sobolev inequality)**.** Let $p \in (1, n)$ and its Sobolev conjugate $p^* = np/(n - p)$. Use the maximal function theorem to prove that 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.

Exercise 2.9. [Pey18, Exercise 1.6] Let $f_1, f_2, \cdots, f_m, \cdots$ be a nondecreasing sequence of nonnegative functions in $L^1(\mathbb{R}^n)$. Let f be the pointwise limit of f_m. Show that, for all $x \in \mathbb{R}^n$,

$$Mf(x) = \lim_{m\to\infty} Mf_m(x).$$

3

Fourier Transform and Tempered Distributions

In this chapter, we introduce the Fourier transform and its elementary properties, approximate identities, the Schwartz space and its dual space. We also give some characterizations of operators commuting with translations and Fourier multipliers as a special class.

§ 3.1 Fourier transform

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

Definition 3.1. If $f \in L^1(\mathbb{R}^n)$, then its ***Fourier transform*** is $\mathscr{F}f$ or $\widehat{f} : \mathbb{R}^n \to \mathbb{C}$ defined by

$$\mathscr{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) đx \tag{3.1}$$

for all $\xi \in \mathbb{R}^n$, where we denote $đx = (2\pi)^{-n/2}dx$ for $x \in \mathbb{R}^n$.

We now continue with some properties of the Fourier transform. Before doing this, we shall introduce some notations. We recall that the space $\mathcal{C}_0(\mathbb{R}^n)$ consists of all continuous functions vanishing at infinity. For a measurable function f on $\mathbb{R}^n$, $x \in \mathbb{R}^n$ and $a \neq 0$, we define the ***translation***, ***dilation*** and ***reflection*** of f by

$$\begin{aligned} \tau^y f(x) =& f(x-y), \\ \delta^a f(x) =& f(ax), \\ \widetilde{f}(x) =& f(-x). \end{aligned}$$

Proposition 3.2. *Given $f, g \in L^1(\mathbb{R}^n)$, $x, y, \xi \in \mathbb{R}^n$, α multi-index, $a, b \in \mathbb{C}$, $\varepsilon \in \mathbb{R}$ and $\varepsilon \neq 0$, we have*

(i) *(Linearity)* $\widehat{(af + bg)} = a\widehat{f} + b\widehat{g}$.

(ii) *(Translation)* $\widehat{\tau^y f}(\xi) = e^{-iy\cdot\xi}\widehat{f}(\xi)$.

(iii) *(Modulation)* $\widehat{(e^{ix\cdot y} f(x))}(\xi) = \tau^y \widehat{f}(\xi)$.

(iv) *(Conjugation)* $\widehat{\overline{f}} = \overline{\widetilde{\widehat{f}}}$.

(v) *(Transformation) If T is an invertible linear transformation of $\mathbb{R}^n$ and*

$S = (T^*)^{-1}$ is its inverse transpose, then $\widehat{f \circ T} = |\det T|^{-1}\widehat{f} \circ S$. In particular, if T is a rotation, then $\widehat{f \circ T} = \widehat{f} \circ T$.

(vi) *(Scaling)* $\widehat{\delta^\varepsilon f}(\xi) = |\varepsilon|^{-n}\delta^{\varepsilon^{-1}}\widehat{f}(\xi)$.

(vii) *(Convolution)* $(2\pi)^{-n/2}\widehat{f * g} = \widehat{f}\widehat{g}$.

(viii) *If $x^\alpha f \in L^1$ for $|\alpha| \leqslant k$, then $\widehat{f} \in \mathcal{C}^k$, and $\partial^\alpha \widehat{f} = \widehat{((-ix)^\alpha f(x))}$.*

(ix) *If $f \in \mathcal{C}^k$, $\partial^\alpha f \in L^1$ for $|\alpha| \leqslant k$, and $\partial^\alpha f \in \mathcal{C}_0$ for $|\alpha| \leqslant k-1$, then $\widehat{\partial^\alpha f}(\xi) = (i\xi)^\alpha \widehat{f}(\xi)$.*

(x) *(Uniform continuity) If $f \in L^1(\mathbb{R}^n)$, then $\widehat{f}$ is uniformly continuous.*

(xi) *(Riemann-Lebesgue lemma) $\mathscr{F}(L^1(\mathbb{R}^n)) \subset \mathcal{C}_0(\mathbb{R}^n)$.*

Proof. (v)

$$\begin{aligned}\widehat{(f \circ T)}(\xi) =& \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(Tx) \, đx \\ =& |\det T|^{-1} \int_{\mathbb{R}^n} e^{-iT^{-1}y\cdot\xi} f(y) \, đy \\ =& |\det T|^{-1} \int_{\mathbb{R}^n} e^{-iy\cdot S\xi} f(y) \, đy \\ =& |\det T|^{-1}\widehat{f}(S\xi).\end{aligned}$$

(x) By

$$\widehat{f}(\xi + h) - \widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}[e^{-ix\cdot h} - 1] f(x) \, đx,$$

we have

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

since $|e^{i\theta} - 1| \leqslant |\theta|$ for any $\theta \geqslant 0$. Given any $\varepsilon > 0$, due to $f \in L^1(\mathbb{R}^n)$, 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 a given $\varepsilon > 0$, there exists a sufficiently small $\delta > 0$ such that $|\widehat{f}(\xi + h) - \widehat{f}(\xi)| < \varepsilon$ when $|h| \leqslant \delta$, where ε is independent of ξ.

(xi) By (ix), if $f \in \mathcal{C}^1 \cap \mathcal{C}_c$, then $|\xi|\widehat{f}(\xi)$ is bounded and hence $\widehat{f} \in \mathcal{C}_0$. However, the set of all such f's is dense in L^1, and $\widehat{f_n} \to \widehat{f}$ uniformly whenever $f_n \to f$ in L^1 by (x). Since $\mathcal{C}_0$ is closed in the uniform norm, the result follows.

The other results are easy to verify. ❑

The Riemann-Lebesgue lemma 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 Exercise 3.5.

We shall need to compute an important specific Fourier transform.

Theorem 3.3. *For all $a > 0$, we have*

$$\widehat{e^{-a|x|^2}}(\xi) = (2a)^{-n/2}e^{-\frac{|\xi|^2}{4a}}. \tag{3.2}$$

In particular,

$$\widehat{e^{-\frac{|x|^2}{2}}}(\xi) = e^{-\frac{|\xi|^2}{2}}.$$

Proof. The integral in question is

$$\int_{\mathbb{R}^n} e^{-ix\cdot\xi}e^{-a|x|^2}\,đx.$$

Note that these factors are a product of one variable integrals. Thus, it is sufficient to prove the case $n = 1$. It is clear that

$$\int_{-\infty}^{\infty} e^{-ix\xi}e^{-ax^2}dx = e^{-\frac{\xi^2}{4a}}\int_{-\infty}^{\infty} e^{-a(x+i\xi/(2a))^2}dx.$$

We observe that the function

$$F(\xi) = \int_{-\infty}^{\infty} e^{-a(x+i\xi/(2a))^2}dx, \quad \xi \in \mathbb{R},$$

defined on the line is constant (and thus equal to $\int_{-\infty}^{\infty} e^{-ax^2}dx$), since its derivative is

$$\begin{aligned}\frac{d}{d\xi}F(\xi) &= -i\int_{-\infty}^{\infty}(x+i\xi/(2a))e^{-a(x+i\xi/(2a))^2}dx\\ &= \frac{i}{2a}\int_{-\infty}^{\infty}\frac{d}{dx}e^{-a(x+i\xi/(2a))^2}dx = 0.\end{aligned}$$

It follows that $F(\xi) = F(0)$ and

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

where we used the formula for the integral of a Gaussian, i.e., the Euler-Poisson integral: $\int_{\mathbb{R}} e^{-\pi x^2}dx = 1$ at the next to last one. ❑

We are ready to invert the Fourier transform. If $f \in L^1$, then we define

$$f^{\vee}(x) = \hat{f}(-x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi}f(\xi)\,đ\xi,$$

and we claim that if $f \in L^1$ and $\widehat{f} \in L^1$ then $(\widehat{f})^\vee = f$. A simple appeal to Fubini's theorem fails because the integrand in

$$\left(\widehat{f}\right)^\vee(x) = \iint e^{ix\cdot\xi} e^{-iy\cdot\xi} f(y) đy đ\xi$$

need not be in $L^1(\mathbb{R}^n \times \mathbb{R}^n)$. The trick is to introduce a convergence factor and then pass it to the limit using Fubini's theorem via the following theorem.

Theorem 3.4 (The multiplication formula). *If $f, g \in L^1(\mathbb{R}^n)$, then*

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

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

Theorem 3.5 (Fourier inversion theorem). *If $f \in L^1$ and $\widehat{f} \in L^1$, then f agrees a.e. with a continuous function f_0, and $\left(\widehat{f}\right)^\vee = \widehat{f^\vee} = f_0$.*

Proof. Given $\varepsilon > 0$ and $x \in \mathbb{R}^n$, let

$$\varphi(\xi) = (2\pi)^{-n/2} \exp(ix\cdot\xi - \frac{\varepsilon^2}{4\pi}|\xi|^2).$$

By (iii) in Proposition 3.2 and Theorem 3.3,

$$\widehat{\varphi}(y) = (2\pi)^{-n/2} \tau^x \widehat{e^{-\frac{1}{4\pi}\varepsilon^2|\xi|^2}}(y) = \varepsilon^{-n} \exp(-\pi|x-y|^2/\varepsilon^2) = g_\varepsilon(x-y),$$

where $g(x) = e^{-\pi|x|^2}$ and the subscript ε has the meaning in (2.8). By Theorem 3.4,

$$\int e^{-\frac{\varepsilon^2}{4\pi}|\xi|^2} e^{ix\cdot\xi} \widehat{f}(\xi) đ\xi = \int \widehat{f} \varphi d\xi = \int f \widehat{\varphi} d\xi = f * g_\varepsilon(x).$$

Since $\int e^{-\pi|x|^2} dx = 1$, by Theorem 2.11, we obtain $f * g_\varepsilon \to f$ a.e. as $\varepsilon \to 0$. However, since $\widehat{f} \in L^1$, and the dominated convergence theorem yields

$$\lim_{\varepsilon\to 0} \int e^{-\frac{\varepsilon^2}{4\pi}|\xi|^2} e^{ix\cdot\xi} \widehat{f}(\xi) đ\xi = \int e^{ix\cdot\xi} \widehat{f}(\xi) đ\xi = \left(\widehat{f}\right)^\vee(x).$$

It follows that $f = \left(\widehat{f}\right)^\vee$ a.e., and similarly $\widehat{f^\vee} = f$ a.e. Since $\left(\widehat{f}\right)^\vee$ and $\widehat{f^\vee}$ are continuous, being Fourier transforms of L^1 functions, the proof is complete. ❑

Corollary 3.6 (Uniqueness). *If $f \in L^1$ and $\widehat{f} = 0$, then $f = 0$ a.e.*

§ 3.2 Schwartz space

We recall the space $\mathscr{D}(\mathbb{R}^n) \equiv \mathcal{C}_c^\infty(\mathbb{R}^n)$ of all smooth functions with compact support, and $\mathcal{C}^\infty(\mathbb{R}^n)$ of all smooth functions on $\mathbb{R}^n$. However, it is not immediately clear that $\mathscr{D}$ is nonempty.

Example 3.7. 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, $f \in \mathcal{C}^\infty$ 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 $|t| < 1$, and zero otherwise. It clearly belongs to $\mathscr{D}(\mathbb{R})$. We can easily obtain n-dimensional variants from φ. For example,

(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 $\eta \in \mathcal{C}^\infty$ 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 $\varepsilon \to 0$.

The other space of $\mathcal{C}^\infty$ functions we shall need is the Schwartz space as follows.

Definition 3.8. The ***Schwartz space*** $\mathscr{S}(\mathbb{R}^n)$ is defined as

$$\mathscr{S}(\mathbb{R}^n) = \left\{\varphi \in \mathcal{C}^\infty(\mathbb{R}^n) : |\varphi|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta \varphi(x)| < \infty,\ \forall \alpha, \beta \in \mathbb{N}_0^n\right\}. \tag{3.3}$$

If $\varphi \in \mathscr{S}$, then $|\varphi(x)| \leqslant C_m(1+|x|)^{-m}$ for any $m \in \mathbb{N}_0$. However, the example $\varphi(x) = e^{-\varepsilon|x|}$ fails to be differential at the origin and, therefore, does not belong to $\mathscr{S}$. Thus, the converse is not true.

Obviously, $\mathscr{D} \subset \mathscr{S}$. The inclusion is strict since $\varphi(x) = e^{-\varepsilon|x|^2} \in \mathscr{S}(\mathbb{R}^n) \setminus \mathscr{D}(\mathbb{R}^n)$ for $\varepsilon > 0$.

Remark 3.9. We observe that the order of multiplication by powers of $x_1, \cdots, x_n$ and differentiation, in (3.3), could have been reversed. That is, for $\varphi \in \mathcal{C}^\infty$,

$$\varphi \in \mathscr{S}(\mathbb{R}^n) \iff \sup_{x\in\mathbb{R}^n} |\partial^\beta(x^\alpha\varphi(x))| < \infty, \forall \alpha, \beta \in \mathbb{N}_0^n.$$

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 $P(x)$).

The following alternative characterization of Schwartz functions is very useful.

Proposition 3.10. *If $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$, then $\varphi \in \mathscr{S}(\mathbb{R}^n)$ iff*

$$\sup_{x\in\mathbb{R}^n} [(1+|x|)^N |\partial^\alpha \varphi(x)|] < \infty, \forall N \in \mathbb{N}_0, \forall \alpha \in \mathbb{N}_0^n.$$

Proof. It is clear that $|x^\gamma| \leqslant (1+|x|)^N$ for $|\gamma| \leqslant N$. However, $\sum_{j=1}^{n} |x_j|^N$ is strictly positive on the unit sphere $|x| = 1$, so it has a positive minimum δ there. It follows that $\sum_{j=1}^{n} |x_j|^N \geqslant \delta |x|^N$ for all x since both sides are homogeneous of degree N, and hence,

$$\begin{aligned}(1+|x|)^N \leqslant & 2^N(1+|x|^N) \leqslant 2^N \left[1+\delta^{-1}\sum_{j=1}^{n}|x_j^N|\right] \\ \leqslant & 2^N\delta^{-1}\sum_{|\gamma|\leqslant N}|x^\beta|.\end{aligned}$$

❑

It is an important observation that if $f \in \mathscr{S}$, then $\partial^\alpha f \in L^p$ for all α and all $p \in [1,\infty]$. Indeed, $|\partial^\alpha f(x)| \leqslant C_N(1+|x|)^{-N}$ for all N, and $(1+|x|)^{-N} \in L^p$ for $N > n/p$. Moreover, we also have the following statement:

Proposition 3.11. *Let $\mathcal{C}^\infty_{poly}(\mathbb{R}^n)$ be the set of all smooth polynomially bounded functions, i.e., the set of all smooth $f : \mathbb{R}^n \to \mathbb{C}$ such that for all $\alpha \in \mathbb{N}_0^n$ there exist $m_\alpha \in \mathbb{N}_0$ and $C_\alpha > 0$ with*

$$|\partial_x^\alpha f(x)| \leqslant C_\alpha (1+|x|)^{m_\alpha} \quad \text{for all } x \in \mathbb{R}^n.$$

Then for every $f \in \mathcal{C}^\infty_{poly}(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$, we have $fg \in \mathscr{S}(\mathbb{R}^n)$.

Proof. This easily follows from the product rule, the Leibniz formula. ❑

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

$$|\varphi|_{\alpha,\beta} = 0 \text{ iff } \varphi = 0$$

fails to hold, for example, for constant function φ.

Proposition 3.12. *$\mathscr{S}$ is a Fréchet space with the topology defined by the seminorms $|\cdot|_{\alpha,\beta}$.*

We leave the proof as an exercise.

Moreover, some easily established properties of $\mathscr{S}(\mathbb{R}^n)$ and its topology are as follows:

Proposition 3.13. (i) *The mapping* $\varphi(x) \mapsto x^\alpha \partial^\beta \varphi(x)$ *is continuous.*
(ii) *If* $\varphi \in \mathscr{S}(\mathbb{R}^n)$*, then* $\lim_{h\to 0} \tau^h \varphi = \varphi$.
(iii) *Suppose* $\varphi \in \mathscr{S}(\mathbb{R}^n)$ *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* $|h| \to 0$.

From part (viii) in Proposition 3.2, we immediately have

Theorem 3.14. $\mathscr{F}$ *maps the Schwartz class* $\mathscr{S}$ *continuously into itself.*

Corollary 3.15. $\mathscr{F}$ *is an isomorphism of* $\mathscr{S}$ *onto itself.*

Proof. By Theorem 3.14, $\mathscr{F}$ maps $\mathscr{S}$ continuously into itself, and hence, so does $f \mapsto f^\vee$ since $f^\vee(x) = \widehat{f}(-x)$. By the Fourier inversion theorem, these maps are inverse to each other. ❑

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 $\widehat{f}$ will also be square-integrable. In fact, we have the following basic result:

Theorem 3.16 (Plancherel theorem). *If* $f \in L^1 \cap L^2$*, then* $\widehat{f} \in L^2$ *and* $\|\widehat{f}\|_2 = \|f\|_2$*; and* $\mathscr{F}|_{(L^1 \cap L^2)}$ *extends uniquely to a unitary isomorphism on* L^2.

Proof. Let $X = \{f \in L^1 : \widehat{f} \in L^1\}$. For any $f \in X$, by Theorem 3.5, we have $\|f\|_\infty = \|\left(\widehat{f}\right)^\vee\|_\infty \leqslant (2\pi)^{-n/2}\|\widehat{f}\|_1$; thus, $f \in L^1 \cap L^\infty \subset L^2$ by Proposition 0.19. Hence, $X \subset L^2$, and X is dense in L^2 because $\mathscr{S} \subset X$ and $\mathscr{S}$ is dense in L^2. Given $f, g \in X$, let $h = \overline{\widehat{g}}$. By the inversion theorem,

$$\widehat{h}(\xi) = \int e^{-ix\cdot\xi}\overline{\widehat{g}(x)}\,đx = \int \overline{e^{ix\cdot\xi}\widehat{g}(x)}\,đx = \overline{g(\xi)}.$$

Hence, by Theorem 3.4,

$$\int f\overline{g} = \int f\widehat{h} = \int \widehat{f}h = \int \widehat{f}\overline{\widehat{g}}.$$

Thus, $\mathscr{F}|_X$ preserves the L^2 inner product; in particular, by taking $g = f$, we obtain $\|\widehat{f}\|_2 = \|f\|_2$. Since $\mathscr{F}(X) = X$ by the inversion theorem, and $\mathscr{F}|_X$ extends by continuity to a unitary isomorphism on L^2.

It remains only to show that this extension agrees with $\mathscr{F}$ on all of $L^1 \cap L^2$. However, if $f \in L^1 \cap L^2$ and $g(x) = e^{-\pi|x|^2}$ as in the proof of the inversion theorem, we have $f * g_\varepsilon \in L^1$ by Young's inequality and $\widehat{f * g_\varepsilon} \in L^1$ because $\widehat{f * g_\varepsilon}(\xi) = e^{-\frac{\varepsilon^2}{4\pi}|\xi|^2}\widehat{f}(\xi)$ and $\widehat{f}$ is bounded. Hence, $f * g_\varepsilon \in X$. Moreover, by Theorem 2.9, $f * g_\varepsilon \to f$ in both the L^1 and L^2 norms. Therefore, $\widehat{f * g_\varepsilon} \to \widehat{f}$ both uniformly and in the L^2 norm, and we are done. ❑

We have thus extended the domain of the Fourier transform from L^1 to $L^1 + L^2$. The Riesz-Thorin interpolation theorem yields the following result for the intermediate L^p spaces:

Theorem 3.17 (Hausdorff-Young inequality). *Let $1 \leqslant p \leqslant 2$ and $1/p + 1/p' = 1$. If $f \in L^p(\mathbb{R}^n)$, then $\widehat{f} \in L^{p'}(\mathbb{R}^n)$ and*

$$\|\widehat{f}\|_{p'} \leqslant (2\pi)^{n(1/p-1/2)}\|f\|_p. \tag{3.4}$$

Proof. It follows from using the Riesz-Thorin interpolation theorem between the $L^1 \to L^\infty$ result $\|\mathscr{F}f\|_\infty \leqslant (2\pi)^{-n/2}\|f\|_1$ (cf. part (x) in Proposition 3.2) and the $L^2 \to L^2$ result, i.e., Plancherel's theorem $\|\mathscr{F}f\|_2 = \|f\|_2$ (cf. Theorem 3.16). ❑

Remark 3.18. (i) Unless $p = 1$ or 2, the constant in the Hausdorff-Young inequality is not the best possible; indeed the best constant is found by testing Gaussian functions. This is much deeper and is due to Babenko [Bab61] when p' is an even integer and to Beckner [Bec75b; Bec75a] in general.

(ii) p' cannot be replaced by some q in (3.4). Namely, if it holds

$$\|\widehat{f}\|_q \leqslant C\|f\|_p, \quad \forall f \in L^p(\mathbb{R}^n), \tag{3.5}$$

then we must have $q = p'$. In fact, we can use the dilation to show it. For $\lambda > 0$, let $f_\lambda(x) = \lambda^{-n}f(x/\lambda)$; then,

$$\begin{aligned}\|f_\lambda\|_p =& \lambda^{-n}\left(\int_{\mathbb{R}^n}|f(x/\lambda)|^p dx\right)^{1/p}\\ =&\lambda^{-n}\left(\int_{\mathbb{R}^n}\lambda^n|f(y)|^p dy\right)^{1/p} = \lambda^{-\frac{n}{p'}}\|f\|_p.\end{aligned}$$

By the property of the Fourier transform, we have $\widehat{f_\lambda} = \lambda^{-n}\widehat{\delta^{\lambda^{-1}}f} = \delta^\lambda\widehat{f}$ and

$$\|\widehat{f_\lambda}\|_q = \left(\int_{\mathbb{R}^n}|\widehat{f}(\lambda\xi)|^q d\xi\right)^{1/q} = \lambda^{-\frac{n}{q}}\|\widehat{f}\|_q.$$

Thus, (3.5) implies $\lambda^{-\frac{n}{q}}\|\widehat{f}\|_q \leqslant C\lambda^{-\frac{n}{p'}}\|f\|_p$, i.e., $\|\widehat{f}\|_q \leqslant C\lambda^{\frac{n}{q}-\frac{n}{p'}}\|f\|_p$, then $q = p'$ by taking λ tending to 0 or ∞.

(iii) Except in the case $p = 2$, inequality (3.4) is not reversible, in the sense that there is no constant C such that $\|\widehat{f}\|_{p'} \geqslant C\|f\|_p$ for $1 < p < 2$ when $f \in \mathscr{D} \equiv \mathcal{C}_c^\infty$. Equivalently, the result cannot be extended to the case $p > 2$ in view of the dual argument and the multiplication formula (Theorem 3.4). To show this, we take $f_\lambda(x) = \phi(x)e^{-\pi(1+i\lambda)|x|^2}$, where $\phi \in \mathscr{D}$ is fixed and λ is a large positive number. Then, $\|f_\lambda\|_p$ is independent of λ for any p. By the Plancherel theorem, $\|\widehat{f_\lambda}\|_2$ is also independent of λ. On the other hand, $\widehat{f_\lambda}$ is the convolution of $\widehat{\phi}$, which is in L^1, with $(2\pi)^{-n}(1+i\lambda)^{-n/2}e^{-[4\pi(1+i\lambda)]^{-1}|x|^2}$ (cf. [Gra14a, Ex.2.3.13, p.133] or [BCD11, Proposition 1.28]),[a] which has L^∞ norm $(2\pi)^{-n}(1+\lambda^2)^{-n/4}$. Accordingly, if $p \in [1,2)$, then

$$\|\widehat{f_\lambda}\|_{p'} \leqslant \|\widehat{f_\lambda}\|_2^{\frac{2}{p'}} \|\widehat{f_\lambda}\|_\infty^{1-\frac{2}{p'}} \leqslant C(1+\lambda^2)^{-\frac{n}{2}(\frac{1}{2}-\frac{1}{p'})} \to 0, \text{ as } \lambda \to \infty.$$

Since $\|f_\lambda\|_p$ is independent of λ, this show that when $p \in [1,2)$, there is no constant C such that $C\|\widehat{f}\|_{p'} \geqslant \|f\|_p$ for all $f \in \mathscr{D}$.

[a] For $0 \neq z \in \mathbb{C}$ and $\operatorname{Re} z \geqslant 0$, one has $\mathscr{F}(e^{-z|x|^2})(\xi) = (2z)^{-n/2}e^{-|\xi|^2/(4z)}$, where $z^{-n/2} := |z|^{-n/2}e^{-in\theta/2}$ if $z = |z|e^{i\theta}$, $\theta \in [-\pi/2, \pi/2]$.

Theorem 3.19. *$\mathscr{D}$ (and hence also $\mathscr{S}$) is dense in L^p ($1 \leqslant p < \infty$) and in $\mathcal{C}_0$.*

Proof. Given $f \in L^p$ and $\sigma > 0$, there exists $g \in \mathcal{C}_c$ with $\|f - g\|_p < \sigma/2$ by the density of $\mathcal{C}_c$ in L^p (cf. [Fol99, Proposition 7.9]). Let $\varphi \in \mathscr{D}$ and $\int \varphi = 1$. Then it is easy to verify $g * \varphi_\varepsilon \in \mathscr{D}$ and $\|g * \varphi_\varepsilon - g\|_p < \sigma/2$ for sufficiently small ε by Theorem 2.9. The same argument applies if L^p is replaced by $\mathcal{C}_0$, $\|\cdot\|_p$ by $\|\cdot\|_\infty$, and the density of $\mathcal{C}_c$ in $\mathcal{C}_0$ (cf. [Fol99, Proposition 4.35]). ❑

Remark 3.20. 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 $\varphi \in \mathscr{S}$.

Theorem 3.21 ($\mathcal{C}^\infty$ Urysohn lemma)**.** *If $K \subset \mathbb{R}^n$ is compact and U is an open set containing K, there exists $f \in \mathscr{D}$ such that $0 \leqslant f \leqslant 1$, $f = 1$ on K, and $\operatorname{supp} f \subset U$.*

Proof. Let $\delta = \operatorname{dist}(K, U^c)$ (the distance from K to U^c, which is positive since K is compact due to [SS05, Lemma 3.1, p.18]), and let $V = \{x : \operatorname{dist}(x, K) < \delta/3\}$. Choose a nonnegative $\varphi \in \mathscr{D}$ such that $\int \varphi = 1$ and

$\varphi(x) = 0$ for $\operatorname{dist}(x, K) \geqslant \delta/3$, and set $f = \chi_V * \varphi$. Then, $f \in \mathscr{D}$, and it is easily checked that $0 \leqslant f \leqslant 1$, $f = 1$ on K, and $\operatorname{supp} f \subset \{x : \operatorname{dist}(x, K) \leqslant 2\delta/3\} \subset U$. ❑

§3.3 Tempered distributions

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

Definition 3.22. 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 $\varphi \in \mathscr{S}$.

Before we discuss some examples, we give alternative characterizations of distributions, which are very useful from the practical point of view. The action of a distribution u on a test function f is represented in either one of the following two ways:

$$\langle u, f\rangle = u(f).$$

Denote

$$\rho_{\alpha,N}(f) = \sup_{|x| \leqslant N} |(\partial^\alpha f)(x)|. \tag{3.6}$$

There exists a simple and important characterization of distributions:

Theorem 3.23. (i) *A linear functional u on $\mathscr{D}(\mathbb{R}^n)$ is a distribution iff for every compact $K \subset \mathbb{R}^n$, there exist $C > 0$ and an integer m such that*

$$|\langle u, f\rangle| \leqslant C \sum_{|\alpha| \leqslant m} \|\partial^\alpha f\|_\infty, \forall f \in \mathcal{C}^\infty(\mathbb{R}^n) \text{ with support in } K. \tag{3.7}$$

(ii) *A linear functional u on $\mathscr{S}(\mathbb{R}^n)$ is a tempered distribution iff there* *exist constant $C > 0$ and integers ℓ and m such that*

$$|\langle u, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant \ell, |\beta|\leqslant m} |\varphi|_{\alpha,\beta}, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{3.8}$$

(iii) *A linear functional u on $\mathcal{C}^\infty(\mathbb{R}^n)$ is a distribution with compact support iff there exist $C > 0$ and integers N, m such that*

$$|\langle u, f\rangle| \leqslant C \sum_{|\alpha|\leqslant m} \rho_{\alpha,N}(f), \quad \forall f \in \mathcal{C}^\infty(\mathbb{R}^n). \tag{3.9}$$

The seminorms $|\cdot|_{\alpha,\beta}$ and $\rho_{\alpha,N}$ are defined in (3.3) and (3.6), respectively.

Proof. We prove only (ii), since the proofs of (i) and (iii) are similar. It is clear that the existence of C, ℓ, m implies the continuity of u.

Suppose u 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 $k \to \infty$ iff $|\varphi_k - \varphi|_{\alpha,\beta} \to 0$ 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 $|\langle u, \varphi\rangle| \leqslant 1$ 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 u, we obtain

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

However, this is the desired inequality with $C = 1/\sigma$. ❑

Example 3.24. 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, it suffices to show that it is continuous at the origin. Then, suppose $\varphi_k \to 0$ in $\mathscr{S}$ as $k \to \infty$. We have for any $q \geqslant 1$ that $\|\varphi_k\|_q$ is dominated by a finite linear combination of seminorms $|\varphi_k|_{\alpha,0}$. 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 3.25. 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

$$|\langle T_f, \varphi\rangle| = \left|\int_{\mathbb{R}} \sum_{k=0}^{m} a_k x^k \varphi(x)dx\right|$$

$$\leqslant \sum_{k=0}^{m} |a_k| \int_{\mathbb{R}} (1+|x|)^{-1-\varepsilon}(1+|x|)^{1+\varepsilon}|x|^k|\varphi(x)|dx$$
$$\leqslant C \sum_{k=0}^{m} |a_k||\varphi|_{k+1+\varepsilon,0} \int_{\mathbb{R}} (1+|x|)^{-1-\varepsilon}dx,$$

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

Example 3.26. The Dirac mass at the origin δ_0. This is defined for $\varphi \in \mathscr{S}$ by

$$\langle \delta_0, \varphi \rangle = \varphi(0).$$

Then, $\delta_0 \in \mathscr{S}'$. The Dirac mass at a point $x_0 \in \mathbb{R}^n$ is defined similarly by

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0).$$

The tempered distributions of Examples 3.24 – 3.26 are called functions or measures. We shall write, in these cases, f and δ_0 instead of T_f and T_{δ_0}. These functions and measures may be considered 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., $\mathscr{B}(\mathbb{R}^n)$.

Suppose that f and g are Schwartz functions and α a multi-index. Integrating by parts $|\alpha|$ times, we obtain

$$\int_{\mathbb{R}^n} (\partial^\alpha f)(x)g(x)dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x)(\partial^\alpha g)(x)dx. \tag{3.10}$$

If we wanted to define the derivative of a tempered distribution u, we would have to give a definition that extends the definition of the derivative of the function and that satisfies (3.10) for $g \in \mathscr{S}'$ and $f \in \mathscr{S}$ if the integrals in (3.10) are interpreted as actions of distributions on functions. We simply use (3.10) to define the derivative of a distribution.

Definition 3.27. Let $u \in \mathscr{S}'$ and α a multi-index. Define

$$\langle \partial^\alpha u, f \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha f \rangle. \tag{3.11}$$

If u is a function, the derivatives of u in the sense of distributions are called ***distributional derivatives***.

In view of Theorem 3.4, it is natural to give the following:

Definition 3.28. Let $u \in \mathscr{S}'$. We define the ***Fourier transform*** $\widehat{u}$ and the ***inverse Fourier transform*** $u^\vee$ of a tempered distribution u by

$$\langle \widehat{u}, f \rangle = \langle u, \widehat{f} \rangle \quad \text{and } \langle u^\vee, f \rangle = \langle u, f^\vee \rangle, \tag{3.12}$$

for all f in $\mathscr{S}$, respectively.

Example 3.29. For $\varphi \in \mathscr{S}$, we have

$$\langle \widehat{\delta_0}, \varphi \rangle = \langle \delta_0, \widehat{\varphi} \rangle = \widehat{\varphi}(0) = \int_{\mathbb{R}^n} e^{-ix\cdot 0} \varphi(x) đx = \left\langle (2\pi)^{-n/2}, \varphi \right\rangle.$$

Thus, $\widehat{\delta_0} = (2\pi)^{-n/2}$ in $\mathscr{S}'$. More generally, since

$$\begin{aligned}\langle \widehat{\partial^\alpha \delta_0}, \varphi \rangle =& \langle \partial^\alpha \delta_0, \widehat{\varphi} \rangle = (-1)^{|\alpha|} \langle \delta_0, \partial^\alpha \widehat{\varphi} \rangle = \langle \delta_0, \widehat{(i\xi)^\alpha \varphi} \rangle \\ =& \langle \widehat{\delta_0}, (i\xi)^\alpha \varphi \rangle = \left\langle (2\pi)^{-n/2} (i\xi)^\alpha, \varphi \right\rangle,\end{aligned}$$

we have $\widehat{\partial^\alpha \delta_0} = (2\pi)^{-n/2}(i\xi)^\alpha$. This calculation indicates that $\widehat{\partial^\alpha \delta_0}$ can be identified with the function $(2\pi)^{-n/2}(i\xi)^\alpha$.

Now observe that the following is true whenever f and g are in $\mathscr{S}$:

$$\begin{aligned}\int_{\mathbb{R}^n} g(x-t) f(x) dx &= \int_{\mathbb{R}^n} f(x+t) g(x) dx, \\ \int_{\mathbb{R}^n} g(ax) f(x) dx &= \int_{\mathbb{R}^n} g(x) a^{-n} f(a^{-1}x) dx, \\ \int_{\mathbb{R}^n} \widetilde{g}(x) f(x) dx &= \int_{\mathbb{R}^n} g(x) \widetilde{f}(x) dx,\end{aligned} \tag{3.13}$$

for all $t \in \mathbb{R}^n$ and $a > 0$, where $\widetilde{\cdot}$ denotes the reflection. Motivated by (3.13), we give the following:

Definition 3.30. The translation $\tau^t u$, the dilation $\delta^a u$, and the reflection $\widetilde{u}$ of a tempered distribution u are defined as follows:

$$\begin{aligned}\langle \tau^t u, f \rangle =& \langle u, \tau^{-t} f \rangle, \\ \langle \delta^a u, f \rangle =& \langle u, a^{-n} \delta^{1/a} f \rangle, \\ \langle \widetilde{u}, f \rangle =& \langle u, \widetilde{f} \rangle,\end{aligned}$$

for all $t \in \mathbb{R}^n$ and $a > 0$. Let A be an invertible matrix, and the composition of a distribution u with an invertible matrix A is the distribution

$$\langle u^A, \varphi \rangle = |\det A|^{-1} \langle u, \varphi^{A^{-1}} \rangle,$$

where $\varphi^{A^{-1}}(x) = \varphi(A^{-1}x)$.

It is easy to see that the operations of translation, dilation, reflection, and differentiation are continuous on tempered distributions.

Example 3.31. The Dirac mass at the origin δ_0 is equal to its reflection, while $\delta^a \delta_0 = a^{-n} \delta_0$ for $a > 0$. Additionally, $\tau^x \delta_0 = \delta_x$ for any $x \in \mathbb{R}^n$.

Now, observe that for f, g and h in $\mathscr{S}$, we have

$$\int_{\mathbb{R}^n} (h * g)(x) f(x) dx = \int_{\mathbb{R}^n} g(x) (\widetilde{h} * f)(x) dx.$$

Motivated by this identity, we define the convolution of a function with a tempered distribution as follows:

Definition 3.32. Let $u \in \mathscr{S}'$ and $h \in \mathscr{S}$. Define the convolution $h * u$ by

$$\langle h * u, f \rangle = \langle u, \widetilde{h} * f \rangle, \qquad f \in \mathscr{S}.$$

Example 3.33. Let $u = \delta_{x_0}$ and $f \in \mathscr{S}$. Then, $f * \delta_{x_0}$ is the function $\tau^{x_0} f$, since for $h \in \mathscr{S}$, we have

$$\langle f * \delta_{x_0}, h \rangle = \langle \delta_{x_0}, \widetilde{f} * h \rangle = (\widetilde{f} * h)(x_0) = \int_{\mathbb{R}^n} f(x - x_0) h(x) dx = \langle \tau^{x_0} f, h \rangle.$$

It follows that convolution with δ_0 is the identity operator by taking $x_0 = 0$.

We now define the product of a function and a distribution.

Definition 3.34. Let $u \in \mathscr{S}'$ and $h \in \mathcal{C}^\infty_{\text{poly}}$. The product hu of h and u is defined by

$$\langle hu, f \rangle = \langle u, hf \rangle, \qquad f \in \mathscr{S}. \tag{3.14}$$

Note that $hf \in \mathscr{S}$, and thus, (3.14) is well-defined. (The product of an arbitrary $\mathcal{C}^\infty$ function with a tempered distribution is not defined.)

Example 3.35. Let $T \in \mathscr{S}'$ and $\varphi \in \mathscr{D}$ 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 $\varphi_k = \varphi(x/k)$.

Next, we give a proposition that extends the properties of the Fourier transform to tempered distributions.

Proposition 3.36. *Given $u, v \in \mathscr{S}'(\mathbb{R}^n)$, $f_j, f \in \mathscr{S}$, $y \in \mathbb{R}^n$, $b \in \mathbb{C}$, $\alpha \in \mathbb{N}_0^n$, and $a > 0$, we have*

(i) $\widehat{u+v} = \widehat{u} + \widehat{v}$, $\widehat{bu} = b\widehat{u}$,

(ii) $\widehat{\widetilde{u}} = \widetilde{\widehat{u}}$,

(iii) $\widehat{\tau^y u}(\xi) = e^{-iy\cdot\xi}\widehat{u}(\xi)$, $\widehat{e^{ix\cdot y}u(x)} = \tau^y \widehat{u}$,

(iv) $\widehat{\delta^a u} = (\widehat{u})_a = a^{-n}\delta^{a^{-1}}\widehat{u}$,

(v) $\widehat{\partial^\alpha u}(\xi) = (i\xi)^\alpha \widehat{u}(\xi)$, $\partial^\alpha \widehat{u} = \widehat{(-ix)^\alpha u(x)}$,

(vi) $(\widehat{u})^\vee = u = \widehat{u^\vee}$,

(vii) $(2\pi)^{-n/2}\widehat{f*u} = \hat{f}\hat{u}$.

Proof. All the statements can be proved easily using duality and the corresponding statements for Schwartz functions. ❑

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

Theorem 3.37. *If $u \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$, then $\varphi * u$ is a $\mathcal{C}^\infty_{poly}$ function and*

$$(\varphi * u)(x) = \langle u, \tau^x \widetilde{\varphi} \rangle, \tag{3.15}$$

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

Proof. We first prove (3.15). Let $\psi \in \mathscr{S}(\mathbb{R}^n)$. We have

$$\begin{aligned} \langle \varphi * u, \psi \rangle =& \langle u, \widetilde{\varphi} * \psi \rangle \\ =& u\left(\int_{\mathbb{R}^n} \widetilde{\varphi}(\cdot - y)\psi(y)dy\right) \\ =& u\left(\int_{\mathbb{R}^n} (\tau^y \widetilde{\varphi})(\cdot)\psi(y)dy\right) \\ =& \int_{\mathbb{R}^n} \langle u, \tau^y \widetilde{\varphi} \rangle \psi(y)dy, \end{aligned} \tag{3.16}$$

where the last step is justified by the continuity of u and by the fact that the Riemann sums of the inner integral in (3.16) converge uniformly to that integral in the topology of $\mathscr{S}$, a fact that will be justified later. This calculation implies (3.15).

We now show that $\varphi * u$ is a $\mathcal{C}^\infty$ function. Let $e_j = (0, \cdots, 1, \cdots, 0)$ with 1 in the jth entry and zero elsewhere. Then by part (iii) in Proposition 3.13,

$$\frac{\tau^{-he_j}\tau^x\widetilde{\varphi} - \tau^x\widetilde{\varphi}}{h} \to \partial_j \tau^x \widetilde{\varphi} = \tau^x \partial_j \widetilde{\varphi},$$

in $\mathscr{S}$ as $h \to 0$. Thus, since u is linear and continuous, we have from (3.15)

$$\frac{\tau^{he_j}(\varphi * u)(x) - (\varphi * u)(x)}{h} = u\left(\frac{\tau^{-he_j}(\tau^x\widetilde{\varphi}) - \tau^x\widetilde{\varphi}}{h}\right) \to \langle u, \tau^x(\partial_j \widetilde{\varphi}) \rangle$$

as $h \to 0$. The same calculation for higher-order derivatives shows that $\varphi * u \in \mathcal{C}^\infty$ and that $\partial^\gamma(\varphi * u) = (\partial^\gamma \varphi) * u$ for all multi-indices γ. It follows from Theorem 3.23 that for some C, m and k we have

$$|\partial^\alpha(\varphi * u)(x)| \leqslant C \sum_{\substack{|\gamma| \leqslant m \\ |\beta| \leqslant k}} \sup_{y \in \mathbb{R}^n} |y^\gamma \tau^x (\partial^{\alpha+\beta}\widetilde{\varphi})(y)|$$

$$
\begin{aligned}
&= C \sum_{\substack{|\gamma|\leqslant m\\|\beta|\leqslant k}} \sup_{y\in\mathbb{R}^n} |(x+y)^\gamma(\partial^{\alpha+\beta}\widetilde{\varphi})(y)| \quad (\text{changing: } y-x\to y)\\
&\leqslant C_m \sum_{|\beta|\leqslant k} \sup_{y\in\mathbb{R}^n} |(1+|x|^m+|y|^m)(\partial^{\alpha+\beta}\widetilde{\varphi})(y)|\\
&\leqslant C_{m,k,\alpha} \sup_{y\in\mathbb{R}^n} \frac{1+|x|^m+|y|^m}{(1+|y|)^N} \quad (\text{taking } N>m)\\
&\leqslant C_{m,k,\alpha}(1+|x|^m),
\end{aligned}
$$

which clearly implies that $\partial^\alpha(\varphi * u)$ grows at most polynomially at infinity.

Next, we return to the point left open concerning the convergence of the Riemann sums in (3.16) in the topology of $\mathscr{S}(\mathbb{R}^n)$. For each $N = 1, 2, \cdots$, consider a partition of $[-N, N]^n$ into $(2N^2)^n$ cubes Q_m of side length $1/N$ and let y_m be the center of each Q_m. For multi-indices α and β, we must show that

$$
D_N(x) = \sum_{m=1}^{(2N^2)^n} x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y_m)\psi(y_m)|Q_m| - \int_{\mathbb{R}^n} x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y)\psi(y)dy
$$

converges to zero in $L^\infty(\mathbb{R}^n)$ as $N\to\infty$. We have by the mean value theorem

$$
\begin{aligned}
& x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y_m)\psi(y_m)|Q_m| - \int_{Q_m} x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y)\psi(y)dy\\
=& \int_{Q_m} x^\alpha [\partial_x^\beta \widetilde{\varphi}(x-y_m)\psi(y_m) - \partial_x^\beta \widetilde{\varphi}(x-y)\psi(y)]dy\\
=& \int_{Q_m} x^\alpha (y_m - y)\cdot(\nabla(\partial_x^\beta \widetilde{\varphi}(x-\cdot)\psi))(\xi)dy\\
=& \int_{Q_m} x^\alpha (y_m - y)\cdot(-\nabla\partial_x^\beta \widetilde{\varphi}(x-\cdot)\psi + \nabla\psi\partial_x^\beta \widetilde{\varphi}(x-\cdot))(\xi)dy
\end{aligned}
$$

for some $\xi = y + \theta(y_m - y)$, where $\theta\in[0,1]$. We see that $|y - y_m| \leqslant \sqrt{n}/2N$ and the last integrand

$$
\begin{aligned}
& |x^\alpha(y-y_m)\cdot(-\nabla\partial_x^\beta\widetilde{\varphi}(x-\xi)\psi(\xi) + \nabla\psi(\xi)\partial_x^\beta\widetilde{\varphi}(x-\xi))|\\
\leqslant& C|x|^{|\alpha|}\frac{\sqrt{n}}{2N}\frac{1}{(2+|\xi|)^M}\frac{1}{(1+|x-\xi|)^{M/2}} \quad (\text{for large } M)\\
\leqslant& C|x|^{|\alpha|}\frac{\sqrt{n}}{N}\frac{1}{(2+|\xi|)^{M/2}}\frac{1}{(1+|x|)^{M/2}}\\
\leqslant& C|x|^{|\alpha|}\frac{\sqrt{n}}{N}\frac{1}{(1+|y|)^{M/2}}\frac{1}{(1+|x|)^{M/2}},
\end{aligned}
$$

since $(1+|x-\xi|)(2+|\xi|) \geqslant 1+|x-\xi|+|\xi| \geqslant 1+|x|$, and $|y| \leqslant |\xi| + \theta|y-y_m| \leqslant |\xi| + \sqrt{n}/2N \leqslant |\xi|+1$ for $N \geqslant \sqrt{n}/2$. Inserting the estimates obtained for the integrand, we obtain

$$
|D_N(x)| \leqslant \frac{C}{N}\frac{|x|^{|\alpha|}}{(1+|x|)^{M/2}}\int_{[-N,N]^n}\frac{dy}{(1+|y|)^{M/2}}
$$

$$+\int_{([-N,N]^n)^c} |x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y)\psi(y)|dy.$$

The first integral in the preceding expression is bounded by

$$\omega_{n-1}\int_0^{\sqrt{n}N} \frac{r^{n-1}dr}{(1+r)^{M/2}} \leqslant \omega_{n-1}\int_0^{\sqrt{n}N} \frac{dr}{(1+r)^{\frac{M}{2}-n+1}} \leqslant \frac{2\omega_{n-1}}{M-2n},$$

where we pick $M > 2n$, while the second integral is bounded by

$$\begin{aligned}
&\int_{([-N,N]^n)^c} \frac{C|x|^{|\alpha|}}{(1+|x-y|)^{M/2}} \frac{dy}{(1+|y|)^M} \\
\leqslant& \frac{C|x|^{|\alpha|}}{(1+|x|)^{M/2}} \int_{([-N,N]^n)^c} \frac{dy}{(1+|y|)^{M/2}} \\
\leqslant& C\omega_{n-1}\int_N^\infty \frac{r^{n-1}dr}{(1+r)^{M/2}} \leqslant C\frac{2\omega_{n-1}}{M-2n}N^{n-M/2},
\end{aligned}$$

for $M > 2\max(n, |\alpha|)$ since $(1+|x-y|)(1+|y|) \geqslant 1+|x-y|+|y| \geqslant 1+|x|$. From these estimates, it follows that

$$\sup_{x\in\mathbb{R}^n} |D_N(x)| \leqslant C(\frac{1}{N} + \frac{1}{N^{\frac{M}{2}-n}}) \to 0, \quad \text{as } N\to\infty.$$

Therefore, $\lim_{N\to\infty}\sup_{x\in\mathbb{R}^n} |D_N(x)| = 0$. ❑

We observe that if a function g is supported in a set K, then for all $f \in \mathscr{D}(K^c)$ we have

$$\int_{\mathbb{R}^n} f(x)g(x)dx = 0. \tag{3.17}$$

Moreover, the support of g is the intersection of all closed sets K with the property (3.17) for all f in $\mathscr{D}(K^c)$. Motivated by this observation we give the following:

Definition 3.38. Let $u \in \mathscr{D}'(\mathbb{R}^n)$. The ***support of u*** ($\operatorname{supp} u$) is the intersection of all closed sets K with the property

$$\varphi \in \mathscr{D}(\mathbb{R}^n), \quad \operatorname{supp}\varphi \subset K^c \Longrightarrow \langle u, \varphi\rangle = 0. \tag{3.18}$$

Example 3.39. $\operatorname{supp}\delta_{x_0} = \{x_0\}$.

Along the same lines, we give the following definition:

Definition 3.40. We say that a ***distribution $u \in \mathscr{D}'(\mathbb{R}^n)$ coincides with the function h*** on an open set Ω if

$$\langle u, f\rangle = \int_{\mathbb{R}^n} f(x)h(x)dx, \quad \forall f \in \mathscr{D}(\Omega). \tag{3.19}$$

When (3.19) occurs, we often say that u agrees with h away from Ω^c.

This definition implies $\operatorname{supp}(u-h) \subset \Omega^c$.

Example 3.41. The distribution $|x|^2 + \delta_{a_1} + \delta_{a_2}$, where $a_1, a_2 \in \mathbb{R}^n$, coincides with the function $|x|^2$ on any open set not containing the points a_1 and a_2.

We have the following characterization of distributions supported at a single point.

Proposition 3.42. *If $u \in \mathscr{S}'(\mathbb{R}^n)$ is supported in the singleton $\{x_0\}$, then there exists an integer k and complex numbers a_α such that*

$$u = \sum_{|\alpha| \leqslant k} a_\alpha \partial^\alpha \delta_{x_0}.$$

Proof. Without loss of generality, we may assume that $x_0 = 0$. By (3.8), we have for some C, m, and k,

$$|\langle u, f\rangle| \leqslant C \sum_{\substack{|\alpha|\leqslant m \\ |\beta| \leqslant k}} \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)|, \quad \forall f \in \mathscr{S}(\mathbb{R}^n).$$

We now prove that if $\varphi \in \mathscr{S}$ satisfies

$$(\partial^\alpha \varphi)(0) = 0, \quad \forall |\alpha| \leqslant k, \tag{3.20}$$

then $\langle u, \varphi\rangle = 0$. To see this, fix a φ satisfying (3.20) and let $\zeta(x)$ be a smooth function on $\mathbb{R}^n$ that is equal to 1 when $|x| \geqslant 2$ and equal to zero for $|x| \leqslant 1$. Let $\zeta^\varepsilon(x) = \zeta(x/\varepsilon)$. Then using (3.20) and the continuity of the derivatives of φ at the origin, it is not hard to show that $|\zeta^\varepsilon\varphi - \varphi|_{\alpha,\beta} \to 0$ as $\varepsilon \to 0$ for all $|\alpha| \leqslant m$ and $|\beta| \leqslant k$. Then,

$$|\langle u, \varphi\rangle| \leqslant |\langle u, \zeta^\varepsilon \varphi\rangle| + |\langle u, \zeta^\varepsilon\varphi - \varphi\rangle| \leqslant 0 + C \sum_{\substack{|\alpha|\leqslant m \\ |\beta| \leqslant k}} |\zeta^\varepsilon\varphi - \varphi|_{\alpha,\beta} \to 0,$$

as $\varepsilon \to 0$. This proves our assertion.

Now, let $f \in \mathscr{S}(\mathbb{R}^n)$. Let $\eta \in \mathscr{D}(\mathbb{R}^n)$ be equal to 1 in a neighborhood of the origin. Write

$$f(x) = \eta(x) \left(\sum_{|\alpha|\leqslant k} \frac{(\partial^\alpha f)(0)}{\alpha!} x^\alpha + h(x) \right) + (1 - \eta(x)) f(x), \tag{3.21}$$

where $h(x) = O(|x|^{k+1})$ as $|x| \to 0$. Then, ηh satisfies (3.20) and hence $\langle u, \eta h\rangle = 0$ by the claim. Additionally,

$$\langle u, (1-\eta) f\rangle = 0$$

by our hypothesis. Applying u to both sides of (3.21), we obtain

$$\langle u, f\rangle = \sum_{|\alpha|\leqslant k} \frac{(\partial^\alpha f)(0)}{\alpha!} \langle u, x^\alpha \eta(x)\rangle = \sum_{|\alpha|\leqslant k} a_\alpha \langle \partial^\alpha \delta_0, f\rangle,$$

with $a_\alpha = (-1)^{|\alpha|} \langle u, x^\alpha \eta(x)\rangle / \alpha!$. This proves the result. ❑

An immediate consequence is the following result.

Corollary 3.43. *Let $u \in \mathscr{S}'(\mathbb{R}^n)$. If $\widehat{u}$ is supported in the singleton $\{\xi_0\}$, then u is a finite linear combination of functions $(-i\xi)^\alpha e^{i\xi\cdot\xi_0}$, where $\alpha \in \mathbb{N}_0^n$. In particular, if $\widehat{u}$ is supported at the origin, then u is a polynomial.*

Proof. Proposition 3.42 gives that $\widehat{u}$ is a linear combination of derivatives of Dirac masses at ξ_0, i.e.,

$$\widehat{u} = \sum_{|\alpha|\leqslant k} a_\alpha \partial^\alpha \delta_{\xi_0}.$$

Then, Proposition 3.36, Example 3.31 and Example 3.29 yield

$$\begin{aligned} u &= \sum_{|\alpha|\leqslant k} a_\alpha \left(\partial^\alpha \delta_{\xi_0}\right)^\vee = \sum_{|\alpha|\leqslant k} a_\alpha \widetilde{\widehat{\partial^\alpha \delta_{\xi_0}}} \\ &= \sum_{|\alpha|\leqslant k} a_\alpha \widetilde{(i\xi)^\alpha \widehat{\delta_{\xi_0}}} = \sum_{|\alpha|\leqslant k} a_\alpha \widetilde{(i\xi)^\alpha \widehat{\tau^{\xi_0}\delta_0}} \\ &= (2\pi)^{-n/2} \sum_{|\alpha|\leqslant k} a_\alpha \widetilde{(i\xi)^\alpha e^{-i\xi\cdot\xi_0}} \\ &= (2\pi)^{-n/2} \sum_{|\alpha|\leqslant k} a_\alpha (-i\xi)^\alpha e^{i\xi\cdot\xi_0}. \end{aligned}$$ ❑

Proposition 3.44. *Distributions with compact support are exactly those whose support is a compact set, i.e.,*

$$u \in \mathscr{E}'(\mathbb{R}^n) \iff \operatorname{supp} u \text{ is compact}.$$

Proof. To prove this assertion, we start with a distribution u with compact support. Then, there exist $C, N, m > 0$ such that (3.9) holds. For a $\mathcal{C}^\infty$ function f whose support is contained in $B(0,N)^c$, the expression on the right in (3.9) vanishes, and we must therefore have $\langle u, f\rangle = 0$. This shows that the support of u is contained in $\overline{B(0,N)}$; hence, it is bounded, and since it is already closed (as an intersection of closed sets), it must be compact.

Conversely, if the support of u as defined in Definition 3.38 is a compact set, then there exists an $N > 0$ such that $\operatorname{supp} u \subset \overline{B(0,N)}$. We take $\eta \in \mathscr{D}$ that is equal to 1 on $\overline{B(0,N)}$ and vanishes off $B(0, N+1)$. Then, for $h \in \mathscr{D}$, the support of $h(1-\eta)$ does not meet the support of u, and we must have

$$\langle u, h\rangle = \langle u, h\eta\rangle + \langle u, h(1-\eta)\rangle = \langle u, h\eta\rangle.$$

The distribution u can be thought of as an element of $\mathscr{E}'$ by defining for $f \in \mathcal{C}^\infty(\mathbb{R}^n)$

$$\langle u, f\rangle = \langle u, f\eta\rangle.$$

Taking m to be the integer that corresponds to the compact set $K = \overline{B(0, N+1)}$ in (3.7) and using that the L^∞ norm of $\partial^\alpha(f\eta)$ is controlled by a finite sum of seminorms $\rho_{\alpha,N+1}(f)$ with $|\alpha| \leqslant m$, we obtain the validity of (3.9) for $f \in \mathcal{C}^\infty$. ❑

For distributions with compact support, we have the following important result.

Theorem 3.45. *If $u \in \mathscr{E}'(\mathbb{R}^n)$, then $\widehat{u}$ is a real analytic function on $\mathbb{R}^n$. In particular, $\widehat{u} \in \mathcal{C}^\infty_{poly}$. Moreover, $\widehat{u}$ has a holomorphic extension on $\mathbb{C}^n$.*

Proof. Since $u \in \mathscr{E}' \subset \mathscr{S}'$, we have for $f \in \mathscr{S}$

$$\begin{aligned}\langle \widehat{u}, f\rangle =&\langle u, \widehat{f}\rangle = u\left(\int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) đx\right)\\ =& \int_{\mathbb{R}^n} u\left(e^{-ix\cdot(\cdot)}\right) f(x) đx,\end{aligned}$$

provided that we can justify the passage of u inside the integral. The reason for this is that the Riemann sums of the integral of $e^{-ix\cdot\xi} f(x)$ over $\mathbb{R}^n$ converge to it in the topology of $\mathcal{C}^\infty$, and thus the linear functional u can be interchanged with the integral. To justify this, we argue as in the proof of Theorem 3.37. For each $N \in \mathbb{N}$, we consider a partition of $[-N, N]^n$ into $(2N^2)^n$ cubes Q_m of side length $1/N$ and let y_m be the center of each Q_m. For $\alpha \in \mathbb{N}_0^n$, let

$$D_N(\xi) = \sum_{m=1}^{(2N^2)^n} e^{-iy_m\cdot\xi}(-iy_m)^\alpha f(y_m)|Q_m| - \int_{\mathbb{R}^n} e^{-ix\cdot\xi}(-ix)^\alpha f(x) dx.$$

We must show that for every $M > 0$, $\sup_{|\xi|\leqslant M} |D_N(\xi)|$ converges to zero as $N \to \infty$. Setting $g(x) = (-ix)^\alpha f(x) \in \mathscr{S}$, we write

$$D_N(\xi) = \sum_{m=1}^{(2N^2)^n} \int_{Q_m} [e^{-iy_m\cdot\xi} g(y_m) - e^{-ix\cdot\xi} g(x)] dx - \int_{([-N,N]^n)^c} e^{-ix\cdot\xi} g(x) dx.$$

Using the mean value theorem, we bound the absolute value of the expression inside the square brackets by

$$(|\nabla g(z_m)| + |\xi||g(z_m)|)\frac{\sqrt{n}}{2N} \leqslant \frac{C_K(1+|\xi|)}{(2+|z_m|)^K}\frac{\sqrt{n}}{N},$$

for some point $z_m = x + \theta(y_m - x)$ in the cube Q_m where $\theta \in [0,1]$. Since $2 + |z_m| \geqslant 1 + |x|$ if $N > \sqrt{n}/2$, and then for $|\xi| \leqslant M$,

$$\begin{aligned}\sum_{m=1}^{(2N^2)^n} \int_{Q_m} \frac{C_K(1+|\xi|)}{(2+|z_m|)^K} dx \leqslant& \sum_{m=1}^{(2N^2)^n} \int_{Q_m} \frac{C_K(1+|\xi|)}{(1+|x|)^K} dx\\ \leqslant& C_K(1+M)\int_0^{\sqrt{n}N} \frac{r^{n-1}dr}{(1+r)^K} \leqslant C_K(1+M) < \infty\end{aligned}$$

provided $K > n$, and for $L > n$,

$$\int_{([-N,N]^n)^c} \frac{dy}{(1+|y|)^L} \leqslant \omega_{n-1} \int_N^\infty \frac{r^{n-1}dr}{(1+r)^L} \leqslant \frac{\omega_{n-1}}{L-n} N^{n-L},$$

it follows that $\sup\limits_{|\xi|\leqslant M} |D_N(\xi)| \to 0$ as $N \to \infty$ by noticing $g \in \mathscr{S}$.

Let $p(\xi)$ be a polynomial; then, the action of $u \in \mathscr{E}'$ on the $\mathcal{C}^\infty$ function $\xi \mapsto p(\xi)e^{-ix\cdot\xi}$ is a well-defined function of x, which we denote by $u(p(\cdot)e^{-ix\cdot(\cdot)})$. Here, $x \in \mathbb{R}^n$, but the same assertion is valid if $x \in \mathbb{R}^n$ is replaced by $z \in \mathbb{C}^n$. In this case, we define the dot product of ξ and z via $\xi \cdot z = \sum\limits_{k=1}^{n} \xi_k z_k$.

It is straightforward to verify that the function of z

$$F(z) = (2\pi)^{-n/2} u(e^{-iz\cdot(\cdot)})$$

defined on $\mathbb{C}^n$ is holomorphic, in fact entire. Indeed, the continuity and linearity of u and the fact that $(e^{-i\xi_j h} - 1)/h \to -i\xi_j$ in $\mathcal{C}^\infty(\mathbb{R})$ as $h \to 0$, $h \in \mathbb{C}$, imply that F is holomorphic in every variable and its derivative with respect to z_j is the action of the distribution u on the $\mathcal{C}^\infty$ function

$$\xi \mapsto (-i\xi_j)e^{-i\sum_{j=1}^n z_j \xi_j}.$$

By induction, it follows that for all $\alpha \in \mathbb{N}_0^n$, we have

$$\partial_{z_1}^{\alpha_1} \cdots \partial_{z_n}^{\alpha_n} F = u\left((-i(\cdot))^\alpha e^{-i\sum_{j=1}^n z_j (\cdot)_j} \right).$$

Since F is entire, its restriction on $\mathbb{R}^n$, i.e., $F(x_1, \cdots, x_n)$, where $x_j = \operatorname{Re} z_j$, is real analytic. Additionally, an easy calculation using (3.9) and Leibniz's rule yields that the restriction F on $\mathbb{R}^n$ and all of its derivatives have polynomial growth at infinity.

Therefore, we conclude that the distribution $\hat{u}(x)$ can be identified with the real analytic function $F(x)$ whose derivatives have polynomial growth at infinity. ❑

§3.4 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 3.46. 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 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 the convolution type. We have the following:

Theorem 3.47. *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*

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

The theorem will be a consequence of the following lemma.

Lemma 3.48. *Let $1 \leqslant p \leqslant \infty$. If $f \in L^p(\mathbb{R}^n)$ has derivatives of all orders $\leqslant n+1$ in the L^p norm, 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 $p = 1$, and we shall show $\hat{f} \in L^1$. By part (viii) and part (x) in Proposition 3.2, we have

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

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 obtain

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

Thus, by Theorem 3.5 due to $f, \hat{f} \in L^1$, f equals almost everywhere a continuous function g and by part (x) in Proposition 3.2,

$$|g(0)| \leqslant \|f\|_\infty \leqslant (2\pi)^{-n/2}\|\widehat{f}\|_1 \leqslant C \sum_{|\alpha|\leqslant n+1} \|\partial^\alpha f\|_1.$$

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, by the above argument, $f\varphi$ 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 $\partial^\alpha(f\varphi) = \sum\limits_{\mu+\nu=\alpha} \frac{\alpha!}{\mu!\nu!}\partial^\mu f \partial^\nu \varphi$, 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 $A \geqslant C'\|\partial^\nu \varphi\|_\infty$, $|\nu| \leqslant |\alpha|$, 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,

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

However, 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. ❑

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

Proof of Theorem 3.47. We first prove that

$$\partial^\beta T f = T \partial^\beta f, \quad \forall f \in \mathscr{S}(\mathbb{R}^n). \tag{3.22}$$

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 3.13, we have $\frac{\tau^h f - f}{h_j} \to -\frac{\partial f}{\partial x_j}$ in $\mathscr{S}$ as $|h| \to 0$ and in the L^p norm. Since T is a 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 obtain (3.22). By Lemma 3.48, Tf equals almost everywhere a continuous function g_f satisfying

$$\begin{aligned}|g_f(0)| \leqslant & C \sum_{|\beta|\leqslant n+1} \|\partial^\beta (Tf)\|_q = C \sum_{|\beta|\leqslant n+1} \|T(\partial^\beta f)\|_q \\ \leqslant & C\|T\| \sum_{|\beta|\leqslant n+1} \|\partial^\beta f\|_p \leqslant C \sum_{|\alpha|\leqslant m, |\beta|\leqslant n+1} |f|_{\alpha,\beta}.\end{aligned}$$

Then, by Theorem 3.23 (ii), the mapping $f \mapsto g_f(0)$ is a continuous linear functional on $\mathscr{S}$, denoted by u_1. We claim that $u = \widetilde{u_1}$ is the linear functional we are seeking. Indeed, if $f \in \mathscr{S}$, using Theorem 3.37, we obtain

$$\begin{aligned}(u * f)(x) =& \langle u, \tau^x \widetilde{f}\rangle = \langle u, \widetilde{\tau^{-x} f}\rangle = \langle \widetilde{u}, \tau^{-x} f\rangle = \langle u_1, \tau^{-x} f\rangle \\ =& (T(\tau^{-x} f))(0) = (\tau^{-x} Tf)(0) = Tf(x).\end{aligned}$$

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

Now, we give a characterization of operators commuting with translations in $p = 2$.

Theorem 3.49. *Let T be a bounded linear transformation mapping $L^2(\mathbb{R}^n)$ to itself. Then T commutes with translation iff there exists an $m \in L^\infty(\mathbb{R}^n)$ such that $Tf = u * f$ with $\widehat{u} = (2\pi)^{-n/2} m$, for all $f \in L^2(\mathbb{R}^n)$. We also have $\|T\| = \|m\|_\infty$.*

Proof. Now, we prove the necessity. 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 3.47, there exists a unique tempered distribution u such that $Tf = u * f$ for all $f \in \mathscr{S}$. The remainder is to prove $\widehat{u} \in L^\infty(\mathbb{R}^n)$.

Let $m = (2\pi)^{n/2}\widehat{u}$, from

$$\|m\widehat{\varphi}\|_2 = \|\widehat{u * \varphi}\|_2 = \|u * \varphi\|_2 \leqslant \|T\|\|\varphi\|_2 = \|T\|\|\widehat{\varphi}\|_2, \quad \forall \varphi \in \mathscr{S}.$$

it follows that

$$\int_{\mathbb{R}^n} \left(\|T\|^2 - |m|^2\right) |\widehat{\varphi}|^2 d\xi \geqslant 0, \quad \forall \varphi \in \mathscr{S}.$$

This implies that $\|T\|^2 - |m|^2 \geqslant 0$ for a.e. $\xi \in \mathbb{R}^n$. Hence, $m \in L^\infty(\mathbb{R}^n)$ and $\|m\|_\infty \leqslant \|T\|$.

Finally, we can show the sufficiency easily. If $\widehat{u} = (2\pi)^{-n/2} m \in L^\infty(\mathbb{R}^n)$, the Plancherel theorem immediately implies that

$$\|Tf\|_2 = \|u * f\|_2 = \|m\widehat{f}\|_2 \leqslant \|m\|_\infty \|f\|_2,$$

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

Thus, if $m = (2\pi)^{n/2}\widehat{u} \in L^\infty$, then $\|T\| = \|m\|_\infty$. ❑

§ 3.5 Fourier multipliers on L^p

We have shown that the translation invariant L^p-L^q operators are convolution-type operators in Theorem 3.47. In this section, we briefly introduce Fourier multipliers on L^p.

Definition 3.50. Let $1 \leqslant p \leqslant \infty$ and $m \in \mathscr{S}'$. m is called a ***Fourier multiplier on*** $L^p(\mathbb{R}^n)$ if the convolution $m^\vee * f \in L^p(\mathbb{R}^n)$ for all $f \in \mathscr{S}(\mathbb{R}^n)$, and

$$\|m\|_{\mathcal{M}_p(\mathbb{R}^n)} = (2\pi)^{-n/2} \sup_{\|f\|_p=1} \|m^\vee * f\|_p$$

is finite. The linear space of all such m is denoted by $\mathcal{M}_p(\mathbb{R}^n)$.

Since $\mathscr{S}$ is dense in L^p $(1 \leqslant p < \infty)$, the mapping from $\mathscr{S}$ to L^p: $f \mapsto m^\vee * f$ can be extended to a mapping from L^p to L^p with the same norm. We write $m^\vee * f$ also for the values of this extended mapping.

For $p = \infty$, we can characterize $\mathcal{M}_p$. Considering the map:

$$f \mapsto m^\vee * f \quad \text{for } f \in \mathscr{S},$$

we have

$$m \in \mathcal{M}_\infty \Leftrightarrow |(m^\vee * f)(0)| \leqslant C\|f\|_\infty, \quad f \in \mathscr{S}. \tag{3.23}$$

Indeed, if $m \in \mathcal{M}_\infty$, we have $m^\vee * f \in \mathcal{C}^\infty_{\text{poly}}$ by Theorem 3.37, and then

$$|(m^\vee * f)(0)| \leqslant \frac{\|m^\vee * f\|_\infty}{\|f\|_\infty}\|f\|_\infty \leqslant C\|f\|_\infty.$$

On the other hand, if $|(m^\vee * f)(0)| \leqslant C\|f\|_\infty$, we can obtain

$$\begin{aligned}\|m^\vee * f\|_\infty = &\sup_{x\in\mathbb{R}^n} |(m^\vee * f)(x)| = \sup_{x\in\mathbb{R}^n} |[m^\vee * (f(x+\cdot))](0)| \\ \leqslant &C\|f(x+\cdot)\|_\infty = C\|f\|_\infty,\end{aligned}$$

which yields $m \in \mathcal{M}_\infty$.

However, (3.23) also means that $m^\vee$ is a bounded measure on $\mathbb{R}^n$ since the dual space of L^∞ is the space of all bounded finitely additive signed measures on $\mathbb{R}^n$ that are absolutely continuous w.r.t. Lebesgue measure. Thus, $\mathcal{M}_\infty$ is equal to the space of all Fourier transforms of bounded measures. Moreover, $\|m\|_{\mathcal{M}_\infty}$ is equal to the total variation norm of $m^\vee$. In view of the inequality above and the Hahn-Banach theorem, we may extend the mapping $f \mapsto m^\vee * f$ from $\mathscr{S}$ to L^∞ to a mapping from L^∞ to L^∞ without increasing its norm. We also write the extended mapping as $f \mapsto m^\vee * f$ for $f \in L^\infty$.

Theorem 3.51. (i) *Let* $1 \leqslant p \leqslant \infty$ *and* $1/p + 1/p' = 1$; *then, we have*

$$\mathcal{M}_p(\mathbb{R}^n) = \mathcal{M}_{p'}(\mathbb{R}^n) \quad \text{(equal norms)}. \tag{3.24}$$

(ii) $\mathcal{M}_1(\mathbb{R}^n) = \{m \in \mathscr{S}'(\mathbb{R}^n) : m^\vee$ *is a bounded measure on* $\mathbb{R}^n\}$, *and* $\|m\|_{\mathcal{M}_1(\mathbb{R}^n)} = (2\pi)^{-n/2}\|m^\vee\|_1$.

(iii) $\mathcal{M}_2(\mathbb{R}^n) = L^\infty(\mathbb{R}^n)$ *with equal norms.*

(iv) *Let* $1 \leqslant p_0, p_1 \leqslant \infty$ *and* $1/p = (1-\theta)/p_0 + \theta/p_1$ $(0 \leqslant \theta \leqslant 1)$, *we have* $\forall m \in \mathcal{M}_{p_0}(\mathbb{R}^n) \cap \mathcal{M}_{p_1}(\mathbb{R}^n)$

$$\|m\|_{\mathcal{M}_p(\mathbb{R}^n)} \leqslant \|m\|^{1-\theta}_{\mathcal{M}_{p_0}(\mathbb{R}^n)} \|m\|^{\theta}_{\mathcal{M}_{p_1}(\mathbb{R}^n)}. \tag{3.25}$$

(v) *The norm* $\|\cdot\|_{M^p(\mathbb{R}^n)}$ *decreases with* $p \in [1,2]$, *and*

$$\mathcal{M}_1 \hookrightarrow \mathcal{M}_p \hookrightarrow \mathcal{M}_q \hookrightarrow \mathcal{M}_2, \quad (1 \leqslant p \leqslant q \leqslant 2). \tag{3.26}$$

Proof. (i) Let $f \in L^p$, $g \in L^{p'}$ and $m \in \mathcal{M}_p$. Then, we have

$$\begin{aligned}
(2\pi)^{n/2}\|m\|_{\mathcal{M}_{p'}} &= \sup_{\|g\|_{p'}=1} \|m^\vee * g\|_{p'} = \sup_{\|f\|_p=\|g\|_{p'}=1} |\langle m^\vee * g, \widetilde{f}\rangle| \\
&= \sup_{\|f\|_p=\|g\|_{p'}=1} |(m^\vee * g * f)(0)| \\
&= \sup_{\|f\|_p=\|g\|_{p'}=1} |(m^\vee * f * g)(0)| \\
&= \sup_{\|f\|_p=\|g\|_{p'}=1} |\langle m^\vee * f, \tilde{g}\rangle| \\
&= \sup_{\|f\|_p=1} \|m^\vee * f\|_p = (2\pi)^{n/2}\|m\|_{\mathcal{M}_p}.
\end{aligned}$$

(ii) It has already been established because of $\mathcal{M}_1 = \mathcal{M}_\infty$.

(iii) It follows from Theorem 3.49 immediately.

(iv) It follows from the Riesz-Thorin theorem that the mapping $f \mapsto m^\vee * f$ maps $L^{p_0} \to L^{p_0}$ with norm $\|m\|_{\mathcal{M}_{p_0}}$ and $L^{p_1} \to L^{p_1}$ with norm $\|m\|_{\mathcal{M}_{p_1}}$.

(v) Since $1/q = (1-\theta)/p + \theta/p'$ for some θ and $p \leqslant q \leqslant 2 \leqslant p'$, by using (3.25) with $p_0 = p$, $p_1 = p'$, we see that

$$\|m\|_{\mathcal{M}_q} \leqslant \|m\|_{\mathcal{M}_p},$$

from which (3.26) follows. ❑

Proposition 3.52. *Let* $1 \leqslant p \leqslant \infty$. *Then,* $\mathcal{M}_p(\mathbb{R}^n)$ *is a Banach algebra under pointwise multiplication.*

Proof. It is clear that $\|\cdot\|_{\mathcal{M}_p}$ is a norm. Note also that $\mathcal{M}_p$ is complete. Indeed, let $\{m_k\}$ be a Cauchy sequence in $\mathcal{M}_p$. So does it in L^∞ because of $\mathcal{M}_p \subset L^\infty$. Thus, it is convergent in L^∞ and we denote the limit by m. From $L^\infty \subset \mathscr{S}'$, we have $m_k^\vee * f \to m^\vee * f$ for any $f \in \mathscr{S}$ in sense of the strong topology on $\mathscr{S}'$. On the other hand, $\{m_k^\vee * f\}$ 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 $g = m^\vee * f$. Thus, $\|m_k - m\|_{\mathcal{M}_p} \to 0$ as $k \to \infty$. Therefore, $\mathcal{M}_p$ is a Banach space.

Let $m_1 \in \mathcal{M}_p$ and $m_2 \in \mathcal{M}_p$. For any $f \in \mathscr{S}$, we have

$$\begin{aligned}(2\pi)^{-n/2}\| ((m_1 m_2))^\vee * f\|_p =&(2\pi)^{-n}\| (m_1)^\vee * (m_2)^\vee * f\|_p \\ \leqslant&(2\pi)^{-n/2}\|m_1\|_{\mathcal{M}_p}\| (m_2)^\vee * f\|_p \\ \leqslant&\|m_1\|_{\mathcal{M}_p}\|m_2\|_{\mathcal{M}_p}\|f\|_p,\end{aligned}$$

which implies $m_1 m_2 \in \mathcal{M}_p$ and

$$\|m_1 m_2\|_{\mathcal{M}_p} \leqslant \|m_1\|_{\mathcal{M}_p}\|m_2\|_{\mathcal{M}_p}.$$

Thus, $\mathcal{M}_p$ is a Banach algebra. ❑

The next theorem states that $\mathcal{M}_p(\mathbb{R}^n)$ is isometrically invariant under affine transforms[1] of $\mathbb{R}^n$.

Theorem 3.53. *Let $a : \mathbb{R}^n \to \mathbb{R}^k$ be a surjective affine transform with $n \geqslant k$, and $m \in \mathcal{M}_p(\mathbb{R}^k)$. Then*

$$\|m \circ a\|_{\mathcal{M}_p(\mathbb{R}^n)} = \|m\|_{\mathcal{M}_p(\mathbb{R}^k)}.$$

In particular, we have

$$\|\delta^c m\|_{\mathcal{M}_p(\mathbb{R}^n)} =\|m\|_{\mathcal{M}_p(\mathbb{R}^n)}, \quad \forall c > 0, \tag{3.27}$$

$$\|\tilde{m}\|_{\mathcal{M}_p(\mathbb{R}^n)} =\|m\|_{\mathcal{M}_p(\mathbb{R}^n)}, \tag{3.28}$$

$$\|m(\langle x, \cdot\rangle)\|_{\mathcal{M}_p(\mathbb{R}^n)} =\|m\|_{\mathcal{M}_p(\mathbb{R})}, \quad \forall x \neq 0, \tag{3.29}$$

where $\langle x, \xi\rangle = \sum_{i=1}^{n} x_i \xi_i$.

Proof. It suffices to consider the case that $a : \mathbb{R}^n \to \mathbb{R}^k$ is a linear transform. Make the coordinate transform

$$\eta_i = a_i(\xi),\ 1 \leqslant i \leqslant k; \quad \eta_j = \xi_j,\ k+1 \leqslant j \leqslant n, \tag{3.30}$$

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, $\eta' = (\eta_1, \cdots, \eta_k)$ and $\eta'' = (\eta_{k+1}, \cdots, \eta_n)$. It is easy to see, for any $f \in \mathscr{S}(\mathbb{R}^n)$, that

$$\begin{aligned}\mathscr{F}^{-1}(m(a(\xi))\widehat{f})(x) =& \int_{\mathbb{R}^n} e^{ix\cdot\xi} m(a(\xi))\widehat{f}(\xi)\,đ\xi \\ =&|\det A| \int_{\mathbb{R}^n} e^{ix\cdot A\eta} m(\eta')\widehat{f}(A\eta)\,đ\eta \\ =&|\det A| \int_{\mathbb{R}^n} e^{iA^\top x\cdot\eta} m(\eta')\widehat{f}(A\eta)\,đ\eta\end{aligned}$$

[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 $\mathbf{q} \in \mathbb{R}^n$.

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} e^{iA^\top x\cdot\eta} m(\eta') \widehat{f((A^\top)^{-1}\cdot)}(\eta) d\!\!\!{}^-\eta \\
&= \int_{\mathbb{R}^k} e^{i(A^\top x)'\cdot\eta'} m(\eta') \\
&\qquad \left(\int_{\mathbb{R}^{n-k}} e^{i(A^\top x)''\cdot\eta''} \widehat{f((A^\top)^{-1}\cdot)}(\eta',\eta'') d\!\!\!{}^-\eta'' \right) d\!\!\!{}^-\eta' \\
&= \int_{\mathbb{R}^k} e^{i(A^\top x)'\cdot\eta'} m(\eta') \left(\mathscr{F}^{-1}_{\eta''} [\widehat{f((A^\top)^{-1}\cdot)}] \right) (\eta', (A^\top x)'') d\!\!\!{}^-\eta' \\
&= \int_{\mathbb{R}^k} e^{i(A^\top x)'\cdot\eta'} m(\eta') \left([\mathscr{F}_{x'}(f((A^\top)^{-1}\cdot))] \right) (\eta', (A^\top x)'') d\!\!\!{}^-\eta' \\
&= \mathscr{F}^{-1}_{\eta'} \left[m(\eta') \left([\mathscr{F}_{x'}(f((A^\top)^{-1}\cdot))] \right) (\eta', (A^\top x)'') \right] ((A^\top x)') \\
&= \int_{\mathbb{R}^k} m^\vee(y') f((A^\top)^{-1}((A^\top x)' - y', (A^\top x)'')) d\!\!\!{}^-y'.
\end{aligned}
$$

It follows from $m \in \mathcal{M}_p(\mathbb{R}^k)$ that for any $f \in \mathscr{S}(\mathbb{R}^n)$

$$
\begin{aligned}
&(2\pi)^{-np/2} \|\mathscr{F}^{-1}(m(a(\xi))) * f\|^p_{L^p(\mathbb{R}^n)} \\
=& \|\mathscr{F}^{-1}(m(a(\xi))\widehat{f})\|^p_{L^p(\mathbb{R}^n)} \\
=& \left\| \int_{\mathbb{R}^k} m^\vee(y') f((A^\top)^{-1}((A^\top x)' - y', (A^\top x)'')) d\!\!\!{}^-y' \right\|^p_{L^p(\mathbb{R}^n)} \\
=& |\det A|^{-1} \left\| \int_{\mathbb{R}^k} m^\vee(y') f((A^\top)^{-1}(x' - y', x'')) d\!\!\!{}^-y' \right\|^p_{L^p(\mathbb{R}^n)} \\
\leqslant& |\det A|^{-1} \|m\|^p_{\mathcal{M}_p(\mathbb{R}^k)} \left\| \|f((A^\top)^{-1}(x', x''))\|_{L^p(\mathbb{R}^k)} \right\|^p_{L^p(\mathbb{R}^{n-k})} \\
=& |\det A|^{-1} \|m\|^p_{\mathcal{M}_p(\mathbb{R}^k)} \|f((A^\top)^{-1}(x))\|^p_{L^p(\mathbb{R}^n)} \\
=& \|m\|^p_{\mathcal{M}_p(\mathbb{R}^k)} \|f\|^p_{L^p(\mathbb{R}^n)}.
\end{aligned}
$$

Thus, we have

$$
\|m(a(\cdot))\|_{\mathcal{M}_p(\mathbb{R}^n)} \leqslant \|m\|_{\mathcal{M}_p(\mathbb{R}^k)}. \tag{3.31}
$$

Taking $f((A^\top)^{-1}x) = f_1(x')f_2(x'')$, one can conclude that the reverse inequality of (3.31) also holds. ❑

The Fourier multipliers can also be defined on certain vector-valued L^p spaces. We will use results for Fourier multipliers on L^p with values in a Hilbert space. Therefore, we consider only this case. Let $\mathcal{H}$ be a Hilbert space, and consider the space $\mathscr{S}(\mathbb{R}^n, \mathcal{H})$ or $\mathscr{S}(\mathcal{H})$ of all mappings $f : \mathbb{R}^n \to \mathcal{H}$, such that $(1+|x|)^N |\partial^\alpha f(x)|_{\mathcal{H}}$ is bounded for each $\alpha \in \mathbb{N}_0^n$ and $N \in \mathbb{N}_0$. The space $\mathcal{L}(\mathscr{S}(\mathcal{H}_0), \mathcal{H}_1)$ consists of all linear continuous mappings from $\mathscr{S}(\mathcal{H}_0)$ to $\mathcal{H}_1$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ are Hilbert spaces. This space is $\mathscr{S}'$ if $\mathcal{H}_0 = \mathcal{H}_1 = \mathbb{C}$. Clearly, we may define the Fourier transform on $\mathscr{S}(\mathcal{H}_0)$ and on $\mathcal{L}(\mathscr{S}(\mathcal{H}_0), \mathcal{H}_1)$ in the same way as before. The integrals converge in $\mathcal{H}_0$, and it is obvious that the inversion formula holds. We shall also use the notation $\mathscr{S}'(\mathcal{H}_0, \mathcal{H}_1)$ for $\mathcal{L}(\mathscr{S}(\mathcal{H}_0), \mathcal{H}_1)$.

Definition 3.54. Let $\mathcal{H}_0$ and $\mathcal{H}_1$ be two Hilbert spaces with norms $|\cdot|_0$ and $|\cdot|_1$, respectively. Consider a mapping $m \in \mathscr{S}'(\mathcal{H}_0, \mathcal{H}_1)$. We write $m \in \mathcal{M}_p(\mathcal{H}_0, \mathcal{H}_1)$ if for all $f \in \mathscr{S}(\mathcal{H}_0)$ we have $m^\vee * f \in L^p(\mathcal{H}_1)$ and if the expression

$$\sup_{\|f\|_{L^p(\mathcal{H}_0)}=1} \|m^\vee * f\|_{L^p(\mathcal{H}_1)}$$

is finite. The last expression is the norm, $\|m\|_{\mathcal{M}_p(\mathcal{H}_0,\mathcal{H}_1)}$, in $\mathcal{M}_p(\mathcal{H}_0, \mathcal{H}_1)$.

Theorems 3.51 and 3.53 have obvious analogues in this general situation. The proofs are the same with trivial changes. Now, we give a simple but very useful theorem for Fourier multipliers.

Theorem 3.55 (Bernstein multiplier theorem)**.** *Assume that $k > n/2$ is an integer, and that $\partial^\alpha m \in L^2(\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1))$ for $|\alpha| = 0$ and k with nonzero norms. Then, we have $m \in \mathcal{M}_p(\mathcal{H}_0, \mathcal{H}_1)$ for $1 \leqslant p \leqslant \infty$, and*

$$\|m\|_{\mathcal{M}_p} \leqslant C\|m\|_2^{1-n/2k} \Big(\sup_{|\alpha|=k} \|\partial^\alpha m\|_2\Big)^{n/2k}.$$

Proof. Clearly, $m \in \mathscr{S}'(\mathcal{H}_0, \mathcal{H}_1)$. Let $t > 0$. By the Cauchy-Schwarz inequality and the Plancherel theorem, we obtain

$$\begin{aligned}&\int_{|x|>t} |m^\vee(x)|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)}dx\\ =&\int_{|x|>t} |x|^{-k}|x|^k|m^\vee(x)|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)}dx \leqslant Ct^{n/2-k} \sup_{|\alpha|=k} \|\partial^\alpha m\|_{L^2(\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1))}.\end{aligned}$$

Similarly, we have

$$\int_{|x|\leqslant t} |m^\vee(x)|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)}dx \leqslant Ct^{n/2}\|m\|_{L^2(\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1))}.$$

Choosing t such that $\|m\|_2 = t^{-k} \sup_{|\alpha|=k} \|\partial^\alpha m\|_2$, we infer, with the help of Theorem 3.51, that

$$\begin{aligned}\|m\|_{\mathcal{M}_p} \leqslant&\|m\|_{\mathcal{M}_1} = (2\pi)^{-n/2} \int_{\mathbb{R}^n} |m^\vee(x)|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)}dx\\ \leqslant&C\|m\|_2^{1-n/2k} \left(\sup_{|\alpha|=k} \|\partial^\alpha m\|_2\right)^{n/2k}.\end{aligned}$$

❑

Remark 3.56. 1) From the proof of Theorem 3.55, we see that $m^\vee \in L^1$, in other words, it is equivalent to the Young inequality for convolution, i.e., $\|m^\vee * f\|_p \leqslant \|m^\vee\|_1\|f\|_p$ for any $1 \leqslant p \leqslant \infty$.

2) It is not valid if the r.h.s. of the inequality is equal to zero because such a $t \in (0, \infty)$ does not exist in this case in view of the proof. For

example, one can consider $m = \chi_{[-1/2,1/2]} \in L^2$, then $\|\partial m\|_2 = 0$ but $m^\vee(\xi) = (2\pi)^{-1/2}\text{sinc}\frac{\xi}{2} \notin L^1$, where the ***sinc*** function is defined by $\text{sinc}(\xi) = \frac{\sin\xi}{\xi}$ for $\xi \neq 0$ and $\text{sinc}(0) = 1$.

Exercises

Exercise 3.1. Prove Proposition 3.12.

Exercise 3.2. Let

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

Prove the space $(\mathscr{S},\rho)$ is a complete metric space.

Exercise 3.3. [Spi74, Exercise 5.3, 5.4 with answers]

(i) In $\mathbb{R}$, find the Fourier transform of

$$f(x) = \begin{cases} 1, & |x| < a, \\ 0, & |x| > a. \end{cases}$$

(ii) Use the result of (i) to evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{\sin(a\xi)\cos(x\xi)}{\xi}d\xi$.

(iii) Deduce the value of $\displaystyle\int_0^\infty \frac{\sin x}{x}dx$.

Exercise 3.4. For all $a > 0$, prove

$$(2\pi)^{-n/2}\widehat{e^{-a|x|}}(\xi) = \frac{c_n a}{(a^2+|\xi|^2)^{(n+1)/2}}, \quad c_n = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}. \tag{3.32}$$

Exercise 3.5. Let $n = 1$ and

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

It is clear that $g(\xi)$ is uniformly continuous on $\mathbb{R}$ and $g(\xi) \to 0$ as $|\xi| \to \infty$. Prove that there is no integrable function whose Fourier transform is g.

Exercise 3.6 (Hardy-Littlewood-Paley theorem on $\mathbb{R}^n$)**.** Let w be a weight 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

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

Assume $p \in (1,2]$. Use the Marcinkiewicz interpolation theorem to prove

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

Exercise 3.7 (Bernstein's inequality[Gra14a, Exercise 2.3.11])**.** Let f be a bounded function on $\mathbb{R}^n$ with $\widehat{f}$ supported in the ball $B(0,R)$. Prove that for all multi-indices α, there exists a constant $C_{\alpha,n}$ (depending only on α and the dimension n) such that

$$\|\partial^\alpha f\|_\infty \leqslant C_{\alpha,n} R^{|\alpha|} \|f\|_\infty.$$

Hint Write $f = f * h_{1/R}$, where $h \in \mathscr{S}(\mathbb{R}^n)$ whose Fourier transform is equal to one on the ball $B(0,1)$ and vanishes outside the ball $B(0,2)$.

Exercise 3.8 (Homogeneous distributions[Gra14a, Exercise 2.3.9])**.** A distribution in $\mathscr{S}'(\mathbb{R}^n)$ is called homogeneous of degree $\gamma \in \mathbb{C}$ if for all $\lambda > 0$ and for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\langle u, \delta^\lambda \varphi \rangle = \lambda^{-n-\gamma} \langle u, \varphi \rangle.$$

(i) Prove that this definition agrees with the usual definition for functions.
(ii) Show that the Dirac mass δ_0 is homogeneous of degree $-n$.
(iii) Prove that if u is homogeneous of degree γ, then $\partial^\alpha u$ is homogeneous of degree $\gamma - |\alpha|$.
(iv) Show that u is homogeneous of degree γ iff $\widehat{u}$ is homogeneous of degree $-n-\gamma$.

Exercise 3.9. [Gra14a, Exercise 2.5.11] Suppose that $u \in C^\infty(\mathbb{R}^n \setminus \{0\})$ is homogeneous of degree $-n + i\tau$, $\tau \in \mathbb{R}$. Prove that the operator given by convolution with u maps $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$.

4

Hilbert Transform

The study of singular integrals is one of the most important topics in harmonic analysis. The Hilbert transform is the prototypical example of a singular integral. This is a particularly important operator for several reasons: it is a model case for the general theory of singular integral operators; it is a link between real and complex analysis; it is related to summability for Fourier integrals in L^p norms. We will derive its L^p boundedness and the maximal Hilbert transform.

§ 4.1 Hilbert transform

The ***Hilbert transform*** is given formally by the ***principal value integral***

$$Hf(x) := \text{p.v.} \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(x-t)}{t} dt := \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|t|>\varepsilon} \frac{f(x-t)}{t} dt. \tag{4.1}$$

It is not immediately obvious that $Hf(x)$ is well-defined even for nice functions f. We first observe that the ***Hilbert kernel*** $k_H(x) = \text{p.v.} \frac{1}{\pi x} \in \mathscr{S}'(\mathbb{R})$. Indeed, we can write for any $\phi \in \mathscr{S}$

$$\begin{aligned} \left\langle \text{p.v.} \frac{1}{x}, \phi \right\rangle &= \lim_{\varepsilon \to 0} \int_{|x|>\varepsilon} \frac{\phi(x)}{x} dx \\ &= \int_{\varepsilon<|x|<1} \frac{\phi(x) - \phi(0)}{x} dx + \int_{|x|\geqslant 1} \frac{\phi(x)}{x} dx, \end{aligned}$$

this holds since the integral of $1/x$ on $\varepsilon < |x| < 1$ is zero. It is now immediate that for any $\phi \in \mathscr{S}(\mathbb{R})$

$$\left| \left\langle \text{p.v.} \frac{1}{x}, \phi \right\rangle \right| \leqslant 2(\|\phi'\|_\infty + \|x\phi\|_\infty),$$

which implies $k_H \in \mathscr{S}'(\mathbb{R})$. Thus, for $f \in \mathscr{S}(\mathbb{R})$, it follows from Theorem 3.37 that $Hf = k_H * f$ is a $\mathcal{C}^\infty$ function, that is, Hf is also well-defined for $f \in \mathscr{S}(\mathbb{R})$, although it does not map $\mathscr{S}$ to itself.

For $f \in \mathscr{S}(\mathbb{R})$, by symmetry and Proposition 3.10, we have for $|x| \geqslant 2$

$$\begin{aligned} &\left| \pi x Hf(x) - \int_{\mathbb{R}} f \right| \\ =& \left| \lim_{\varepsilon \to 0} \int_{\varepsilon<|t|<1} \frac{xf(x-t) - xf(x)}{t} dt \right. \end{aligned}$$

$$
\begin{aligned}
&+\int_{|t|\geqslant 1}\frac{xf(x-t)}{t}dt-\int_{\mathbb{R}}f(x-t)dt\Bigg|\\
=&\left|\lim_{\varepsilon\to 0}\int_{\varepsilon<|t|<1}\frac{(x-t)f(x-t)-xf(x)}{t}dt+\int_{|t|\geqslant 1}\frac{(x-t)f(x-t)}{t}dt\right|\\
\leqslant&\int_{|t|<1}\int_0^1|(xf)'(x-\theta t)|d\theta dt+C\int_{|t|\geqslant 1}\frac{dt}{|t|(1+|x-t|)^3}\\
\leqslant& C\int_{|t|<1}\int_0^1\frac{d\theta dt}{(1+|x-\theta t|)}+C\int_{|t|\geqslant 1}\frac{dt}{|x|(1+|x-t|)^2}\\
\leqslant& C\int_{|t|<1}\int_0^1\frac{d\theta dt}{(1+|x|/2)}+\frac{C}{|x|}\int_{\mathbb{R}}\frac{dt}{(1+|x-t|)^2}\\
\leqslant&\frac{C}{|x|},
\end{aligned}
$$

since $|x-\theta t|\geqslant |x|-|t|\geqslant |x|/2$ for $|t|<1$ and $|t|(1+|x-t|)\geqslant |t|+|x-t|\geqslant |x|$ for $|t|\geqslant 1$. Thus, for $f\in\mathscr{S}$, we obtain the asymptotic

$$\lim_{|x|\to\infty} xHf(x)=\frac{1}{\pi}\int_{\mathbb{R}}f. \tag{4.2}$$

Hence, if f has nonzero mean then Hf only decays like $\frac{1}{|x|}$ at infinity. In particular, we already see that H is not bounded on L^1.

If $f\in\mathcal{C}_c^1$, then we can restrict the integral to a compact interval $t\in[-R,R]$ for some large R depending on x and f, and use symmetry to write

$$\int_{|t|>\varepsilon}\frac{f(x-t)}{t}dt=\int_{\varepsilon<|t|<R}\frac{f(x-t)-f(x)}{t}dt.$$

The mean-value theorem then shows that $\frac{f(x-t)-f(x)}{t}$ is uniformly bounded on the interval $t\in[-R,R]$ for fixed x and f, and thus, the limit actually exists from the dominated convergence theorem.

Moreover, from the above arguments, it follows that H at least maps $\mathscr{S}(\mathbb{R})$ to $L^2(\mathbb{R})$. It follows from the definition in (4.1) that H commutes with translations. It is also formally skew-adjoint, i.e., the adjoint operator $H^*=-H$, indeed by symmetry we have for $f,g\in\mathscr{S}$,

$$
\begin{aligned}
\int_{\mathbb{R}}Hf(x)\overline{g(x)}dx=&\int_{\mathbb{R}}\int_{\mathbb{R}}k_H(x-y)f(y)dy\overline{g(x)}dx\\
=&-\int_{\mathbb{R}}\int_{\mathbb{R}}f(y)k_H(y-x)\overline{g(x)}dxdy\\
=&-\int_{\mathbb{R}}f(y)\overline{Hg(y)}dy.
\end{aligned}
$$

Similarly, the dual operator $H'=-H$, i.e.,

$$\int_{\mathbb{R}}Hf(x)g(x)dx=-\int_{\mathbb{R}}f(y)Hg(y)dy. \tag{4.3}$$

Now, we give an example.

Example 4.1. Consider the characteristic function $\chi_{[a,b]}$ of an interval $[a,b]$. It is a simple calculation to show that

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}. \tag{4.4}$$

Let us verify this identity. By the definition, we have

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{|y|>\varepsilon} \frac{\chi_{[a,b]}(x-y)}{y} dy = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\substack{|y|>\varepsilon \\ x-b \leqslant y \leqslant x-a}} \frac{1}{y} dy.$$

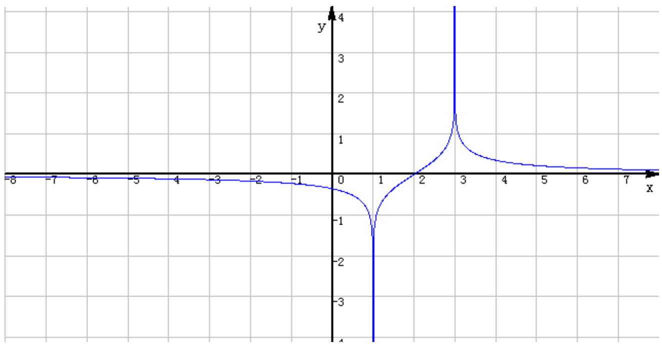
It is clear that it will be $-\infty$ and $+\infty$ at $x = a$ and $x = b$, respectively. Thus, we only need to consider three cases: $x - b > 0$, $x - a < 0$ and $x - b < 0 < x - a$. For the first two cases, we have

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \int_{x-b}^{x-a} \frac{1}{y} dy = \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}.$$

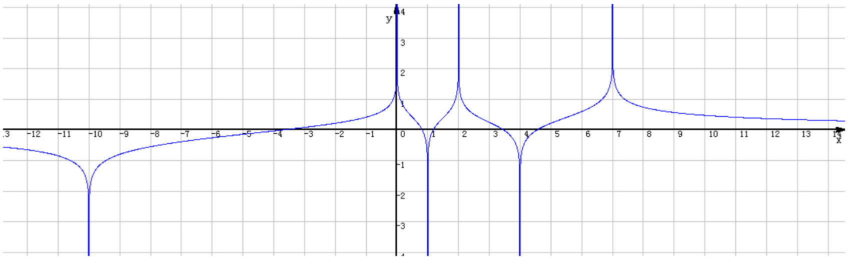
For the third case, we obtain (without loss of generality, we can assume $\varepsilon < \min(|x-a|, |x-b|)$)

$$\begin{aligned} H(\chi_{[a,b]})(x) =& \frac{1}{\pi} \lim_{\varepsilon \to 0} \left(\int_{x-b}^{-\varepsilon} \frac{1}{y} dy + \int_{\varepsilon}^{x-a} \frac{1}{y} dy \right) \\ =& \frac{1}{\pi} \lim_{\varepsilon \to 0} \left(\ln \frac{\varepsilon}{|x-b|} + \ln \frac{|x-a|}{\varepsilon} \right) \\ =& \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}, \end{aligned}$$

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 $x \to \pm\infty$ by (4.2) (or a direct computation with the help of L'Hospital's rule). See the following graph with $a = 1$ and $b = 3$:



The following is a graph of the function $H(\chi_{[-10,0]\cup[1,2]\cup[4,7]})$:



The Hilbert transform is connected to complex analysis (and in particular to Cauchy integrals) by the following identities.

Proposition 4.2 (Plemelj formula, [Tao06]). *Let $f \in \mathcal{C}^1(\mathbb{R})$ obey a qualitative decay bound $f(x) = O_f(\langle x \rangle^{-1})$ (say, these conditions are needed just to make Hf to be well-defined) where $\langle x \rangle = \sqrt{1+x^2}$. Then, for any $x \in \mathbb{R}$,*

$$\frac{1}{2\pi i} \lim_{\varepsilon \to 0} \int_{\mathbb{R}} \frac{f(y)}{y-(x \pm i\varepsilon)} dy = \frac{\pm f(x) + iHf(x)}{2}.$$

Proof. By translation invariance, we can take $x = 0$. By taking complex conjugates we may assume that the $\pm$ sign is $+$. Our task is then to show that

$$\lim_{\varepsilon \to 0} \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(y)}{y - i\varepsilon} dy - \frac{1}{2} f(0) - \frac{i}{2\pi} \int_{|y|>\varepsilon} \frac{f(y)}{0-y} dy = 0.$$

Multiplying by $2\pi i$ and taking the change of variables $y = \varepsilon w$, it reduces to showing

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}} f(\varepsilon w) \left(\frac{1}{w-i} - \chi_{\{|w|>1\}} \frac{1}{w} \right) dw - \pi i f(0) = 0.$$

$$\begin{aligned}
&\int_{\mathbb{R}} \left(\frac{1}{w-i} - \chi_{\{|w|>1\}} \frac{1}{w} \right) dw \\
=& \int_{|w|>1} \left(\frac{1}{w-i} - \frac{1}{w} \right) dw + \int_{|w|\leqslant 1} \frac{1}{w-i} dw \\
=& i \int_{|w|>1} \frac{dw}{1+w^2} - \int_{|w|>1} \frac{dw}{w(1+w^2)} \\
&+ \int_{|w|\leqslant 1} \frac{w}{1+w^2} dw + i \int_{|w|\leqslant 1} \frac{dw}{1+w^2} \\
=& i \int_{\mathbb{R}} \frac{dw}{1+w^2} = \pi i.
\end{aligned}$$

Thus, we only need to show

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}} (f(\varepsilon w) - f(0)) \left(\frac{1}{w-i} - \chi_{\{|w|>1\}} \frac{1}{w} \right) dw = 0.$$

Since f is bounded and

$$\begin{aligned}\int_{\mathbb{R}}\left|\frac{1}{w-i}-\chi_{\{|w|>1\}}\frac{1}{w}\right|dw \leqslant& \int_{|w|>1}\left|\frac{1}{w-i}-\frac{1}{w}\right|dw+\int_{|w|\leqslant 1}\left|\frac{1}{w-i}\right|dw\\ =&\int_{|w|>1}\frac{dw}{w\sqrt{1+w^2}}+\int_{|w|\leqslant 1}\frac{dw}{\sqrt{1+w^2}}\\ \leqslant&\int_{|w|>1}\frac{dw}{w^2}+\int_{|w|\leqslant 1}dw\\ =&2+2=4,\end{aligned}$$

the claim follows from the dominated convergence theorem. ❑

Now suppose that f not only obeys the hypotheses of the Plemelj formula but also extends holomorphically to the upper half-plane $\{z\in\mathbb{C}:\operatorname{Im} z\geqslant 0\}$ and obeys the decay bound $f(x)=O_f(\langle x\rangle^{-1})$ in this region. Then, Cauchy's formula gives

$$\frac{1}{2\pi i}\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\frac{f(y)}{y-(x+i\varepsilon)}dy=f(x)$$

and

$$\frac{1}{2\pi i}\lim_{\varepsilon\to 0}\int_{\mathbb{R}}\frac{f(y)}{y-(x-i\varepsilon)}dy=0.$$

Thus, by either of the Plemelj formulae we see that $f(x)+iHf(x)=2f(x)$, i.e., $Hf=-if$ in this case. In particular, comparing real and imaginary parts we conclude that $\operatorname{Im} f=H\operatorname{Re} f$ and $\operatorname{Re} f=-H\operatorname{Im} f$, so $\operatorname{Re} f=-H^2\operatorname{Re} f$. Thus, for reasonably decaying holomorphic functions on the upper half-plane, the real and imaginary parts of the boundary value are connected via the Hilbert transform. In particular, this shows that such functions are uniquely determined by just the real part of the boundary value.

The above discussion also strongly suggests the identity $H^2=-1$. This can be made more manifest by the following Fourier representation of the Hilbert transform.

Proposition 4.3. *If $f\in\mathscr{S}(\mathbb{R})$, then*

$$\widehat{Hf}(\xi)=-i\operatorname{sgn}(\xi)\hat{f}(\xi),\ \text{a.e. }\xi\in\mathbb{R}. \tag{4.5}$$

Proof. Since the Hilbert transform is odd, a symmetry argument allows us to reduce to the case $\xi>0$. Then, it suffices to show that $\frac{-f+iHf}{2}$ has a vanishing Fourier transform in this half-line. Define the ***Cauchy integral operator***

$$C_\varepsilon f(x)=\frac{1}{2\pi i}\int_{\mathbb{R}}\frac{f(y)}{y-(x-i\varepsilon)}dy. \tag{4.6}$$

The Plemelj formulae show that these converge pointwise to $\frac{-f+iHf}{2}$ as

$\varepsilon \to 0$. Since $f \in \mathscr{S}$, it is also not hard to show via the dominated convergence theorem that they also converge in L^2. Thus, by the L^2 boundedness of the Fourier transform, it suffices to show that each of the $C_\varepsilon f$ also has a vanishing Fourier transform on the half-line. Fix $\varepsilon > 0$, we can truncate and define

$$C_{\varepsilon,R}f(x) = \frac{1}{2\pi i}\int_{\mathbb{R}} \frac{f(y)}{y-(x-i\varepsilon)}\chi_{\{|y-x|<R\}}dy.$$

By the dominated convergence theorem, we can show that $C_{\varepsilon,R}f$ converges to $C_\varepsilon f$ in L^2 as $R \to \infty$, so it will suffice to show that the Fourier transforms of $C_{\varepsilon,R}f$ converge pointiwse to zeros as $R \to \infty$ on the half-line $\xi > 0$. From the Fubini theorem, we easily compute

$$\begin{aligned}\widehat{C_{\varepsilon,R}f}(\xi) =& \frac{1}{2\pi i}\int_{\mathbb{R}} e^{-ix\xi}\int_{\mathbb{R}} \frac{f(y)}{y-(x-i\varepsilon)}\chi_{\{|y-x|<R\}}dy\,đx\\ =& \frac{1}{2\pi i}\int_{\mathbb{R}} e^{-iy'\xi}\left(\int_{\mathbb{R}} e^{-i(x-y')\xi}f(x-y')đx\right)\frac{\chi_{\{|y'|<R\}}}{-y'+i\varepsilon}dy'\\ =& \frac{1}{2\pi i}\hat{f}(\xi)\int_{\mathbb{R}}\frac{e^{-iy\xi}}{-(y-i\varepsilon)}\chi_{\{|y|<R\}}dy.\end{aligned}$$

Then, by shifting the contour to the lower semicircle of radius R and then letting $R \to \infty$, we obtain the result. ❑

From this proposition and Plancherel's theorem, we conclude that H is an isometry:

$$\|Hf\|_2 = \|f\|_2, \quad \forall f \in \mathscr{S}(\mathbb{R}). \tag{4.7}$$

Because of this, H has a unique dense extension to $L^2(\mathbb{R})$, and formula (4.5) is valid for all $f \in L^2(\mathbb{R})$.

It is obvious, for the dilation operator δ^ε with $\varepsilon > 0$, by changes of variables ($\varepsilon y \to y$), that

$$\begin{aligned}(H\delta^\varepsilon)f(x) =& \lim_{\sigma\to 0}\frac{1}{\pi}\int_{|y|>\sigma}\frac{f(\varepsilon x-\varepsilon y)}{y}dy\\ =& \lim_{\sigma\to 0}\int_{|y|>\varepsilon\sigma}\frac{f(\varepsilon x-y)}{y}dy = (\delta^\varepsilon H)f(x),\end{aligned}$$

Therefore, $H\delta^\varepsilon = \delta^\varepsilon H$; and it follows obviously that $H\delta^\varepsilon = -\delta^\varepsilon H$ if $\varepsilon < 0$.

These simple considerations of dilation "invariance" and translation invariance characterize the Hilbert transform.

Proposition 4.4 (Characterization of Hilbert transform)**.** *Suppose T is a bounded linear operator on $L^2(\mathbb{R})$ which satisfies the following properties:*

(i) *T commutes with translations;*

(ii) *T commutes with positive dilations;*

(iii) *T anticommutes with the reflections.*

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 3.49, there is a bounded function $m(\xi)$ such that $\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi)$. Assumptions (ii) and (iii) may be written as $T\delta^\varepsilon f = \operatorname{sgn}(\varepsilon)\delta^\varepsilon Tf$ for all $f \in L^2(\mathbb{R})$. By part (vi) in Proposition 3.2, we have

$$\begin{aligned}\widehat{T\delta^\varepsilon f}(\xi) =& m(\xi)\widehat{\delta^\varepsilon f}(\xi) = m(\xi)|\varepsilon|^{-1}\widehat{f}(\xi/\varepsilon),\\ \operatorname{sgn}(\varepsilon)\widehat{\delta^\varepsilon Tf}(\xi) =& \operatorname{sgn}(\varepsilon)|\varepsilon|^{-1}\widehat{Tf}(\xi/\varepsilon) = \operatorname{sgn}(\varepsilon)|\varepsilon|^{-1}m(\xi/\varepsilon)\widehat{f}(\xi/\varepsilon),\end{aligned}$$

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 proven. ❑

§ 4.2 L^p boundedness of Hilbert transform

Kolmogorov's theorem asserts that the Hilbert transform satisfies a weak type $(1,1)$ estimate. He proved this using complex analysis. Here, we shall give the real analysis proof based on the Calderon-Zygmund decomposition since it goes over to dimension $d > 1$. The next theorem shows that the Hilbert transform, now defined for functions in L^2, can be extended to functions in L^p, $1 \leqslant p < \infty$.

Theorem 4.5. *The following assertions hold:*

(i) (Kolmogorov's theorem) *H is of weak type $(1,1)$: for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$*

$$|\{x \in \mathbb{R} : |Hf(x)| > \alpha\}| \leqslant \frac{C}{\alpha}\|f\|_1.$$

(ii) (Riesz's theorem) *H is of type (p,p), $1 < p < \infty$: for $f \in L^p(\mathbb{R}) \cap L^2(\mathbb{R})$*

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

Proof. (i) Fix $\alpha > 0$. From the Calderón-Zygmund decomposition of f at height α (Theorem 2.17), there exist two functions g and b such that $f = g + b$ and

(1) $\|g\|_1 \leqslant \|f\|_1$ and $\|g\|_\infty \leqslant 2\alpha$, thus $g \in L^1 \cap L^2$ and so is b.

(2) $b = \sum\limits_j b_j$, where each b_j is supported in a dyadic interval I_j satisfying $\int_{I_j} b_j(x)dx = 0$ and $\|b_j\|_1 \leqslant 4\alpha|I_j|$. Furthermore, the intervals I_j and I_k have disjoint interiors when $j \neq k$.

(3) $\sum\limits_j |I_j| \leqslant \alpha^{-1}\|f\|_1$.

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, $|\Omega^*| \leqslant 2|\Omega| \leqslant 2\alpha^{-1}\|f\|_1$.

Since $Hf = Hg + Hb$, from parts (iv) and (vi) of Proposition 1.15, (4.7) and (1), we have

$$\begin{aligned}
&(Hf)_*(\alpha) \leqslant (Hg)_*(\alpha/2) + (Hb)_*(\alpha/2)\\
\leqslant&(\alpha/2)^{-2}\int_{\mathbb{R}} |Hg(x)|^2 dx + |\Omega^*| + |\{x \notin \Omega^* : |Hb(x)| > \alpha/2\}|\\
\leqslant&\frac{4}{\alpha^2}\int_{\mathbb{R}} |g(x)|^2 dx + 2\alpha^{-1}\|f\|_1 + 2\alpha^{-1}\int_{\mathbb{R}\setminus\Omega^*} |Hb(x)|dx\\
\leqslant&\frac{8}{\alpha}\int_{\mathbb{R}} |g(x)|dx + \frac{2}{\alpha}\|f\|_1 + \frac{2}{\alpha}\int_{\mathbb{R}\setminus\Omega^*}\sum_j |Hb_j(x)|dx\\
\leqslant&\frac{8}{\alpha}\|f\|_1 + \frac{2}{\alpha}\|f\|_1 + \frac{2}{\alpha}\sum_j\int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx.
\end{aligned}$$

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 $|x-y| \geqslant |I_j|/2$ for $y \in I_j$. Denote the center of I_j by c_j; then, since b_j is mean value zero, we have

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

The last inequality follows from the fact that $|y-c_j| < |I_j|/2$ and $|x-y| > |x-c_j|/2$ due to $|x-c_j| > |I_j|$. The inner integral is bounded by

$$2|I_j|\int_{|I_j|}^{\infty}\frac{1}{r^2}dr = 2|I_j|\frac{1}{|I_j|} = 2.$$

Thus, by (2) and (3),

$$\begin{aligned}
(Hf)_*(\alpha) \leqslant&\frac{10}{\alpha}\|f\|_1 + \frac{4}{\alpha\pi}\sum_j\int_{I_j}|b_j(y)|dy \leqslant \frac{10}{\alpha}\|f\|_1 + \frac{4}{\alpha\pi}\sum_j 4\alpha|I_j|\\
\leqslant&\frac{10}{\alpha}\|f\|_1 + \frac{16}{\pi}\frac{1}{\alpha}\|f\|_1 = \frac{10+16/\pi}{\alpha}\|f\|_1.
\end{aligned}$$

(ii) Since H is of weak type $(1,1)$ and of type $(2,2)$, by the Marcinkiewicz interpolation theorem, we have the strong type (p,p) inequality for $1 < p < 2$. If $p > 2$, we apply the dual estimates with the help of (4.3) and the result for $p' < 2$ (where $1/p + 1/p' = 1$):

$$\|Hf\|_p = \sup_{\|g\|_{p'}\leqslant 1}|\langle Hf, g\rangle| = \sup_{\|g\|_{p'}\leqslant 1}|\langle f, Hg\rangle|$$

$$\leqslant \|f\|_p \sup_{\|g\|_{p'}\leqslant 1} \|Hg\|_{p'} \leqslant C_{p'}\|f\|_p.$$

This completes the proof. ❑

Remark 4.6. i) Recall from the proof of the Marcinkiewicz interpolation theorem that the coefficient is

$$C_p = \begin{cases} \dfrac{2}{p}\left[\dfrac{(10+16/\pi)p}{p-1} + \dfrac{2p}{2-p}\right]^{1/p}, & 1 < p < 2, \\ \dfrac{2(p-1)}{p}\left[(10+16/\pi)p + \dfrac{2p}{p-2}\right]^{1/p'}, & p > 2. \end{cases}$$

Therefore, the constant C_p tends to infinity as p tends to 1 or ∞. More precisely,

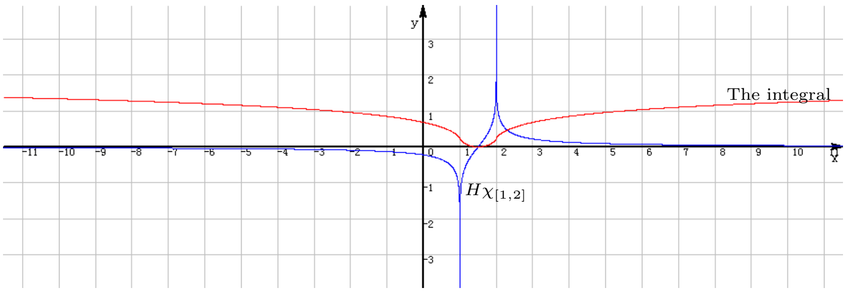
$$C_p = O(p^{1/p'}) \text{ as } p \to \infty, \text{ and } C_p = O((p-1)^{-1/p}) \text{ as } p \to 1.$$

The best constant C_p is given by

$$C_p = \begin{cases} \tan \dfrac{\pi}{2p}, & 1 < p \leqslant 2, \\ \cot \dfrac{\pi}{2p}, & 2 < p < \infty, \end{cases}$$

which is due to [Pic72], see also [Gra14a, Remark 5.1.8].

ii) The strong (p,p) inequality is false if $p = 1$ or $p = \infty$, which can be easily 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 4.5, 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 $\varepsilon > 0$,

$$\lim_{m,n\to\infty} |\{x \in \mathbb{R} : |(Hf_n - Hf_m)(x)| > \varepsilon\}| = 0.$$

Therefore, it converges in measure to a measurable function, which we define as 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 a.e. to Hf as defined.

§ 4.3 The maximal Hilbert transform and L^p boundedness

Definition 4.7. The ***truncated Hilbert transform*** (at height ε) of a function $f \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$, is defined by

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

Observe that $H^{(\varepsilon)}f$ is well-defined for all $f \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$. This follows from Hölder's inequality since $1/x$ is integrable to the power p' on the set $|x| > \varepsilon$.

Clearly, the Hilbert transform of $f \in \mathscr{S}$ can be given by

$$Hf(x) = \lim_{\varepsilon \to 0} H^{(\varepsilon)}f(x). \tag{4.8}$$

We now introduce the maximal Hilbert transform.

Definition 4.8. The ***maximal Hilbert transform*** is the operator

$$H^{(*)}f(x) = \sup_{\varepsilon>0} |H^{(\varepsilon)}f(x)| \tag{4.9}$$

defined for all $f \in L^p$, $1 \leqslant p < \infty$.

Since $H^{(\varepsilon)}f$ is well-defined, $H^{(*)}f$ makes sense for $f \in L^p(\mathbb{R})$, although for some values of x, $H^{(*)}f(x)$ may be infinite.

Example 4.9. Using the result of Example 4.1, we obtain that

$$H^{(*)}\chi_{[a,b]}(x) = \frac{1}{\pi}\left|\ln\frac{|x-a|}{|x-b|}\right| = \left|H\chi_{[a,b]}(x)\right|.$$

However, in general, $H^{(*)}f(x) \neq |Hf(x)|$ by taking f to be the characteristic function of the union of two disjointed closed intervals. (We leave the calculation to the readers.)

The definition of H gives that $H^{(\varepsilon)}f$ converges pointwise to Hf whenever $f \in \mathscr{D}(\mathbb{R})$. If we had the estimate $\|H^{(*)}f\|_p \leqslant C_p\|f\|_p$ for $f \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$, then it follows that $H^{(\varepsilon)}f$ converges to Hf a.e. as $\varepsilon \to 0$ for any $f \in L^p(\mathbb{R})$. This limit a.e. provides a way to describe Hf for general $f \in L^p(\mathbb{R})$. Note that Theorem 4.5 implies only that H has a (unique) bounded extension on L^p, but it does not provide a way to describe Hf when f is a general L^p function.

The next theorem is a simple consequence of these ideas.

Theorem 4.10. *There exists a constant C such that for all $p \in (1, \infty)$, we have*

$$\|H^{(*)} f\|_p \leqslant C \max(p, (p-1)^{-2}) \|f\|_p. \tag{4.10}$$

Moreover, for all $f \in L^p(\mathbb{R})$, $H^{(\varepsilon)} f$ converges to Hf a.e. and in L^p.

Proof. Consider the ***Poisson kernel*** and ***conjugate Poisson kernel*** with $\varepsilon > 0$

$$P_\varepsilon = \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2}, \quad Q_\varepsilon = \frac{1}{\pi} \frac{x}{x^2 + \varepsilon^2},$$

which satisfies the identity with

$$-\frac{1}{i\pi(x + i\varepsilon)} = P_\varepsilon(x) + iQ_\varepsilon(x).$$

Since

$$\int_0^\infty e^{i(x+i\varepsilon)\xi} d\xi = -\frac{1}{i(x + i\varepsilon)},$$

and by Fubini's theorem for $f \in L^2(\mathbb{R})$

$$\begin{aligned}\int_0^\infty e^{i(x+i\varepsilon)\xi} d\xi * f(x) &= \int_{\mathbb{R}} \int_0^\infty e^{i(y+i\varepsilon)\xi} d\xi f(x-y) dy \\ &= \int_0^\infty \int_{\mathbb{R}} e^{i(-y'+x+i\varepsilon)\xi} f(y') dy' d\xi \\ &= (2\pi)^{1/2} \int_0^\infty e^{i(x+i\varepsilon)\xi} \widehat{f}(\xi) d\xi,\end{aligned}$$

it follows that

$$\int_0^\infty e^{i(x+i\varepsilon)\xi} \widehat{f}(\xi) đ\xi = \frac{1}{2}[(P_\varepsilon * f)(x) + i(Q_\varepsilon * f)(x)]. \tag{4.11}$$

Because P_ε and Q_ε are even and odd functions of x, respectively, since $\widetilde{P_\varepsilon * f} = P_\varepsilon * \widetilde{f}$, $\widetilde{Q_\varepsilon * f} = -Q_\varepsilon * \widetilde{f}$ and $\widehat{\widetilde{f}} = \widetilde{\widehat{f}}$, using (4.11) with f and $\widetilde{f}$, we obtain

$$\begin{aligned}(P_\varepsilon * f)(x) &= \int_{\mathbb{R}} \widehat{f}(\xi) e^{ix\xi} e^{-\varepsilon|\xi|} đ\xi = \left(e^{-\varepsilon|\xi|} \widehat{f}\right)^\vee, \\ (Q_\varepsilon * f)(x) &= -i \int_{\mathbb{R}} \widehat{f}(\xi) e^{ix\xi} e^{-\varepsilon|\xi|} \operatorname{sgn}(\xi) đ\xi \\ &= -i \left(\operatorname{sgn}(\xi) e^{-\varepsilon|\xi|} \widehat{f}\right)^\vee.\end{aligned}$$

As a result, we obtain that the Fourier transforms of P_ε and Q_ε (taken in L^2) are given by

$$\widehat{P_\varepsilon}(\xi) = (2\pi)^{-1/2} e^{-\varepsilon|\xi|}, \quad \widehat{Q_\varepsilon}(\xi) = -i \operatorname{sgn}(\xi) (2\pi)^{-1/2} e^{-\varepsilon|\xi|}.$$

It also follows that $\int_{\mathbb{R}^n} P_\varepsilon(x) dx = (2\pi)^{1/2} \widehat{P_\varepsilon}(0) = 1$ and then $\{P_\varepsilon\}_{\varepsilon > 0}$ is an approximate identity.

Thus,

$$\widehat{f * Q_\varepsilon} = -i \operatorname{sgn}(\xi) e^{-\varepsilon|\xi|}\widehat{f} = e^{-\varepsilon|\xi|}\widehat{Hf} = \widehat{P_\varepsilon * Hf},$$

which implies for all $f \in L^2 \cap L^p$

$$f * Q_\varepsilon = Hf * P_\varepsilon, \quad \varepsilon > 0. \tag{4.12}$$

Then, we have

$$H^{(\varepsilon)} f = H^{(\varepsilon)} f - f * Q_\varepsilon + Hf * P_\varepsilon. \tag{4.13}$$

Using the identity

$$\begin{aligned} H^{(\varepsilon)} f(x) - (f * Q_\varepsilon)(x) &= -\frac{1}{\pi}\left[\int_{-\infty}^{\infty} \frac{t f(x-t)}{t^2+\varepsilon^2} dt - \int_{|t|>\varepsilon} \frac{f(x-t)}{t} dt\right] \\ &= -\frac{1}{\pi}\int_{\mathbb{R}} f(x-t)\psi_\varepsilon(t) dt, \end{aligned} \tag{4.14}$$

where $\psi_\varepsilon(x) = \varepsilon^{-1}\psi(\varepsilon^{-1}x)$ and

$$\psi(t) = \begin{cases} \dfrac{t}{t^2+1} - \dfrac{1}{t}, & \text{if } |t| > 1, \\ \dfrac{t}{t^2+1}, & \text{if } |t| \leqslant 1. \end{cases}$$

Note that ψ has an integral of zero since ψ is an odd function and is integrable over the line. Indeed,

$$\begin{aligned} \int_{\mathbb{R}} |\psi(t)| dt &= \int_{|t|>1} \left|\frac{t}{t^2+1} - \frac{1}{t}\right| dt + \int_{|t|\leqslant 1} \frac{|t|}{t^2+1} dt \\ &= \int_{|t|>1} \frac{1}{(t^2+1)|t|} dt + \int_{|t|\leqslant 1} \frac{|t|}{t^2+1} dt \\ &= \int_1^\infty \frac{dt^2}{(t^2+1)t^2} + \int_0^1 \frac{dt^2}{t^2+1} \\ &= \int_1^\infty \frac{ds}{(s+1)s} + \int_0^1 \frac{ds}{s+1} \\ &= \int_1^\infty \left(\frac{1}{s} - \frac{1}{s+1}\right) ds + \int_0^1 \frac{ds}{s+1} \\ &= \left[\ln\left|\frac{s}{s+1}\right|\right]_1^\infty + [\ln|s+1|]_0^1 \\ &= 2\ln 2. \end{aligned}$$

The least decreasing radial majorant of ψ is

$$\Psi(t) = \sup_{|s|\geqslant|t|} |\psi(s)| = \begin{cases} \dfrac{1}{(t^2+1)|t|}, & \text{if } |t| > 1, \\ \dfrac{1}{2}, & \text{if } |t| \leqslant 1, \end{cases}$$

since the function $g(x) = \frac{x}{x^2+1}$ is increasing for $x \in [0,1]$ and decreasing for $x \in (1, \infty)$. It is easy to see that $\|\Psi\|_1 = \ln 2 + 1$. It follows from

Theorem 2.10 that

$$\sup_{\varepsilon>0} |H^{(\varepsilon)}f(x) - (f * Q_\varepsilon)(x)| \leqslant \frac{\ln 2 + 1}{\pi} Mf(x). \tag{4.15}$$

In view of (4.13) and (4.15), from Theorem 2.10 we obtain for $f \in L^2(\mathbb{R}) \cap L^p(\mathbb{R})$ that

$$\begin{aligned} |H^{(*)}f(x)| &= \sup_{\varepsilon>0} |H^{(\varepsilon)}f(x)| \\ &\leqslant \sup_{\varepsilon>0} |H^{(\varepsilon)}f(x) - (f * Q_\varepsilon)(x)| + \sup_{\varepsilon>0} |Hf * P_\varepsilon| \\ &\leqslant \frac{\ln 2 + 1}{\pi} Mf(x) + M(Hf)(x). \end{aligned}$$

It follows immediately from Theorem 2.6 and Theorem 4.5 that $H^{(*)}$ is L^p bounded with norm at most $C \max(p, (p-1)^{-2})$ since $L^2(\mathbb{R}) \cap L^p(\mathbb{R})$ is dense in $L^p(\mathbb{R})$.

Applying Theorem 2.9 to (4.14), we have $\lim_{\varepsilon\to 0} \|H^{(\varepsilon)}f - (f * Q_\varepsilon)\|_p = 0$ since ψ has an integral of zero. By Theorem 2.9, we also have $\lim_{\varepsilon\to 0} \|Hf * P_\varepsilon - Hf\|_p = 0$. Thus, from (4.13), it follows that $\lim_{\varepsilon\to 0} \|H^{(\varepsilon)}f - Hf\|_p = 0$, and therefore, we also have $H^{(\varepsilon)}f \to Hf$ a.e. as $\varepsilon \to 0$. ❑

Exercises

Exercise 4.1. [Zho99, Exercise 1, on p.143] Let $f \in L^2(\mathbb{R})$, $xf \in L^2(\mathbb{R})$ and $\int_{\mathbb{R}} f(x)dx = 0$. To show that $Hf \in L^1(\mathbb{R})$, where H is the Hilbert transform.

Exercise 4.2. [Gra14a, Exercise 5.1.3]

(i) Calculate the Hilbert transform of the following functions: e^{ix}, $\cos x$, $\sin x$, $\sin(\pi x)/\pi x$, where $x \in \mathbb{R}$.

(ii) Show that the operators given by convolution with the smooth function $\sin t/t$ and the distribution p.v. $\cos t/t$ are bounded on $L^p(\mathbb{R})$ whenever $1 < p < \infty$.

Exercise 4.3. [Gra14a, Exercise 5.1.4] Calculate the distribution function of the Hilbert transform of the characteristic function of a measurable subset E of the real line of finite measure, i.e., $(H\chi_E)_*(\alpha)$.

5

Calderón-Zygmund Singular Integral Operators

In this chapter, we consider singular integrals whose kernels have the same essential properties as the kernel of the Hilbert transform.

§ 5.1 Calderón-Zygmund singular integrals

We can generalize Theorem 4.5 to obtain the following result.

Theorem 5.1 (Calderón-Zygmund Theorem). *Let K be a tempered distribution in $\mathbb{R}^n$ that coincides with a locally integrable function on $\mathbb{R}^n \setminus \{0\}$ and satisfies*

$$|\widehat{K}(\xi)| \leqslant (2\pi)^{-n/2}B, \tag{5.1}$$

and the **Hörmander condition**

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \quad y \in \mathbb{R}^n. \tag{5.2}$$

Then we have the strong type (p, p) estimate for $1 < p < \infty$

$$\|K * f\|_p \leqslant C_p \|f\|_p, \tag{5.3}$$

and the weak type $(1, 1)$ estimate

$$(K * f)_*(\alpha) \leqslant \frac{C}{\alpha}\|f\|_1. \tag{5.4}$$

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. The Hörmander condition (5.2) is often deduced from another stronger condition as follows.

Proposition 5.2. *The Hörmander condition (5.2) holds if*

$$|\nabla K(x)| \leqslant \frac{C}{|x|^{n+1}}, \forall x \neq 0. \qquad (\textbf{gradient condition}) \tag{5.5}$$

Proof. By the integral mean value theorem and (5.5), we have

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant \int_{|x|\geqslant 2|y|} \int_0^1 |\nabla K(x - \theta y)||y|d\theta dx$$

$$\leqslant \int_0^1 \int_{|x|\geqslant 2|y|} \frac{C|y|}{|x-\theta y|^{n+1}} dx d\theta \leqslant \int_0^1 \int_{|x|\geqslant 2|y|} \frac{C|y|}{(|x|/2)^{n+1}} dx d\theta$$
$$\leqslant 2^{n+1} C|y|\omega_{n-1} \int_{2|y|}^{\infty} \frac{1}{r^2} dr = 2^{n+1} C|y|\omega_{n-1} \frac{1}{2|y|} = 2^n C\omega_{n-1}.$$

This completes the proof. ❑

Proof of Theorem 5.1. Let $f \in \mathscr{S}$ and $Tf = K * f$. From (5.1), it follows that

$$\begin{aligned} \|Tf\|_2 =&\|\widehat{Tf}\|_2 = (2\pi)^{n/2}\|\widehat{K}\widehat{f}\|_2 \\ \leqslant&(2\pi)^{n/2}\|\widehat{K}\|_\infty \|\widehat{f}\|_2 \leqslant B\|\widehat{f}\|_2 \\ =&B\|f\|_2, \end{aligned} \tag{5.6}$$

by the Plancherel theorem (Theorem 3.16) and part (vii) in Proposition 3.2.

It suffices 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 (5.1) and (5.2). 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 from the Calderón-Zygmund decomposition of f at height α, then as in Theorem 4.5, we can write $f = g + b$ satisfying

(i) $\|g\|_1 \leqslant \|f\|_1$ and $\|g\|_\infty \leqslant 2^n\alpha$;

(ii) $b = \sum\limits_j b_j$, where each b_j is supported in a dyadic cube Q_j satisfying

$$\int_{Q_j} b_j(x)dx = 0 \text{ and } \|b_j\|_1 \leqslant 2^{n+1}\alpha|Q_j|,$$

where the cubes Q_j and Q_k have disjoint interiors when $j \neq k$;

(iii) $\sum\limits_j |Q_j| \leqslant \alpha^{-1}\|f\|_1$.

The argument now proceeds as in Theorem 4.5, and the proof reduces to showing that

$$\int_{\mathbb{R}^n\setminus Q_j^*} |Tb_j(x)|dx \leqslant C\int_{Q_j} |b_j(x)|dx, \tag{5.7}$$

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 (5.7) follows from the Hörmander condition (5.2). Indeed, since each b_j has zero average, we have for $x \notin Q_j^*$

$$Tb_j(x) = \int_{Q_j} K(x-y)b_j(y)dy = \int_{Q_j} [K(x-y) - K(x-c_j)]b_j(y)dy,$$

then,

$$\int_{\mathbb{R}^n \setminus Q_j^*} |Tb_j(x)|dx \leqslant \int_{Q_j} \left(\int_{\mathbb{R}^n \setminus Q_j^*} |K(x-y) - K(x-c_j)|dx \right) |b_j(y)|dy.$$

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 (5.2), we obtain

$$\int_{\mathbb{R}^n \setminus Q_j^*} |K(x-y) - K(x-c_j)|dx \leqslant \int_{|x'| \geqslant 2|y'|} |K(x'-y') - K(x')|dx' \leqslant B.$$

Since the remainder proof is (essentially) a repetition of the proof of Theorem 4.5, we omit the details and complete the proof. ❑

There is still an element that may be considered unsatisfactory in our formulation 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) The results do not directly treat the "principal-value" singular integrals, which exist because of the cancellation of positive and negative values. However, from what we have done, it is now a relatively simple matter to obtain a theorem that covers the cases of interest.

Definition 5.3. Suppose that $K \in L^1_{loc}(\mathbb{R}^n \setminus \{0\})$ satisfies the following conditions:

$$\begin{aligned} &|K(x)| \leqslant B|x|^{-n}, \quad \forall x \neq 0, \\ &\int_{|x| \geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \quad \forall y \neq 0, \end{aligned} \tag{5.8}$$

and

$$\int_{R_1 < |x| < R_2} K(x)dx = 0, \quad \forall 0 < R_1 < R_2 < \infty. \tag{5.9}$$

Then, K is called the ***Calderón-Zygmund kernel***, where B is a constant independent of x and y.

For L^2 boundedness, we have the following lemma.

Lemma 5.4. *Suppose that K satisfies conditions (5.8) and (5.9) of the above* *definition with bound B. Let*

$$K_\varepsilon(x) = \begin{cases} K(x), & |x| \geqslant \varepsilon, \\ 0, & |x| < \varepsilon. \end{cases}$$

Then, we have the estimate

$$\sup_\xi |\widehat{K_\varepsilon}(\xi)| \leqslant (2\pi)^{-n/2} CB, \quad \varepsilon > 0, \tag{5.10}$$

where C depends only on the dimension n.

Proof. First, we prove inequality (5.10) for the special case $\varepsilon = 1$. Since $\widehat{K_1}(0) = 0$, we can assume $\xi \neq 0$ and have

$$\begin{aligned} \widehat{K_1}(\xi) =& \int_{\mathbb{R}^n} e^{-ix\cdot\xi} K_1(x) đx \\ =& \int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x) đx \\ &+ \int_{2\pi/|\xi| \leqslant |x|} e^{-ix\cdot\xi} K_1(x) đx \\ =:& I_1 + I_2. \end{aligned}$$

By condition (5.9), $\int_{1<|x|<2\pi/|\xi|} K(x)dx = 0$ for $|\xi| < 2\pi$, which implies

$$\int_{|x|<2\pi/|\xi|} K_1(x)dx = 0, \quad \forall \xi \neq 0.$$

Thus, $\int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x) đx = \int_{|x|<2\pi/|\xi|} [e^{-ix\cdot\xi} - 1] K_1(x) đx$. Hence, from the fact $|e^{i\theta} - 1| \leqslant |\theta|$ and the first condition in (5.8), we obtain

$$\begin{aligned} |I_1| \leqslant& \int_{|x|<2\pi/|\xi|} |x||\xi||K_1(x)| đx \leqslant B|\xi| \int_{|x|<2\pi/|\xi|} |x|^{-n+1} đx \\ =& (2\pi)^{-n/2} \omega_{n-1} B|\xi| \int_0^{2\pi/|\xi|} dr = (2\pi)^{1-n/2} \omega_{n-1} B. \end{aligned}$$

To estimate I_2, choose $z = z(\xi)$ such that $e^{-i\xi\cdot z} = -1$. This choice can be realized if $z = \pi\xi/|\xi|^2$, with $|z| = \pi/|\xi|$. By changing variables $x + z = y$, we obtain

$$\begin{aligned} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} K_1(x) đx =& -\int_{\mathbb{R}^n} e^{-i(x+z)\cdot\xi} K_1(x) đx = -\int_{\mathbb{R}^n} e^{-iy\cdot\xi} K_1(y-z) đy \\ =& -\int_{\mathbb{R}^n} e^{-ix\cdot\xi} K_1(x-z) đx, \end{aligned}$$

which implies $\int_{\mathbb{R}^n} e^{-ix\cdot\xi} K_1(x) đx = \frac{1}{2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} [K_1(x) - K_1(x-z)] đx$, then we have

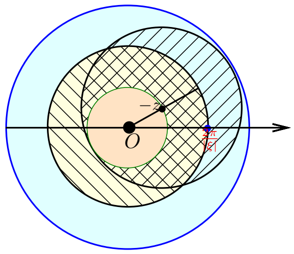
$$\begin{aligned} I_2 =& \left(\int_{\mathbb{R}^n} - \int_{|x|<2\pi/|\xi|} \right) e^{-ix\cdot\xi} K_1(x) đx \\ =& \frac{1}{2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} [K_1(x) - K_1(x-z)] đx \end{aligned}$$

$$
\begin{aligned}
&-\int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x) \bar{d}x \\
=&\frac{1}{2}\lim_{R\to\infty}\int_{2\pi/|\xi|\leqslant|x|\leqslant R} e^{-ix\cdot\xi}[K_1(x)-K_1(x-z)]\bar{d}x \\
&-\frac{1}{2}\int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x) \bar{d}x \\
&-\frac{1}{2}\int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x-z) \bar{d}x.
\end{aligned}
$$

The last two integrals are equal to, in view of integration by parts,

$$
\begin{aligned}
&-\frac{1}{2}\int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x)\bar{d}x - \frac{1}{2}\int_{|y+z|<2\pi/|\xi|} e^{-i(y+z)\cdot\xi} K_1(y)\bar{d}y \\
=&-\frac{1}{2}\int_{|x|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x)\bar{d}x + \frac{1}{2}\int_{|x+z|<2\pi/|\xi|} e^{-ix\cdot\xi} K_1(x)\bar{d}x \\
=&-\frac{1}{2}\int_{|x|<2\pi/|\xi|\leqslant|x+z|} e^{-ix\cdot\xi} K_1(x)\bar{d}x \\
&+\frac{1}{2}\int_{|x+z|<2\pi/|\xi|\leqslant|x|} e^{-ix\cdot\xi} K_1(x)\bar{d}x.
\end{aligned}
$$

For the first integral, we have $2\pi/|\xi| > |x| \geqslant |x+z| - |z| > 2\pi/|\xi| - \pi/|\xi| = \pi/|\xi|$, and for the second one, $2\pi/|\xi| \leqslant |x| \leqslant |x+z| + |z| < 3\pi/|\xi|$. These two integrals are taken over a region contained in the spherical shell, $\pi/|\xi| < |x| < 3\pi/|\xi|$ (see the figure), and are bounded by $\frac{1}{2}B\omega_{n-1}\ln 3$ since $|K_1(x)| \leqslant B|x|^{-n}$. By $|z| = \pi/|\xi|$ and condition (5.8), the first integral of I_2 is majorized by



$$
\begin{aligned}
&\frac{1}{2}\int_{|x|\geqslant 2\pi/|\xi|} |K_1(x-z)-K_1(x)|dx \\
=&\frac{1}{2}\int_{|x|\geqslant 2|z|} |K_1(x-z)-K_1(x)|dx \leqslant \frac{1}{2}B.
\end{aligned}
$$

Thus, we have obtained

$$
|\widehat{K_1}(\xi)| \leqslant (2\pi)^{-n/2}\left(2\pi\omega_{n-1}B + \frac{1}{2}B + \frac{1}{2}B\omega_{n-1}\ln 3\right) \leqslant C_n(2\pi)^{-n/2}B,
$$

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 $\varepsilon^{-n}K(\varepsilon^{-1}x)$. ***Note that if K satisfies the assumptions of the 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 theorems in this chapter.) Now, with a given K, let $K' = \varepsilon^n K(\varepsilon x)$. Then, K' satisfies the conditions of the 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 $|\widehat{K_1'}(\xi)| \leqslant (2\pi)^{-n/2}CB$. The Fourier transform of $\varepsilon^{-n}K_1'(\varepsilon^{-1}x)$ is $\widehat{K_1'}(\varepsilon\xi)$ which is again bounded by $(2\pi)^{-n/2}CB$; however, $\varepsilon^{-n}K_1'(\varepsilon^{-1}x) = K_\varepsilon(x)$; therefore, the lemma is completely proved. ❑

Theorem 5.5. *Suppose that K is a Calderón-Zygmund kernel. For $\varepsilon > 0$ and $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, let*

$$T^{(\varepsilon)} f(x) = \int_{|y| \geqslant \varepsilon} f(x-y) K(y) dy. \tag{5.11}$$

Then, the following conclusions hold:

(i) *We have*

$$\|T^{(\varepsilon)} f\|_p \leqslant A_p \|f\|_p \tag{5.12}$$

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 the 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* $f \in L^p(\mathbb{R}^n)$.

Proof. Since K satisfies conditions (5.8) and (5.9), then $K_\varepsilon(x)$ satisfies the same conditions with bounds not greater than CB. By Lemma 5.4 and Theorem 5.1, we have that the L^p boundedness of the operators $\{K_\varepsilon\}_{\varepsilon>0}$ is uniform. Thus, (i) holds.

Next, we prove that $\{T^{(\varepsilon)} f_1\}_{\varepsilon>0}$ is a Cauchy sequence in L^p provided $f_1 \in \mathcal{C}_c^1(\mathbb{R}^n)$. In fact, we have

$$T^{(\varepsilon)} f_1(x) - T^{(\eta)} f_1(x) = \int_{|y|\geqslant\varepsilon} K(y) f_1(x-y) dy - \int_{|y|\geqslant\eta} K(y) f_1(x-y) dy$$

$$= \operatorname{sgn}(\eta - \varepsilon) \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} K(y)[f_1(x-y) - f_1(x)]dy,$$

because of the cancellation condition (5.9). For $p \in (1, \infty)$, we obtain, by the mean value theorem with some $\theta \in [0, 1]$, Minkowski's inequality and (5.8), 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 (5.12) for f_2 to obtain $\|T^{(\varepsilon)} f_2\|_p \leqslant C\|f_2\|_p$; then, we see that $\lim\limits_{\varepsilon\to 0} T^{(\varepsilon)} f$ exists in the L^p norm; and that the limiting operator T also satisfies the inequality (5.12). Thus, we complete the proof. ❑

Remark 5.6. 1) The linear operator T defined by (ii) of Theorem 5.5 is called the ***Calderón-Zygmund singular integral operator***. $T^{(\varepsilon)}$ is also called the ***truncated operator*** of T.

2) The cancellation property alluded to is contained in condition (5.9). This hypothesis, together with (5.8), allows us to prove the L^2 boundedness and the L^p convergence of the truncated integrals (5.12).

3) We should note that the kernel $K(x) = \frac{1}{\pi x}$, $x \in \mathbb{R}$ clearly satisfies the hypotheses of Theorem 5.5. Therefore, the Hilbert transform exists 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 shown in Theorem 4.5.

§ 5.2 The method of rotations and singular integral with odd kernels

We first introduce the homogeneous singular integrals.

Definition 5.7. Let $\Omega \in L^1(\mathbb{S}^{n-1})$ have a mean value of zero. For $0 < \varepsilon < N < \infty$ and $f \in \bigcup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$, we define the ***truncated singular integral***

$$T_\Omega^{(\varepsilon,N)} f(x) = \int_{\varepsilon\leqslant|y|\leqslant N} f(x-y)\frac{\Omega(y/|y|)}{|y|^n}dy. \tag{5.13}$$

Definition 5.8. Let $\Omega \in L^1(\mathbb{S}^{n-1})$ have a mean value of zero. We denote by T_Ω the ***singular integral operator*** whose kernel is p.v. $\frac{\Omega(x/|x|)}{|x|^n}$, i.e., for $f \in \mathscr{S}(\mathbb{R}^n)$

$$T_\Omega f(x) = \text{p.v.} \frac{\Omega(\cdot/|\cdot|)}{|\cdot|^n} * f(x) = \lim_{\substack{\varepsilon\to0\\ N\to\infty}} T_\Omega^{(\varepsilon,N)} f(x).$$

The associated ***maximal singular integral*** is defined by

$$T_\Omega^{(**)} f = \sup_{0<N<\infty} \sup_{0<\varepsilon<N} |T_\Omega^{(\varepsilon,N)} f|. \tag{5.14}$$

We note that if Ω is bounded, there is no need to use the upper truncation in the definition of $T_\Omega^{(\varepsilon,N)}$ given in (5.13). In this case, the maximal singular integrals can be defined as

$$T_\Omega^{(*)} f = \sup_{\varepsilon>0} |T_\Omega^{(\varepsilon)} f|, \tag{5.15}$$

where for $f \in \bigcup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$, $\varepsilon > 0$ and $x \in \mathbb{R}^n$, $T_\Omega^{(\varepsilon)} f(x)$ is defined in terms of absolutely convergent integral

$$T_\Omega^{(\varepsilon)} f(x) = \int_{|y|\geqslant\varepsilon} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy.$$

To examine the relationship between $T_\Omega^{(*)}$ and $T_\Omega^{(**)}$ for $\Omega \in L^\infty(\mathbb{S}^{n-1})$, notice that

$$\left|T_\Omega^{(\varepsilon,N)} f(x)\right| \leqslant \sup_{0<N<\infty} \left|T_\Omega^{(\varepsilon,N)} f(x)\right|. \tag{5.16}$$

Then, for $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, we let $N \to \infty$ on the l.h.s. in (5.16), and we note that the limit exists in view of the absolute convergence of the integral, which is $|T_\Omega^{(\varepsilon)} f(x)|$. Then we take the supremum over $\varepsilon > 0$ to deduce that $T_\Omega^{(*)}$ is pointwise bounded by $T_\Omega^{(**)}$. Since $T_\Omega^{(\varepsilon,N)} = T_\Omega^{(\varepsilon)} -$

$T_\Omega^{(N)}$, it also follows that $T_\Omega^{(**)} \leqslant 2T_\Omega^{(*)}$. Thus, $T_\Omega^{(*)}$ and $T_\Omega^{(**)}$ are pointwise comparable when Ω lies in $L^\infty(\mathbb{S}^{n-1})$.

A simple procedure called the ***method of rotations*** plays a crucial role in the study of operators with odd kernels. This method is based on the use of directional Hilbert transforms.

Fix a unit vector $\theta \in \mathbb{R}^n$. For $f \in \mathscr{S}(\mathbb{R}^n)$, let

$$H_\theta f(x) = \frac{1}{\pi}\,\text{p.v.} \int_{-\infty}^{\infty} f(x - t\theta)\frac{dt}{t}. \tag{5.17}$$

We call $H_\theta f$ the ***directional Hilbert transform*** of f in the direction θ. For functions $f \in \mathscr{S}(\mathbb{R}^n)$, the integral in (5.17) is well-defined since it converges rapidly at infinity, and by subtracting $f(x)$, it also converges near zero.

Now, we define the ***directional maximal Hilbert transforms***. For a function $f \in \bigcup\limits_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$ and $0 < \varepsilon < N < \infty$, let

$$\begin{aligned} H_\theta^{(\varepsilon,N)} f(x) =& \frac{1}{\pi}\int_{\varepsilon\leqslant|t|\leqslant N} f(x - t\theta)\frac{dt}{t}, \\ H_\theta^{(**)} f(x) =& \sup_{0<\varepsilon<N<\infty} \left|H_\theta^{(\varepsilon,N)} f(x)\right|. \end{aligned}$$

We observe that for any fixed $0 < \varepsilon < N < \infty$ and $f \in L^p(\mathbb{R}^n)$, $H_\theta^{(\varepsilon,N)} f$ is well-defined a.e. Indeed, by Minkowski's integral inequality, we obtain

$$\left\|H_\theta^{(\varepsilon,N)} f\right\|_{L^p(\mathbb{R}^n)} \leqslant \frac{2}{\pi}\|f\|_{L^p(\mathbb{R}^n)} \ln\frac{N}{\varepsilon} < \infty,$$

which implies that $H_\theta^{(\varepsilon,N)} f(x)$ is finite a.e. Thus, $H_\theta^{(**)} f$ is well-defined for $f \in \bigcup\limits_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$.

Note that for $f \in L^p(\mathbb{R}^n)$, we have by Minkowski's inequality

$$\begin{aligned} \|T_\Omega^{(\varepsilon,N)} f\|_p \leqslant& \|f\|_p \int_{\varepsilon\leqslant|y|\leqslant N} \frac{|\Omega(y/|y|)|}{|y|^n} dy \\ =& \|f\|_p \int_\varepsilon^N \int_{\mathbb{S}^{n-1}} \frac{|\Omega(y')|}{r^n} r^{n-1} d\sigma(y') dr \\ =& \|f\|_p \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \ln\frac{N}{\varepsilon}, \end{aligned}$$

which implies that (5.13) is finite a.e. and therefore well-defined a.e.

Theorem 5.9. *If Ω is odd and integrable over $\mathbb{S}^{n-1}$, then T_Ω and $T_\Omega^{(**)}$ are L^p bounded for all $1 < p < \infty$. More precisely, T_Ω initially defined on Schwartz functions has a bounded extension on $L^p(\mathbb{R}^n)$ (which is also denoted by T_Ω).*

Proof. Let e_j be the usual unit vectors in $\mathbb{S}^{n-1}$. The operator H_{e_1} is the directional Hilbert transform in direction e_1. Clearly, H_{e_1} is bounded on $L^p(\mathbb{R}^n)$ with norm bounded by that of the Hilbert transform on $L^p(\mathbb{R})$. Indeed, by Theorem 4.5, we have

$$\begin{aligned}
\|H_{e_1}f\|_{L^p(\mathbb{R}^n)}^p =& \left\| \frac{1}{\pi} \lim_{\varepsilon\to 0} \int_{|t|\geqslant\varepsilon} f(x-te_1)\frac{dt}{t} \right\|_{L^p(\mathbb{R}^n)}^p \\
=& \left\| \frac{1}{\pi} \lim_{\varepsilon\to 0} \int_{|t|\geqslant\varepsilon} f(x_1-t, x_2, \cdots, x_n)\frac{dt}{t} \right\|_{L^p(\mathbb{R}^n)}^p \\
\leqslant& \left\| \|H\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})} \|f(x_1,x')\|_{L^p_{x_1}(\mathbb{R})} \right\|_{L^p_{x'}(\mathbb{R}^{n-1})}^p \\
=& \|H\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})}^p \|f\|_{L^p(\mathbb{R}^n)}^p.
\end{aligned}$$

Next, observe that the following identity is valid for all matrices $A \in O(n)$ (the set of all $n \times n$ orthogonal matrices):

$$\begin{aligned}
H_{Ae_1}f(x) =& \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} f(x-tAe_1)\frac{dt}{t} \\
=& \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} f(A(A^{-1}x-te_1))\frac{dt}{t} \\
=& H_{e_1}(f\circ A)(A^{-1}x). \qquad (5.18)
\end{aligned}$$

This implies that the L^p boundedness of H_θ can be reduced to that of H_{e_1}. We conclude that H_θ is L^p bounded for $1 < p < \infty$ with norm bounded by the norm of the Hilbert transform on $L^p(\mathbb{R})$ for every $\theta \in \mathbb{S}^{n-1}$.

Identity (5.18) is also valid for $H_\theta^{(\varepsilon,N)}$ and $H_\theta^{(**)}$. Consequently, $H_\theta^{(**)}$ is bounded on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$ with norm at most that of $H^{(**)}$ on $L^p(\mathbb{R})$ (or twice the norm of $H^{(*)}$ on $L^p(\mathbb{R})$).

Next, we realize a general singular integral T_Ω with Ω odd as an average of the directional Hilbert transforms H_θ. We start with $f \in \mathscr{S}(\mathbb{R}^n)$ and the following identities:

$$\begin{aligned}
\int_{\varepsilon\leqslant|y|\leqslant N} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy =& \int_{\mathbb{S}^{n-1}} \Omega(\theta) \int_\varepsilon^N f(x-r\theta)\frac{dr}{r} d\sigma(\theta) \\
=& - \int_{\mathbb{S}^{n-1}} \Omega(\theta) \int_\varepsilon^N f(x+r\theta)\frac{dr}{r} d\sigma(\theta) \\
=& \int_{\mathbb{S}^{n-1}} \Omega(\theta) \int_{-N}^{-\varepsilon} f(x-r\theta)\frac{dr}{r} d\sigma(\theta),
\end{aligned}$$

where the first one follows by switching to polar coordinates, the second one is a consequence of the first one and the fact that Ω is odd via the change variables $\theta \mapsto -\theta$, and the third one follows from the second one by changing variables $r \mapsto -r$. Averaging the first and third identities, we obtain

$$\int_{\varepsilon\leqslant|y|\leqslant N} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy$$

$$=\frac{1}{2}\int_{\mathbb{S}^{n-1}}\Omega(\theta)\int_{\varepsilon\leqslant|r|\leqslant N}\frac{f(x-r\theta)}{r}drd\sigma(\theta) \tag{5.19}$$

$$=\frac{\pi}{2}\int_{\mathbb{S}^{n-1}}\Omega(\theta)H_{\theta}^{(\varepsilon,N)}f(x)d\sigma(\theta). \tag{5.20}$$

Since Ω is odd and thus has a mean value of zero, we can obtain

$$\begin{aligned}(5.19)=&\frac{1}{2}\int_{\mathbb{S}^{n-1}}\Omega(\theta)\int_{\varepsilon\leqslant|r|\leqslant 1}\frac{f(x-r\theta)-f(x)}{r}drd\sigma(\theta)\\&+\frac{1}{2}\int_{\mathbb{S}^{n-1}}\Omega(\theta)\int_{1<|r|\leqslant N}\frac{f(x-r\theta)}{r}drd\sigma(\theta).\end{aligned}$$

Because $f\in\mathscr{S}$, the inner integrals are uniformly bounded, so we can apply the dominated convergence theorem to obtain

$$T_{\Omega}f(x)=\frac{\pi}{2}\int_{\mathbb{S}^{n-1}}\Omega(\theta)H_{\theta}f(x)d\sigma(\theta). \tag{5.21}$$

From (5.20), we conclude that

$$T_{\Omega}^{(**)}f(x)\leqslant\frac{\pi}{2}\int_{\mathbb{S}^{n-1}}|\Omega(\theta)|H_{\theta}^{(**)}f(x)d\sigma(\theta). \tag{5.22}$$

The L^p boundedness of T_Ω and $T_\Omega^{(**)}$ for Ω odd are then trivial consequences of (5.22) and (5.21) via Minkowski's integral inequality. ❑

Remark 5.10. It follows from the proof of Theorem 5.9 and from Theorem 4.5 and Theorem 4.10 that whenever Ω is an odd function on $\mathbb{S}^{n-1}$, we have

$$\|T_\Omega\|_{L^p\to L^p},\|T_\Omega^{(**)}\|_{L^p\to L^p}\leqslant\|\Omega\|_1\begin{cases}O(p^{1/p'}), & \text{as } p\to\infty,\\ O((p-1)^{-1/p}), & \text{as } p\to 1.\end{cases}$$

We now define the n ***Riesz transforms***. For $f\in L^p(\mathbb{R}^n)$, $1\leqslant p<\infty$, we set

$$R_j=T_{\Omega_j},\quad j=1,\cdots,n. \tag{5.23}$$

with $\Omega_j(x)=c_n\frac{x_j}{|x|}$ where $c_n=\frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}$. We can also define the maximal Riesz transforms $R_j^{(*)}$ and the maximal singular integral $R_j^{(**)}$ as $T_{\Omega_j}^{(*)}$ and $T_{\Omega_j}^{(**)}$, respectively. Then, we have:

Corollary 5.11. *The Riesz transforms R_j and the maximal Riesz transforms $R_j^{(*)}$ are bounded on $L^p(\mathbb{R}^n)$ for $1<p<\infty$.*

Proof. The assertion follows from the fact that the Riesz transforms have odd kernels. Since the kernel of R_j decays similar to $|x|^{-n}$ near infinity, it follows that $R_j^{(*)}f$ is well-defined for $f\in L^p(\mathbb{R}^n)$. Since $R_j^{(*)}$ is pointwise bounded by $R_j^{(**)}$, and the conclusion follows from Theorem 5.9. ❑

§ 5.3 L^2 boundedness of homogeneous singular integrals

In this section, we compute the Fourier transform of

$$\text{p.v.}\,\Omega(x/|x|)/|x|^n.$$

This provides information on whether the operator T_Ω is L^2 bounded. We have the following result.

Theorem 5.12. *Let $\Omega \in L^1(\mathbb{S}^{n-1})$ have a mean value of zero. Then, the Fourier transform of*

$$(2\pi)^{n/2}\,\text{p.v.}\,\frac{\Omega(x/|x|)}{|x|^n}$$

is a finite a.e. and a homogeneous function of degree 0 given by

$$m(\xi) = \int_{\mathbb{S}^{n-1}} \left[\ln \frac{1}{|\xi \cdot x|} - \frac{\pi i}{2}\,\text{sgn}\,(\xi \cdot x)\right] \Omega(x)d\sigma(x). \tag{5.24}$$

Proof. Since $K(x) = \Omega(x/|x|)/|x|^n$ is not integrable, we first consider its truncated function. Let $0 < \varepsilon < \eta < \infty$, and

$$K_{\varepsilon,\eta}(x) = \begin{cases} \dfrac{\Omega(x/|x|)}{|x|^n}, & \varepsilon \leqslant |x| \leqslant \eta, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $K_{\varepsilon,\eta} \in L^1(\mathbb{R}^n)$. If $f \in L^2(\mathbb{R}^n)$, then $\widehat{K_{\varepsilon,\eta} * f}(\xi) = (2\pi)^{n/2}\widehat{K_{\varepsilon,\eta}}(\xi)\hat{f}(\xi)$.

It is convenient to introduce polar coordinates. Let $x = rx'$, $r = |x|$, $x' = x/|x| \in \mathbb{S}^{n-1}$, and $\xi = R\xi'$, $R = |\xi|$, $\xi' = \xi/|\xi| \in \mathbb{S}^{n-1}$. Then we have

$$\begin{aligned}(2\pi)^{n/2}\widehat{K_{\varepsilon,\eta}}(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} K_{\varepsilon,\eta}(x)dx = \int_{\varepsilon\leqslant|x|\leqslant\eta} e^{-ix\cdot\xi}\frac{\Omega(x/|x|)}{|x|^n}dx \\ &= \int_{\mathbb{S}^{n-1}} \Omega(x') \left(\int_\varepsilon^\eta e^{-iRrx'\cdot\xi'} r^{-n} r^{n-1} dr\right) d\sigma(x') \\ &= \int_{\mathbb{S}^{n-1}} \Omega(x') \left(\int_\varepsilon^\eta e^{-iRrx'\cdot\xi'} \frac{dr}{r}\right) d\sigma(x').\end{aligned}$$

Since

$$\int_{\mathbb{S}^{n-1}} \Omega(x')d\sigma(x') = 0,$$

we can introduce the factor $\cos(Rr)$ (which does not depend on x') in the integral defining $\widehat{K_{\varepsilon,\eta}}(\xi)$. We shall also need the auxiliary integral

$$I_{\varepsilon,\eta}(\xi, x') = \int_\varepsilon^\eta [e^{-iRrx'\cdot\xi'} - \cos(Rr)]\frac{dr}{r}, \quad R > 0.$$

Thus, it follows

$$(2\pi)^{n/2}\widehat{K_{\varepsilon,\eta}}(\xi) = \int_{\mathbb{S}^{n-1}} I_{\varepsilon,\eta}(\xi, x')\Omega(x')d\sigma(x').$$

Now, we first consider $I_{\varepsilon,\eta}(\xi, x')$. For its imaginary part, by changing the variable $Rr(x'\cdot\xi') = t$, we have

$$\begin{aligned}\operatorname{Im} I_{\varepsilon,\eta}(\xi, x') =& -\int_{\varepsilon}^{\eta} \frac{\sin(Rr(x'\cdot\xi'))}{r}dr\\ =& -\operatorname{sgn}(x'\cdot\xi')\int_{R\varepsilon|x'\cdot\xi'|}^{R\eta|x'\cdot\xi'|}\frac{\sin t}{t}dt\end{aligned}$$

is uniformly bounded (i.e., $|\operatorname{Im} I_{\varepsilon,\eta}(\xi, x')| \leqslant 4$) and converges to

$$-\operatorname{sgn}(x'\cdot\xi')\int_0^\infty \frac{\sin t}{t}dt = -\frac{\pi}{2}\operatorname{sgn}(x'\cdot\xi'),$$

as $\varepsilon \to 0$ and $\eta \to \infty$.

For its real part, since $\cos r$ is an even function, we have

$$\operatorname{Re} I_{\varepsilon,\eta}(\xi, x') = \int_{\varepsilon}^{\eta} [\cos(Rr|x'\cdot\xi'|) - \cos(Rr)]\frac{dr}{r}.$$

If $x'\cdot\xi' = \pm 1$, then $\operatorname{Re} I_{\varepsilon,\eta}(\xi, x') = 0$. Next, we assume $x'\cdot\xi' \neq \pm 1$. By the fundamental theorem of calculus, we can write

$$\begin{aligned}&\int_{\varepsilon}^{\eta}\frac{\cos(\lambda r)-\cos(\mu r)}{r}dr = -\int_{\varepsilon}^{\eta}\int_{\mu}^{\lambda}\sin(tr)dtdr = -\int_{\mu}^{\lambda}\int_{\varepsilon}^{\eta}\sin(tr)drdt\\
=&\int_{\mu}^{\lambda}\int_{\varepsilon}^{\eta}\frac{\partial_r\cos(tr)}{t}drdt = \int_{\mu}^{\lambda}\frac{\cos(t\eta)-\cos(t\varepsilon)}{t}dt\\
=&\int_{\mu\eta}^{\lambda\eta}\frac{\cos s}{s}ds - \int_{\mu}^{\lambda}\frac{\cos(t\varepsilon)}{t}dt = \frac{\sin s}{s}\Big|_{\mu\eta}^{\lambda\eta} + \int_{\mu\eta}^{\lambda\eta}\frac{\sin s}{s^2}ds - \int_{\mu}^{\lambda}\frac{\cos(t\varepsilon)}{t}dt\\
\to& 0 - \int_{\mu}^{\lambda}\frac{1}{t}dt = -\ln(\lambda/\mu) = \ln(\mu/\lambda), \text{ as } \eta\to\infty,\ \varepsilon\to 0,\end{aligned}$$

by the dominated convergence theorem with

$$\int_{\mu}^{\lambda}\left|\frac{\cos(t\varepsilon)}{t}\right|dt \leqslant \int_{\mu}^{\lambda}\frac{1}{t}dt = \ln(\lambda/\mu).$$

Take $\lambda = R|x'\cdot\xi'|$, and $\mu = R$. Therefore,

$$\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}} \operatorname{Re}(I_{\varepsilon,\eta}(\xi, x')) = \int_0^\infty [\cos Rr(x'\cdot\xi') - \cos Rr]\frac{dr}{r} = \ln(1/|x'\cdot\xi'|).$$

Next, we need to show that $\widehat{K_{\varepsilon,\eta}}(\xi)$ is finite for a.e. $\xi \in \mathbb{R}^n$. Now we assume $0 < \varepsilon < 1 < \eta$. For the case $x'\cdot\xi' \neq \pm 1$, we obtain from the previous identity

$$\begin{aligned}|\operatorname{Re} I_{\varepsilon,\eta}(\xi, x')| =& \left|\int_{R\eta|\xi'\cdot x'|}^{R\eta}\frac{\cos s}{s}ds - \int_{R|\xi'\cdot x'|}^{R}\frac{\cos\varepsilon s}{s}ds\right|\\
\leqslant& \int_{R\eta|\xi'\cdot x'|}^{R\eta}\frac{1}{s}ds + \int_{R|\xi'\cdot x'|}^{R}\frac{1}{s}ds\end{aligned}$$

$$=2\ln(1/|\xi'\cdot x'|).$$

By the properties of $I_{\varepsilon,\eta}$ just proved, we have

$$(2\pi)^{n/2}|\widehat{K_{\varepsilon,\eta}}(\xi)| \leqslant \int_{\mathbb{S}^{n-1}} \left[4+2\ln(1/|\xi'\cdot x'|)\right] |\Omega(x')|d\sigma(x') \tag{5.25}$$
$$\leqslant 4\|\Omega\|_{L^1(\mathbb{S}^{n-1})} + 2\int_{\mathbb{S}^{n-1}} \ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x').$$

For $n=1$, we have $\mathbb{S}^0=\{-1,1\}$ and then $\int_{\mathbb{S}^0}\ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x')=0$. For $n\geqslant 2$, if we can show

$$\int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} \ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x')d\sigma(\xi')<\infty,$$

then, $\widehat{K_{\varepsilon,\eta}}(\xi)$ is finite a.e. on $\mathbb{S}^{n-1}$. We can select an orthogonal matrix A such that $Ae_1=x'$, and thus, by changing the variables,

$$\begin{aligned}
&\int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} \ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x')d\sigma(\xi')\\
=&\int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} \ln(1/|\xi'\cdot Ae_1|)d\sigma(\xi')|\Omega(x')|d\sigma(x')\\
=&\int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} \ln(1/|e_1\cdot A^{-1}\xi'|)d\sigma(\xi')|\Omega(x')|d\sigma(x')\\
\overset{A^{-1}\xi'=y}{=\!=\!=\!=}&\|\Omega\|_{L^1(\mathbb{S}^{n-1})}\int_{\mathbb{S}^{n-1}} \ln(1/|y_1|)d\sigma(y).
\end{aligned}$$

If for $\phi_j\in[0,\pi]$ $(j=1,\cdots,n-2)$ and $\phi_{n-1}\in[0,2\pi]$, let

$$y_1=\cos\phi_1,\ y_k=\cos\phi_k\prod_{j=1}^{k-1}\sin\phi_j,\ (k=2,\cdots,n-1),\ y_n=\prod_{j=1}^{n-1}\sin\phi_j,$$

then the volume element $d\sigma_{\mathbb{S}^{n-1}}(y)$ of the $(n-1)$-sphere is given by

$$\begin{aligned}
d\sigma_{\mathbb{S}^{n-1}}(y)=&\prod_{j=1}^{n-2}\sin^{n-1-j}\phi_j\,d\phi_1\,d\phi_2\cdots d\phi_{n-1}\\
=&\sin^{n-3}(\phi_1)\,dy_1\prod_{j=2}^{n-2}\sin^{n-1-j}\phi_j\,d\phi_2\cdots d\phi_{n-1}\\
=&(1-y_1^2)^{(n-3)/2}dy_1d\sigma_{\mathbb{S}^{n-2}}(\bar{y}),
\end{aligned}$$

due to $dy_1=\sin(\phi_1)d\phi_1$ and $\sin\phi_1=\sqrt{1-y_1^2}$. Then, we obtain with the notation $\bar{y}=(y_2,y_3,\ldots,y_n)$,

$$\begin{aligned}
&\int_{\mathbb{S}^{n-1}}\ln(1/|y_1|)d\sigma(y)\\
=&\int_{-1}^{1}\ln(1/|y_1|)\int_{\mathbb{S}^{n-2}}(1-y_1^2)^{(n-3)/2}d\sigma(\bar{y})dy_1\\
=&\omega_{n-2}\int_{-1}^{1}\ln(1/|y_1|)(1-y_1^2)^{(n-3)/2}dy_1\\
=&2\omega_{n-2}\int_{0}^{1}\ln(1/|y_1|)(1-y_1^2)^{(n-3)/2}dy_1
\end{aligned}$$

$$
\begin{aligned}
&=2\omega_{n-2}\int_0^{\pi/2}\ln(1/\cos\theta)(\sin\theta)^{n-2}d\theta \quad (\text{let } y_1=\cos\theta)\\
&=2\omega_{n-2}I_2.
\end{aligned}
$$

For $n\geqslant 3$, by integration by parts, we have

$$
I_2\leqslant\int_0^{\pi/2}\ln(1/\cos\theta)\sin\theta d\theta=\int_0^{\pi/2}\sin\theta d\theta=1.
$$

For $n=2$, we have by changing variables

$$
\begin{aligned}
I_2=&\int_0^{\pi/2}\ln(1/\cos\theta)d\theta=-\int_0^{\pi/2}\ln(\cos\theta)d\theta\\
=&-\int_0^{\pi/2}\ln\sin\left(\frac{\pi}{2}-\theta\right)d\theta=-\int_0^{\pi/2}\ln(\sin\theta)d\theta\\
=&-\int_0^{\pi/2}\ln\left(2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right)d\theta\\
=&-\int_0^{\pi/2}\left(\ln 2+\ln\sin\frac{\theta}{2}+\ln\cos\frac{\theta}{2}\right)d\theta\\
=&-\frac{\pi}{2}\ln 2-2\int_0^{\pi/4}\ln\sin x dx-2\int_0^{\pi/4}\ln\cos x dx\\
=&-\frac{\pi}{2}\ln 2-2\int_0^{\pi/4}\ln\sin x dx-2\int_{\pi/4}^{\pi/2}\ln\sin x dx\\
=&-\frac{\pi}{2}\ln 2+2I_2,
\end{aligned}
$$

which yields $I_2=\frac{\pi}{2}\ln 2$.

In view of the limit of $I_{\varepsilon,\eta}(\xi,x')$ as $\varepsilon\to 0$, $\eta\to\infty$ just proved, we obtain

$$
(2\pi)^{n/2}\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}}\widehat{K_{\varepsilon,\eta}}(\xi)=m(\xi),\quad \text{a.e.}
$$

By the Plancherel theorem, if $f\in L^2(\mathbb{R}^n)$, $K_{\varepsilon,\eta}*f$ converges in the L^2 norm as $\varepsilon\to 0$ and $\eta\to\infty$, and the Fourier transform of this limit is $m(\xi)\hat{f}(\xi)$. From the formula of the multiplier $m(\xi)$, it is homogeneous of degree 0 in view of the mean zero property of Ω. Thus, we obtain the conclusion. ❑

Remark 5.13. 1) In the theorem, the condition that Ω is mean value zero on $\mathbb{S}^{n-1}$ is necessary and cannot be neglected. Since in the estimate

$$
\int_{\mathbb{R}^n}\frac{\Omega(y/|y|)}{|y|^n}f(x-y)dy=\left[\int_{|y|\leqslant 1}+\int_{|y|>1}\right]\frac{\Omega(y/|y|)}{|y|^n}f(x-y)dy,
$$

the main difficulty lies in the first integral. For instance, if we assume $\Omega(x)\equiv 1\in L^1(\mathbb{S}^{n-1})$, $f(x)=\chi_{|x|\leqslant 1}(x)\in L^2(\mathbb{R}^n)$, then this integral is divergent for $|x|\leqslant 1/2$ since

$$
\int_{\mathbb{R}^n}\frac{\Omega(y/|y|)}{|y|^n}f(x-y)dy=\int_{|x-y|\leqslant 1}\frac{1}{|y|^n}dy\geqslant\int_{|y|\leqslant 1/2}\frac{1}{|y|^n}dy=\infty.
$$

2) The proof 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., $\operatorname{Im} I_{\varepsilon,\eta}(\xi, x')$, we required only $\int_{\mathbb{S}^{n-1}} |\Omega_o(x')| d\sigma(x') < \infty$ for the odd part; and for the even part, the proof requires the uniform boundedness of

$$\int_{\mathbb{S}^{n-1}} |\Omega_e(x')| \ln(1/|\xi' \cdot x'|) d\sigma(x').$$

In addition, $\ln(1/|\xi' \cdot x'|)$ is not bounded but any power ($\geqslant 1$) of it is integrable. We immediately obtain the following corollary and leave the proof to readers.

Corollary 5.14. *Given a function $\Omega = \Omega_e + \Omega_o$ with mean value zero on $\mathbb{S}^{n-1}$, suppose that the odd part $\Omega_o \in L^1(\mathbb{S}^{n-1})$ and the even part $\Omega_e \in L^q(\mathbb{S}^{n-1})$ for some $q > 1$. Then, the Fourier transform of* p.v. $\Omega(x')/|x|^n$ *is bounded.*

If $\Omega \in L^1(\mathbb{S}^{n-1})$ is odd, i.e., $\Omega(-x) = -\Omega(x)$ for all $x \in \mathbb{S}^{n-1}$, then

$$\int_{\mathbb{S}^{n-1}} \Omega(x) \ln(1/|\xi \cdot x|) d\sigma(x) = 0, \quad \forall \xi \in \mathbb{S}^{n-1}.$$

Thus, $m \in L^\infty(\mathbb{R}^n)$ in view of Theorem 5.12. We have the following result by Theorem 3.49.

Corollary 5.15. *Given an odd function $\Omega \in L^1(\mathbb{S}^{n-1})$, then the singular integral $T_\Omega f(x) :=$ p.v. $\displaystyle\int_{\mathbb{R}^n} \frac{\Omega(y/|y|)}{|y|^n} f(x-y) dy$ is always L^2 bounded.*

§ 5.4 Singular integral operators with Dini-type condition

In this section, we shall consider those operators that commute not only with translations but also with dilations. Among these we shall study the class of singular integral operators, falling under the scope of Theorem 5.5.

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 $\varepsilon^{-n} K(\varepsilon^{-1} x)$. Therefore, 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

$$K(x) = \frac{\Omega(x)}{|x|^n}, \tag{5.26}$$

with Ω homogeneous of degree 0, i.e., $\Omega(\varepsilon x) = \Omega(x)$, $\varepsilon > 0$. This condition on Ω is equivalent to 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 $\mathbb{S}^{n-1}$.

Let us try to reinterpret the conditions of Theorem 5.5 in terms of Ω.

1) By (5.8), $\Omega(x)$ must be bounded and consequently integrable on $\mathbb{S}^{n-1}$; 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 (5.8):

$$\text{if } w(\eta) := \sup_{\substack{|x-x'|\leqslant\eta\\ |x|=|x'|=1}} |\Omega(x)-\Omega(x')|, \quad \text{then } \int_0^1 \frac{w(\eta)}{\eta}d\eta < \infty. \tag{5.27}$$

Of course, any Ω that is of class $\mathcal{C}^1$ or even merely Lipschitz continuous satisfies the condition (5.27).

2) The cancellation condition (5.9) is then the same as the mean value zero of Ω on $\mathbb{S}^{n-1}$.

Theorem 5.16. *Let $\Omega \in L^\infty(\mathbb{S}^{n-1})$ be homogeneous of degree 0 with mean value zero on $\mathbb{S}^{n-1}$, and suppose that Ω satisfies the smoothness property (5.27). For $1 < p < \infty$ and $f \in L^p(\mathbb{R}^n)$, let*

$$T^{(\varepsilon)}f(x) = \int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy. \tag{5.28}$$

(i) *Then, there exists a bound A_p (independent of f and ε) such that*

$$\|T^{(\varepsilon)}f\|_p \leqslant A_p\|f\|_p.$$

(ii) $\lim_{\varepsilon\to 0} T^{(\varepsilon)}f = Tf$ *exists in L^p, and*

$$\|Tf\|_p \leqslant A_p\|f\|_p.$$

(iii) *If $f \in L^2(\mathbb{R}^n)$, then the Fourier multiplier m corresponding to T is a homogeneous function of degree 0 expressed in (5.24).*

Proof. Conclusions (i) and (ii) are immediately consequences of Theorem 5.5, once we have shown that any $K(x)$ of the form $\frac{\Omega(x)}{|x|^n}$ satisfies

$$\int_{|x|\geqslant 2|y|} |K(x-y)-K(x)|dx \leqslant B, \tag{5.29}$$

if Ω is as in condition (5.27). 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

$$\int_{|x|\geqslant 2|y|}\left|\frac{1}{|x-y|^n}-\frac{1}{|x|^n}\right|dx=\int_{|x|\geqslant 2|y|}\left|\frac{|x|^n-|x-y|^n}{|x-y|^n|x|^n}\right|dx$$

$$=\int_{|x|\geqslant 2|y|}\frac{||x|-|x-y||\sum_{j=0}^{n-1}|x|^{n-1-j}|x-y|^j}{|x-y|^n|x|^n}dx$$

$$\leqslant\int_{|x|\geqslant 2|y|}|y|\sum_{j=0}^{n-1}|x|^{-j-1}|x-y|^{j-n}dx$$

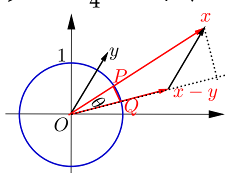
$$\leqslant\int_{|x|\geqslant 2|y|}|y|\sum_{j=0}^{n-1}|x|^{-j-1}(|x|/2)^{j-n}dx\quad(\because |x-y|\geqslant|x|-|y|\geqslant|x|/2)$$

$$=\int_{|x|\geqslant 2|y|}|y|\sum_{j=0}^{n-1}2^{n-j}|x|^{-n-1}dx=2(2^n-1)|y|\int_{|x|\geqslant 2|y|}|x|^{-n-1}dx$$

$$=2(2^n-1)|y|\omega_{n-1}\frac{1}{2|y|}=(2^n-1)\omega_{n-1}.$$

Now, we estimate the first group of terms. Let θ be the angle with sides x and $x-y$ whose opposite side is y in the triangle formed by vectors x, y and $x-y$.

Since $|y|\leqslant|x|/2\leqslant|x|$, we have $\theta\leqslant\frac{\pi}{2}$ and so $\cos\frac{\theta}{2}\geqslant\cos\frac{\pi}{4}=1/\sqrt{2}$. Moreover, by the sine theorem, we have $\sin\theta\leqslant\frac{|y|}{|x|}$. On the other hand, in the triangle formed by $\overrightarrow{OP}:=\frac{x}{|x|}$, $\overrightarrow{OQ}:=\frac{x-y}{|x-y|}$ and $\overrightarrow{PQ}:=\frac{x-y}{|x-y|}-\frac{x}{|x|}$, it is clear that $\theta=\angle POQ$ and $\frac{\sin\theta}{|\overrightarrow{PQ}|}=\frac{\sin\frac{\pi-\theta}{2}}{|\overrightarrow{OP}|}$ by the sine theorem. Then, we have



$$\left|\frac{x-y}{|x-y|}-\frac{x}{|x|}\right|=|\overrightarrow{PQ}|=\frac{\sin\theta}{\sin(\frac{\pi}{2}-\frac{\theta}{2})}=\frac{\sin\theta}{\cos\frac{\theta}{2}}$$

$$\leqslant\sqrt{2}\frac{|y|}{|x|}\leqslant 2\frac{|y|}{|x|}.$$

Thus, the integral corresponding to the first group of terms is dominated by

$$2^n\int_{|x|\geqslant 2|y|}w\left(2\frac{|y|}{|x|}\right)\frac{dx}{|x|^n}=2^n\omega_{n-1}\int_{2|y|}^{\infty}w(2|y|/r)\frac{dr}{r}$$

$$=2^n\omega_{n-1}\int_0^1\frac{w(\eta)d\eta}{\eta}<\infty$$

in view of changes of variables $\eta=2|y|/r$ and the Dini-type condition (5.27).

Part (iii) is the same as the proof of Theorem 5.12 with minor modifications. Indeed, we only need to simplify the proof of (5.25) due to

$\Omega \in L^\infty(\mathbb{S}^{n-1})$. We can control (5.25) by

$$4\omega_{n-1}\|\Omega\|_{L^\infty(\mathbb{S}^{n-1})} + 2\|\Omega\|_{L^\infty(\mathbb{S}^{n-1})} \int_{\mathbb{S}^{n-1}} \ln(1/|\xi' \cdot x'|) d\sigma(x'),$$

where the integral in the last term is equal to

$$\int_{\mathbb{S}^{n-1}} \ln(1/|y_1|) d\sigma(y) \leqslant C_n$$

which have been estimated in Theorem 5.12. Thus, we have completed the proof. ❑

Theorem 5.16 guarantees the existence of the singular integral

$$\lim_{\varepsilon\to 0} \int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy \tag{5.30}$$

in the sense of convergence in the L^p norm. The natural counterpart of this result is that of a.e. convergence. For the questions involving a.e. convergence, it is best to also consider the corresponding maximal function.

Theorem 5.17. *Suppose that Ω satisfies the conditions of Theorem 5.16. 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.)

(i) $\lim_{\varepsilon\to 0} T^{(\varepsilon)} f(x)$ *exists for almost every x.*

(ii) *Let $T^{(*)} f(x) = \sup_{\varepsilon>0} |T^{(\varepsilon)} f(x)|$. If $f \in L^1(\mathbb{R}^n)$, then the mapping $T^{(*)}$ is of weak type $(1,1)$.*

(iii) *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 is the proof of inequality (iii), which can be obtained as a relatively easy consequence of the existence of $\lim_{\varepsilon\to 0} T^{(\varepsilon)}$ in L^p, already proven in Theorem 5.16, and certain general properties of "approximations to the identity" as follows.

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 5.16. We shall prove this part by showing the following ***Cotlar inequality*** for $f \in L^p(\mathbb{R}^n)$ with $p \in (1,\infty)$

$$T^{(*)} f(x) \leqslant M(Tf)(x) + CMf(x).$$

Let $\varphi \in \mathscr{D}(\mathbb{R}^n)$ be a smooth nonnegative decreasing radial function, which is supported in the unit ball with integral one. 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

$$\Phi = \varphi * K - K_1, \tag{5.31}$$

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 Ψ of Φ is integrable (so as to apply Theorem 2.10).

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 the mean value zero of Ω on $\mathbb{S}^{n-1}$ 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 for $\varphi * K$ and the boundedness of K in this case. If $|x| \geqslant 2$, we have

$$\Phi(x) = \int_{\mathbb{R}^n} K(x-y)\varphi(y)dy - K(x) = \int_{|y|\leqslant 1} [K(x-y) - K(x)]\varphi(y)dy.$$

Similar to (5.29), we can obtain the bound for $|y| \leqslant 1$ and thus $|x| \geqslant 2|y|$,

$$\begin{aligned} |K(x-y) - K(x)| \leqslant& 2^n w\left(\frac{2|y|}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_{L^\infty(\mathbb{S}^{n-1})}|y||x|^{-(n+1)} \\ \leqslant& 2^n w\left(\frac{2}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_{L^\infty(\mathbb{S}^{n-1})}|x|^{-(n+1)}, \end{aligned}$$

as in the proof of Theorem 5.16 since w is increasing. Thus, due to $\|\varphi\|_1 = 1$, we obtain for $|x| \geqslant 2$

$$|\Phi(x)| \leqslant 2^n w\left(\frac{2}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_{L^\infty(\mathbb{S}^{n-1})}|x|^{-(n+1)}.$$

Therefore, we obtain $|\Psi| \leqslant C$ for $|x| < 2$, and

$$|\Psi(x)| \leqslant 2^n w\left(\frac{2}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_{L^\infty(\mathbb{S}^{n-1})}|x|^{-(n+1)},$$

for $|x| \geqslant 2$, and then we can prove that $\Psi \in L^1(\mathbb{R}^n)$ with the help of the Dini-type condition.

From (5.31), it follows, because the singular integral operator $T : \varphi \to \varphi * K$ commutes with dilations, that

$$\varphi_\varepsilon * K - K_\varepsilon = \Phi_\varepsilon, \quad \text{with } \Phi_\varepsilon(x) = \varepsilon^{-n}\Phi(x/\varepsilon). \tag{5.32}$$

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

$$(\varphi_\varepsilon * K) * f(x) = Tf * \varphi_\varepsilon(x), \tag{5.33}$$

where the identity holds for every x. In fact, we notice first that

$$(\varphi_\varepsilon * K_\delta) * f(x) = T_\delta f * \varphi_\varepsilon(x), \quad \forall \delta > 0 \tag{5.34}$$

because both sides of (5.34) are equal for each x to the absolutely convergent double integral $\int_{\mathbb{R}^n} \int_{|y| \geqslant \delta} K(y) f(z-y) \varphi_\varepsilon(x-z) dy dz$. Moreover, $\nabla \varphi_\varepsilon \in L^{p'}(\mathbb{R}^n)$, with $1/p + 1/p' = 1$, so $\varphi_\varepsilon * K_\delta \to \varphi_\varepsilon * K$ in $L^{p'}$ since

$$\begin{aligned}
\|\varphi_\varepsilon * (K_\delta - K)\|_{p'} =& \left\| \int_{|y|<\delta} \frac{\Omega(y)}{|y|^n} \varphi_\varepsilon(x-y) dy \right\|_{p'} \\
=& \left\| \int_{|y|<\delta} \frac{\Omega(y)}{|y|^n} (\varphi_\varepsilon(x-y) - \varphi_\varepsilon(x)) dy \right\|_{p'} \\
\leqslant& \left\| \int_{|y|<\delta} \frac{\Omega(y)}{|y|^{n-1}} |\nabla \varphi_\varepsilon(x-\theta y)| dy \right\|_{p'} \\
\leqslant& \delta \omega_{n-1} \|\Omega\|_{L^\infty(\mathbb{S}^{n-1})} \|\nabla \varphi_\varepsilon\|_{p'} \to 0, \text{ as } \delta \to 0,
\end{aligned}$$

by Minkowski's inequality. We also have $T_\delta f \to Tf$ in L^p, as $\delta \to 0$, by Theorem 5.16. It follows (5.33), and so by (5.32)

$$T^{(\varepsilon)} f = K_\varepsilon * f = \varphi_\varepsilon * K * f - \Phi_\varepsilon * f = Tf * \varphi_\varepsilon - f * \Phi_\varepsilon.$$

Passing to the supremum over ε and applying Theorems 2.10, 2.6 and 5.16, we obtain the Cotlar inequality and

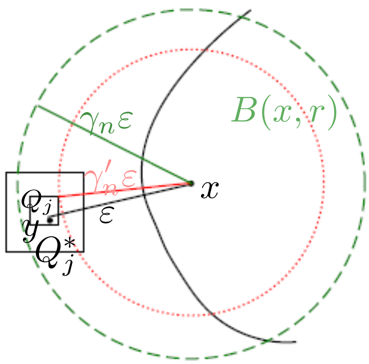
$$\begin{aligned}
\|T^{(*)} f\|_p \leqslant& \|\sup_{\varepsilon>0} |Tf * \varphi_\varepsilon|\|_p + \|\sup_{\varepsilon>0} |f * \Phi_\varepsilon|\|_p \\
\leqslant& C\|M(Tf)\|_p + C\|Mf\|_p \leqslant C\|Tf\|_p + C\|f\|_p \leqslant C\|f\|_p.
\end{aligned}$$

Thus, we have proved (iii).

The second and most difficult stage of the proof is conclusion (ii). Here, the argument proceeds in the main as in the proof of the weak type (1,1) result for singular integrals in Theorem 5.1. We review it with deliberate brevity to avoid a repetition of details already examined.

For a given $\alpha > 0$, we split $f = g + b$ as in the proof of Theorem 5.1. It is easy to check the part for g with the help of the Cotlar inequality. Therefore, we only consider the part for b.

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.



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 $x \in \mathbb{R}^n \setminus \cup_j Q_j^*$ and assume that for some $y \in Q_j$, $|x-y| = \varepsilon$. Then, we have $(\sqrt{n} - \frac{1}{2})\ell(Q) \leqslant \varepsilon \leqslant \frac{\sqrt{n}}{2}(2\sqrt{n}+1)\ell(Q_j) = (n + \frac{\sqrt{n}}{2})\ell(Q_j)$. It follows that $\sup_{z \in Q_j} |x-z| \leqslant \varepsilon + \sqrt{n}\ell(Q_j) \leqslant \varepsilon + \sqrt{n}\frac{\varepsilon}{\sqrt{n}-\frac{1}{2}}$. If we take $\gamma_n = 1 + \frac{\sqrt{n}}{\sqrt{n}-\frac{1}{2}}$ and $r = \gamma_n \varepsilon$, then the closed ball centered at x of radius $\gamma_n \varepsilon$ contains Q_j, i.e., $B(x,r) \supset Q_j$.

iii) Under the same hypotheses as ii), we have $\inf_{z \in Q_j} |x-z| \geqslant (\sqrt{n} - \frac{1}{2})\ell(Q) \geqslant \gamma_n' \varepsilon$ by taking $\gamma_n' = \frac{\sqrt{n}-\frac{1}{2}}{n+\frac{\sqrt{n}}{2}}$. Thus, we obtain $|x-y| \geqslant \gamma_n' \varepsilon$ for every $y \in Q_j$.

With these observations and following the development in the proof of Theorem 5.1, we shall prove for $x \in \mathbb{R}^n \setminus \cup_j Q_j^*$,

$$\sup_{\varepsilon > 0} |T^{(\varepsilon)} b(x)| \leqslant \sum_j \int_{Q_j} |K(x-y) - K(x-c_j)||b_j(y)| dy + C \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |b(y)| dy, \tag{5.35}$$

with $K(x) = \frac{\Omega(x)}{|x|^n}$.

The addition of the maximal function to the r.h.s of (5.35) is the main new element of the proof.

To prove (5.35), 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

$$T^{(\varepsilon)} b(x) = \sum_j \int_{Q_j} K_\varepsilon(x-y) b_j(y) dy. \tag{5.36}$$

Case 1). $K_\varepsilon(x-y) = 0$ if $|x-y| < \varepsilon$, and thus, the integral over these cubes Q_j in (5.36) is zero.

Case 2). $K_\varepsilon(x-y) = K(x-y)$, if $|x-y| > \varepsilon$, and therefore the integral over Q_j equals

$$\int_{Q_j} K(x-y) b_j(y) dy = \int_{Q_j} [K(x-y) - K(x-c_j)] b_j(y) dy.$$

This term is majorized in absolute value by

$$\int_{Q_j} |K(x-y)-K(x-c_j)||b_j(y)|dy,$$

which expression appears in the r.h.s. of (5.35).

Case 3). We write simply

$$\left|\int_{Q_j} K_\varepsilon(x-y)b_j(y)dy\right| \leqslant \int_{Q_j} |K_\varepsilon(x-y)||b_j(y)|dy = \int_{Q_j\cap B(x,r)} |K_\varepsilon(x-y)||b_j(y)|dy,$$

by ii), with $r=\gamma_n\varepsilon$. However, by iii) and because $\Omega(x)$ 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,

$$\left|\int_{Q_j} K_\varepsilon(x-y)b_j(y)dy\right| \leqslant \frac{C}{|B(x,r)|}\int_{Q_j\cap B(x,r)} |b_j(y)|dy.$$

If we sum over all cubes Q_j, we finally obtain, for $r=\gamma_n\varepsilon$,

$$|T^{(\varepsilon)}b(x)| \leqslant \sum_j \int_{Q_j} |K(x-y)-K(x-c_j)||b_j(y)|dy + \frac{C}{|B(x,r)|}\int_{B(x,r)} |b(y)|dy.$$

Taking the supremum over ε gives (5.35). This inequality can be written in the form

$$T^{(*)}b(x) \leqslant \Sigma(x) + CMb(x), \quad x\in\mathbb{R}^n\setminus\cup_j Q_j^*,$$

and so

$$\begin{aligned}&|\{x\in\mathbb{R}^n\setminus\cup_j Q_j^* : T^{(*)}b(x)>\alpha/2\}|\\ \leqslant&|\{x\in\mathbb{R}^n\setminus\cup_j Q_j^* : \Sigma(x)>\alpha/4\}| + |\{x\in\mathbb{R}^n\setminus\cup_j Q_j^* : CMb(x)>\alpha/4\}|.\end{aligned}$$

The first term in the r.h.s. is similar to (5.7), and we can obtain

$$\int_{\mathbb{R}^n\setminus\cup_j Q_j^*} \Sigma(x)dx \leqslant C\|b\|_1$$

which implies $|\{x\in\mathbb{R}^n\setminus\cup_j Q_j^* : \Sigma(x)>\alpha/4\}| \leqslant \frac{4C}{\alpha}\|b\|_1$ by Chebyshev's inequality. For the second one, by Theorem 2.6, i.e., the weak type estimate for the maximal function M, we obtain $|\{x\in\mathbb{R}^n\setminus\cup_j Q_j^* : CMb(x)>\alpha/4\}| \leqslant \frac{C}{\alpha}\|b\|_1$. The weak type $(1,1)$ property of $T^{(*)}$ then follows as in the proof of the same property for T, in Theorem 5.1 for more details.

The final stage of the proof, i.e., (i), the passage from the inequalities of $T^{(*)}$ to the existence of the a.e. limits, follows the familiar pattern described in the proof of the Lebesgue differential theorem (i.e., Theorem 2.12).

More precisely, for any $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, let

$$\Lambda f(x) = \left| \limsup_{\varepsilon \to 0} T^{(\varepsilon)} f(x) - \liminf_{\varepsilon \to 0} T^{(\varepsilon)} f(x) \right|.$$

Clearly, $\Lambda f(x) \leqslant 2T^{(*)} f(x)$. Now for any $\delta > 0$, write $f = f_1 + f_2$ where $f_1 \in \mathcal{C}_c^1$, and $\|f_2\|_p \leqslant \delta$.

We have already proved in the proof of Theorem 5.5 that $T^{(\varepsilon)} f_1$ converges uniformly as $\varepsilon \to 0$, so $\Lambda f_1(x) \equiv 0$. By (5.12), we have $\|\Lambda f_2\|_p \leqslant 2A_p \|f_2\|_p \leqslant 2A_p \delta$ if $1 < p < \infty$. This shows that $\Lambda f_2 = 0$ a.e.; thus, by $\Lambda f(x) \leqslant \Lambda f_1(x) + \Lambda f_2(x)$, we have $\Lambda f = 0$ almost everywhere. Therefore, $\lim_{\varepsilon \to 0} T^{(\varepsilon)} f$ exists a.e. if $1 < p < \infty$.

In the case $p = 1$, we obtain similarly

$$|\{x : \Lambda f(x) > \alpha\}| \leqslant \frac{A}{\alpha} \|f_2\|_1 \leqslant \frac{A\delta}{\alpha},$$

and so again $\Lambda f(x) = 0$ a.e., which implies that $\lim_{\varepsilon \to 0} T^{(\varepsilon)} f(x)$ exists a.e. ❑

§ 5.5 Vector-valued analogues

It is interesting to note that the results of this chapter, where our functions were assumed 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 quickly reviewing certain aspects of integration theory in this context.

Let $\mathcal{H}$ be a separable Hilbert space. Then, a function $f(x)$, from $\mathbb{R}^n$ to $\mathcal{H}$, is ***measurable*** if the scalar valued functions $(f(x), \varphi)$ are measurable, where $(\cdot, \cdot)$ denotes the inner product of $\mathcal{H}$, and φ denotes an arbitrary vector of $\mathcal{H}$.

If $f(x)$ is such a measurable function, then $|f(x)|$ is also measurable (as a function with nonnegative values), where $|\cdot|$ denotes the norm of $\mathcal{H}$.

Thus, $L^p(\mathbb{R}^n, \mathcal{H})$ is defined as the equivalent classes of measurable functions $f(x)$ from $\mathbb{R}^n$ to $\mathcal{H}$, 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 $\mathcal{H}_1$ and $\mathcal{H}_2$ be two separable Hilbert spaces, and let $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ denote the Banach space of bounded linear operators from $\mathcal{H}_1$ to $\mathcal{H}_2$, with the usual operator norm.

We say that a function $f(x)$, from $\mathbb{R}^n$ to $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, is measurable if

$f(x)\varphi$ is an $\mathcal{H}_2$-valued measurable function for every $\varphi \in \mathcal{H}_1$. In this case, $|f(x)|$ is also measurable, and we can define the space $L^p(\mathbb{R}^n, \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2))$, as before; here again $|\cdot|$ denotes the norm, this time in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$.

The usual facts about convolution hold in this setting. For example, let $f \in L^p(\mathbb{R}^n, \mathcal{H}_1)$ and $K \in L^q(\mathbb{R}^n, \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2))$, then $g(x) = \int_{\mathbb{R}^n} K(x-y)f(y)dy$ converges in the norm of $\mathcal{H}_2$ for almost every x, and by the Cauchy-Schwarz inequality

$$|g(x)| \leqslant \int_{\mathbb{R}^n} |K(x-y)f(y)|dy \leqslant \int_{\mathbb{R}^n} |K(x-y)||f(y)|dy.$$

Additionally, $\|g\|_r \leqslant \|K\|_q \|f\|_p$, if $1/r = 1/p + 1/q - 1$, with $1 \leqslant r \leqslant \infty$.

Suppose that $f \in L^1(\mathbb{R}^n, \mathcal{H})$. Then we can define its Fourier transform

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) đx,$$

which is an element of $L^\infty(\mathbb{R}^n, \mathcal{H})$. If $f \in L^1(\mathbb{R}^n, \mathcal{H}) \cap L^2(\mathbb{R}^n, \mathcal{H})$, then $\widehat{f} \in L^2(\mathbb{R}^n, \mathcal{H})$ with $\|\widehat{f}\|_2 = \|f\|_2$. The Fourier transform can then be extended by continuity to a unitary mapping of the Hilbert space $L^2(\mathbb{R}^n, \mathcal{H})$ to itself.

These facts can be obtained easily from the scalar-valued case by introducing an arbitrary orthonormal basis in $\mathcal{H}$.

Now suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are two given Hilbert spaces. Assume that $f(x)$ takes values in $\mathcal{H}_1$, and $K(x)$ takes values in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then

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

whenever defined, takes values in $\mathcal{H}_2$.

Theorem 5.18. *The results in this chapter, in particular Theorems 5.1, 5.5, 5.16 and 5.17, and Proposition 5.2 are valid in the more general context where f takes its value in $\mathcal{H}_1$, K takes its values in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and Tf and $T^{(\varepsilon)}f$ take their values in $\mathcal{H}_2$, and throughout, the absolute value $|\cdot|$ is replaced by the appropriate norm in $\mathcal{H}_1$, $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $\mathcal{H}_2$, respectively.*

This theorem is not a corollary of the scalar-valued case treated in any obvious way. However, its proof consists of nothing but an identical repetition of the arguments given for the scalar-valued cases if we take into account the remarks made in the above paragraphs. Therefore, we leave the proof to the interested reader.

Remark 5.19. 1) The final bounds obtained do not depend on the Hilbert spaces $\mathcal{H}_1$ or $\mathcal{H}_2$ but only on B, p, and n, as in the scalar-valued cases.

2) Most of the argument goes through in the even greater generality of Banach space-valued functions, appropriately defined, one can refer to [Gra14a, pp. 385-414]. The Hilbert space structure is used only in L^2 theory when applying the variant of Plancherel's formula.

The Hilbert space structure also enters in the following corollary.

Corollary 5.20. *With the same assumptions as in Theorem 5.18, if in addition*

$$\|Tf\|_2 = c\|f\|_2, \quad c > 0, \quad f \in L^2(\mathbb{R}^n, \mathcal{H}_1),$$

then $\|f\|_p \leqslant A'_p\|Tf\|_p$, *if* $f \in L^p(\mathbb{R}^n, \mathcal{H}_1)$, $1 < p < \infty$.

Proof. We remark that $L^2(\mathbb{R}^n, \mathcal{H}_j)$ are Hilbert spaces. In fact, let $(\cdot,\cdot)_j$ denote the inner product of $\mathcal{H}_j$, $j = 1, 2$, and let $\langle \cdot,\cdot\rangle_j$ denote the corresponding inner product in $L^2(\mathbb{R}^n, \mathcal{H}_j)$; 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 $L^2(\mathbb{R}^n, \mathcal{H}_1)$ to the Hilbert space $L^2(\mathbb{R}^n, \mathcal{H}_2)$, and thus, by the general theory of inner products (see the theory of Hilbert spaces, e.g., [Din07, Chapter 6, p279]), there exists a unique adjoint transformation T^*, from $L^2(\mathbb{R}^n, \mathcal{H}_2)$ to $L^2(\mathbb{R}^n, \mathcal{H}_1)$, which satisfies the characterizing property

$$\langle Tf_1, f_2\rangle_2 = \langle f_1, T^* f_2\rangle_1, \quad \text{with } f_j \in L^2(\mathbb{R}^n, \mathcal{H}_j).$$

However, in view of the polarization identity, our assumption is equivalent to the identity

$$\langle Tf, Tg\rangle_2 = c^2\langle f, g\rangle_1, \quad \text{for all } f, g \in L^2(\mathbb{R}^n, \mathcal{H}_1).$$

Thus, using the definition of the adjoint, $\langle T^*Tf, g\rangle_1 = c^2\langle f, g\rangle_1$, the assumption can be restated as

$$T^*Tf = c^2 f, \quad f \in L^2(\mathbb{R}^n, \mathcal{H}_1). \tag{5.37}$$

T^* is again an operator of the same kind as T, but it takes a function with values in $\mathcal{H}_2$ to functions with values in $\mathcal{H}_1$, with the kernel $\widetilde{K^*}(x) = K^*(-x)$, where $*$ denotes the adjoint of an element in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$.

This is obvious on the formal level since

$$\begin{aligned}\langle Tf_1, f_2\rangle_2 &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (K(x-y)f_1(y), f_2(x))_2 dy dx \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (f_1(y), K^*(-(y-x))f_2(x))_1 dx dy = \langle f_1, T^* f_2\rangle_1.\end{aligned}$$

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 (5.37), for $f \in (L^2 \cap L^p)(\mathbb{R}^n, \mathcal{H}_1)$,

$$c^2\|f\|_p = \|T^*Tf\|_p \leqslant A_p\|Tf\|_p.$$

This proves the corollary with $A'_p = A_p/c^2$ in view of the density argument of $L^2 \cap L^p$ in L^p for $1 < p < \infty$. ❑

Remark 5.21. This corollary applies in particular to the singular integrals commuted with dilations; then, the condition required is that the multiplier $m(\xi)$ has a constant absolute value. This is the case, for example, when T is the Hilbert transform, $K(x) = \frac{1}{\pi x}$, and $m(\xi) = -i \operatorname{sgn}(\xi)$.

§ 5.6 Littlewood-Paley square function theorem

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 $1 < p < \infty$, in which the Littlewood-Paley square function theorem is a fundamental result.

Definition 5.22. Let ϕ be a real-valued function in $\mathscr{D}(\mathbb{R}^n)$ that is supported in $\mathbb{A} = \{\xi : 2^{-1} \leqslant |\xi| \leqslant 2\}$ and satisfies $\sum_{k\in\mathbb{Z}} \phi_k^2(\xi) = 1$ in $\mathbb{R}^n \setminus \{0\}$, where $\phi_k(\xi) = \phi(2^{-k}\xi)$, we call ϕ a ***Littlewood-Paley function***.

It is not completely obvious that such a function exists.

Lemma 5.23. *A Littlewood-Paley function exists.*

Proof. By the $\mathscr{C}^\infty$ Urysohn lemma (i.e., Theorem 3.21), there exists a function $\tilde{f} \in \mathscr{D}$ such that $\tilde{f} \in [0,1]$, $\tilde{f} = 1$ on $\{\xi : |\xi| \leqslant 1/2\}$ and $\operatorname{supp} \tilde{f} \subset \{\xi : |\xi| < 1\}$. Thus, we can take $f(\xi) = \tilde{f}(\xi/2) - \tilde{f}(\xi)$, which is nonnegative, supported in $\mathbb{A}$. Then,

$$\operatorname{supp} f(2^{-k}\xi) \subset \{\xi : 2^{k-1} \leqslant |\xi| \leqslant 2^{k+1}\}.$$

Therefore, the sum

$$F(\xi) = \sum_{k\in\mathbb{Z}} f^2(2^{-k}\xi)$$

contains at most five nonvanishing terms for each $\xi \neq 0$. Clearly, $F \in \mathscr{S}$, and $F(\xi) > 0$ for $\xi \neq 0$. We set $\phi(\xi) = f(\xi)/F^{1/2}(\xi)$. Obviously, $\phi \in \mathscr{S}$, and satisfies the conditions since $F(2^{-j}\xi) = F(\xi)$. ❑

Definition 5.24. For $f \in L^p$, we can define $Q_k f = (2\pi)^{-n/2} (\phi_k)^\vee * f = \left(\phi_k \hat{f}\right)^\vee$. We define the ***square function*** Sf by

$$Sf(x) = \left(\sum_{k\in\mathbb{Z}} |Q_k f(x)|^2\right)^{1/2}.$$

From the Tonelli theorem and the Plancherel theorem, it is easy to see that

$$\|f\|_2 = \|Sf\|_2 \tag{5.38}$$

and of course this depends on the identity $\sum_{k\in\mathbb{Z}} \phi_k^2(\xi) = 1$. We are interested in this operator because we can characterize the L^p spaces similarly.

Theorem 5.25 (Littlewood-Paley square function theorem)**.** *Let $1 < p < \infty$. There is a finite nonzero constant $C = C(p, n, \phi)$ such that if $f \in L^p$, then*

$$C_{p'}^{-1}\|f\|_p \leqslant \|Sf\|_p \leqslant C_p\|f\|_p.$$

Proof. We prove this theorem via the Calderón-Zygmund Theorem (i.e., Theorem 5.1) and Proposition 5.2 by considering a vector-valued singular integral with the kernel

$$K(x) = (\cdots, (2\pi)^{-n/2}2^{nk}\phi^\vee(2^k x), \cdots),$$

i.e., $K * f = (\cdots, Q_k f, \cdots)$. Clearly, $K \in \mathscr{S}'(\mathbb{R}^n) \cap L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then, it follows

$$|\widehat{K}(\xi)|^2 = (2\pi)^{-n} \sum_{k\in\mathbb{Z}} |\phi_k(\xi)|^2 = (2\pi)^{-n}.$$

We write out the norm of K

$$|K(x)|^2 = (2\pi)^{-n} \sum_{k\in\mathbb{Z}} 2^{2nk}|\phi^\vee(2^k x)|^2.$$

We choose N such that $2^N \leqslant |x| \leqslant 2^{N+1}$ and split the sum above at $-N$. Recall that $\phi^\vee \in \mathscr{S}(\mathbb{R}^n)$ and decays faster than the reciprocal of any polynomial. To estimate the gradient, we observe that $\nabla K(x) = (\cdots, (2\pi)^{-n/2}2^{(n+1)k}\nabla\phi^\vee(2^k x), \cdots)$. Near 0, i.e., for $k \leqslant -N$, we use that $|\nabla\phi^\vee(2^k x)| \leqslant C$. For $k > -N$, we use that $|\nabla\phi^\vee(x)| \leqslant C|x|^{-n-2}$. Thus, we have

$$|\nabla K(x)|^2 \leqslant C\left(\sum_{k\leqslant -N} 2^{2k(n+1)} + \sum_{k>-N} 2^{2k(n+1)}(2^{k+N})^{-2(n+2)}\right) = C2^{-2N(n+1)}.$$

Recalling that $2^N \sim |x|$, we obtain the desired upper-bound for $|\nabla K(x)| \leqslant C|x|^{-(n+1)}$. Therefore, by the vector-valued version of Theorem 5.1 and Proposition 5.2, i.e., Theorem 5.18, we obtain the right-hand inequality in the theorem

$$\|Sf\|_p = \left\| \left(\sum_{k\in\mathbb{Z}} |Q_k f|^2\right)^{1/2} \right\|_p = \|K * f\|_p \leqslant C_p\|f\|_p.$$

For the converse inequality, it follows from Corollary 5.20 due to (5.38). ❑

§ 5.7 Mikhlin and Hörmander multiplier theorem

We introduce a partition of unity to be frequently used later.

Lemma 5.26. *There exists a function* $\varphi(\mathbb{R}^n)$, *such that*
(i) $\operatorname{supp}\varphi = \{\xi : 2^{-1} \leqslant |\xi| \leqslant 2\}$;
(ii) $\varphi(\xi) > 0$ *for* $2^{-1} < |\xi| < 2$;
(iii) $\sum\limits_{k\in\mathbb{Z}} \varphi(2^{-k}\xi) = 1$ *for* $\xi \neq 0$.

Proof. The proof is similar to that of Lemma 5.23. Choose any function $f \in \mathscr{S}$ such that (i) and (ii) are satisfied. Then,

$$\operatorname{supp} f(2^{-k}\xi) = \{\xi : 2^{k-1} \leqslant |\xi| \leqslant 2^{k+1}\}.$$

Therefore, the sum

$$F(\xi) = \sum_{k\in\mathbb{Z}} f(2^{-k}\xi)$$

contains at most five nonvanishing terms for each $\xi \neq 0$. Clearly, $F \in \mathscr{S}$, and $F(\xi) > 0$ for $\xi \neq 0$. Let $\varphi = f/F$. Obviously, $\varphi \in \mathscr{S}$, and satisfies (i) and (ii). Since $F(2^{-j}\xi) = F(\xi)$, φ also satisfies (iii). ❑

Theorem 5.27 (Mikhlin multiplier theorem)**.** *Let* $\mathcal{H}_0$ *and* $\mathcal{H}_1$ *be Hilbert spaces. Assume that* m *is a mapping from* $\mathbb{R}^n$ *to* $\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$ *and that*

$$|\xi|^{|\alpha|}|\partial_\xi^\alpha m(\xi)|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)} \leqslant A, \quad |\alpha| \leqslant k, \tag{5.39}$$

for some integer $k > n/2$. *Then* $m \in \mathcal{M}_p(\mathcal{H}_0, \mathcal{H}_1)$, $1 < p < \infty$, *and*

$$\|m\|_{\mathcal{M}_p} \leqslant C_p A.$$

Proof. We use the vector version of the Calderón-Zygmund theorem (i.e., Theorem 5.1) to prove it. Denote $T_m f = \left(m\widehat{f}\right)^\vee = (2\pi)^{-n/2} m^\vee * f =: K * f$. It is clear that $K \in \mathscr{S}'(\mathbb{R}^n) \cap L^1_{\text{loc}}(\mathbb{R}^n)$ by the assumption. For convenience, we denote $|\cdot|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)}$ by $|\cdot|$. Thus, taking $k = 0$ in (5.39), we have $|\widehat{K}| = (2\pi)^{-n/2}|m(\xi)| \leqslant (2\pi)^{-n/2}A$. Thus, we only need to verify

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant C \quad \text{uniformly in } y. \tag{5.40}$$

Denote $\varphi_j(\xi) = \varphi(2^{-j}\xi)$, where φ is given by Lemma 5.26. We write

$$m(\xi) = \sum_{j\in\mathbb{Z}} m_j(\xi), \quad \text{where } m_j = \varphi_j m.$$

Let us now prove (5.40) assuming (5.39) holds for $|\alpha| \leqslant k$. By the

Plancherel theorem and Leibniz' rule, we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} |x^\alpha m_j^\vee(x)|^2 dx =& C\int_{\mathbb{R}^n} |\partial_\xi^\alpha m_j(\xi)|^2 d\xi \\ =& C\int_{\mathbb{R}^n} \left| \sum_{\alpha_1+\alpha_2=\alpha} c_{\alpha_1,\alpha_2} \partial_\xi^{\alpha_1} m(\xi) \partial_\xi^{\alpha_2} \varphi_j(\xi) \right|^2 d\xi \\ \leqslant& CA^2 \sum_{\alpha_1+\alpha_2=\alpha} \int_{|\xi|\sim 2^j} 2^{-2j|\alpha_2|} |\xi|^{-2|\alpha_1|} \\ \leqslant& CA^2 2^{j(n-2|\alpha|)}.\end{aligned}$$

Using the Cauchy-Schwarz inequality and applying the above with $|\alpha| = 0$, we have

$$\int_{|x|\leqslant R} |m_j^\vee(x)| dx \leqslant CA(2^j R)^{n/2}.$$

By applying the above with $|\alpha| = k$, we obtain

$$\begin{aligned}\int_{|x|>R} |m_j^\vee(x)| dx \leqslant& \left(\int_{\mathbb{R}^n} |x^\alpha m_j^\vee(x)|^2 dx \right)^{1/2} \left(\int_{|x|>R} |x|^{-2|\alpha|} dx \right)^{1/2} \\ \leqslant& CA(2^j R)^{\frac{n}{2}-k}.\end{aligned} \tag{5.41}$$

Choosing $R \sim 2^{-j}$, we find

$$\int_{\mathbb{R}^n} |m_j^\vee(x)| dx \leqslant CA \quad \text{uniformly in } j.$$

Arguing in the same way, we have

$$\int_{\mathbb{R}^n} |\nabla m_j^\vee(x)| dx \leqslant CA2^j.$$

In particular, by the mean value theorem, this shows

$$\int_{\mathbb{R}^n} |m_j^\vee(x-y) - m_j^\vee(x)| dx \leqslant CA2^j |y|. \tag{5.42}$$

Thus, we have from (5.42) and (5.41)

$$\begin{aligned}\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)| dx \leqslant& \sum_{j\in\mathbb{Z}} \int_{|x|\geqslant 2|y|} |m_j^\vee(x-y) - m_j^\vee(x)| dx \\ \leqslant& CA \sum_{2^j \leqslant |y|^{-1}} 2^j |y| + C \sum_{2^j > |y|^{-1}} \int_{|x|\geqslant |y|} |m_j^\vee(x)| dx \\ \leqslant& CA + CA \sum_{2^j > |y|^{-1}} (2^j |y|)^{\frac{n}{2}-k} \\ \leqslant& CA.\end{aligned}$$

This completes the proof by the vector version of the Calderón-Zygmund theorem (Theorem 5.1), i.e., Theorem 5.18. ❑

Remark 5.28. This result is sharp in the sense that the L^1 and L^∞ bounds can fail. To see this, let us consider the Hilbert transform, which essentially corresponds to taking $m^\vee(x) = \frac{C}{x}$ in $n = 1$. We know that

$m(\xi) = c\,\text{sgn}\,(\xi)$, which satisfies the condition (5.39) with $k = 1$. However, as we have shown before, the Hilbert transform is not bounded on L^1 or L^∞.

Example 5.29. One application of the Mikhlin multiplier theorem is the following "Schauder" type estimate, which is useful in the setting of elliptic PDEs: for any $i, j = 1, \cdots, n$, $1 < p < \infty$, and any $f \in \mathscr{S}(\mathbb{R}^n)$

$$\left\| \frac{\partial^2 f}{\partial x_i \partial x_j} \right\|_p \leqslant C \|\Delta f\|_p.$$

Indeed, this is equivalent to $m_{ij}(\xi) := \frac{\xi_i \xi_j}{|\xi|^2}$, which is a consequence of the Mikhlin multiplier theorem.

The proof of the theorem leads to a generalization of its statement, which we formulate as a corollary. We leave the proof as an exercise.

Corollary 5.30 (Hörmander multiplier theorem). *Let $\mathcal{H}_0$ and $\mathcal{H}_1$ be Hilbert spaces. Assume that m is a mapping from $\mathbb{R}^n$ to $\mathcal{L}(\mathcal{H}_0, \mathcal{H}_1)$ and that*

$$\begin{aligned} |m(\xi)|_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)} &\leqslant A, \\ \sup_{0<R<\infty} R^{2|\alpha|-n} \int_{R \leqslant |\xi| \leqslant 2R} \left| \partial_\xi^\alpha m(\xi) \right|^2_{\mathcal{L}(\mathcal{H}_0,\mathcal{H}_1)} d\xi &\leqslant A, \quad |\alpha| \leqslant k, \end{aligned} \tag{5.43}$$

for some integer $k > n/2$. Then, $m \in \mathcal{M}_p(\mathcal{H}_0, \mathcal{H}_1)$, $1 < p < \infty$, and

$$\|m\|_{\mathcal{M}_p} \leqslant C_p A.$$

Exercises

Exercise 5.1. Let Ω be an integrable function with mean value zero on the sphere $\mathbb{S}^{n-1}$. Suppose that Ω satisfies a Hölder condition of order $0 < \alpha < 1$ on $\mathbb{S}^{n-1}$. This means that

$$|\Omega(x) - \Omega(y)| \leqslant B_0 |x - y|^\alpha$$

for all $x, y \in \mathbb{S}^{n-1}$. Prove that the function $K(x) = \Omega(x/|x|)/|x|^n$ satisfies Hörmander's condition with a constant at most a multiple of $B_0 + \|\Omega\|_\infty$.

Exercise 5.2. [Gra14a, Exercise 5.1.8] Let $Q_y^{(j)}$ be the jth conjugate Poisson kernel of P_y defined by

$$Q_y^{(j)}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \frac{x_j}{(|x|^2 + y^2)^{\frac{n+1}{2}}}.$$

(i) Calculate the Fourier transform of $Q_y^{(j)}$.

(ii) Conclude that $R_j P_y = Q_y^{(j)}$ and for $f \in L^2(\mathbb{R}^n)$, we have $R_j f * P_y = f * Q_y^{(j)}$.

Exercise 5.3. [Gra14a, Exercise 5.1.5] Let $1 \leqslant p < \infty$ and let T be a linear operator defined on $\mathscr{S}$ that commutes with dilations. Suppose that there exists a constant $C > 0$ such that for all $f \in \mathscr{S}(\mathbb{R}^n)$ with L^p norm one we have

$$|\{x : |Tf(x)| > 1\}| \leqslant C.$$

Prove that T admits a bounded extension from $L^p(\mathbb{R}^n)$ to $L^{p,\infty}(\mathbb{R}^n)$ with norm at most $C^{1/p}$.

Hint Try functions of the form $\lambda^{-n/p} f(\lambda^{-1}x)/\|f\|_p$ with $\lambda > 0$.

Exercise 5.4. [Gra14a, Exercise 5.2.6] For $\Omega \in L^1(\mathbb{S}^{n-1})$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, define

$$M_\Omega f(x) = \sup_{R>0} \frac{1}{V_n R^n} \int_{|y|\leqslant R} |\Omega(y/|y|)||f(x-y)|dy.$$

Apply the method of rotations to prove that M_Ω maps $L^p(\mathbb{R}^n)$ to itself for $1 < p < \infty$.

Exercise 5.5. Prove Corollary 5.30 .

Exercise 5.6. [Gra14b, Exercise 6.2.5] Suppose that $\varphi(\xi)$ is a smooth function on $\mathbb{R}^n$ that vanishes in a neighborhood of the origin and is equal to 1 in a neighborhood of infinity. Prove that the function $e^{i\xi_j|\xi|^{-1}}\varphi(\xi)$ is in $\mathcal{M}_p(\mathbb{R}^n)$ for $1 < p < \infty$ for every ξ_j.

Exercise 5.7. [Gra14b, Exercise 6.2.7] Let $\zeta(\xi)$ be a smooth function on the line that is supported in a compact set that does not contain the origin and let a_j be a bounded sequence of complex numbers. Prove that the function $m(\xi) = \sum\limits_{j\in\mathbb{Z}} a_j \zeta(2^{-j}\xi)$ is in $\mathcal{M}_p(\mathbb{R})$ for all $1 < p < \infty$.

Exercise 5.8. Let P be a polynomial with complex coefficients, of degree greater than or equal to 1, without a real root.

i) Prove that $\sup\limits_{x\in\mathbb{R}} 1/|P(x)|$ and $\sup\limits_{x\in\mathbb{R}} |P'(x)/P(x)|$ are finite.

ii) Prove that, for all $p > 1$, $1/P(\ln|x|) \in \mathcal{M}_p(\mathbb{R})$.

Exercise 5.9. Let $x \in \mathbb{R}$. Is $(1 + |\ln|x||)^{-1/2}$ a Fourier multiplier?

Exercise 5.10. [Gra14b, Exercise 6.1.3, 6.1.2, 6.1.4] Let Ψ be an integrable function on $\mathbb{R}^n$ with mean value zero that satisfies

$$|\Psi(x)| \leqslant B(1+|x|)^{-n-\varepsilon}, \qquad \int_{\mathbb{R}^n} |\Psi(x-y) - \Psi(x)|dx \leqslant B|y|^{\varepsilon'},$$

for some $B, \varepsilon, \varepsilon' > 0$ and for all $y \neq 0$. Let $\Psi_t = t^{-n}\Psi(x/t)$.

(i) Prove that $|\widehat{\Psi}(\xi)| \leqslant c_{n,\varepsilon,\varepsilon'} B \min(|\xi|^{\min(1,\varepsilon/2)}, |\xi|^{-\varepsilon})$ for some constant $c_{n,\varepsilon,\varepsilon'}$ and conclude that

$$\left\| \left(\sum_{k\in\mathbb{Z}} |\Psi_{2^{-k}} * f|^2 \right)^{1/2} \right\|_2 \leqslant C_n B \|f\|_2.$$

(ii) Prove that for some constant $C_{\varepsilon,\varepsilon'} < \infty$ we have

$$\sup_{\xi\in\mathbb{R}^n}\left(\int_0^\infty |\widehat{\Psi}(t\xi)|^2\frac{dt}{t}\right)^{1/2} + \sup_{\xi\in\mathbb{R}^n}\left(\sum_{k\in\mathbb{Z}} |\widehat{\Psi}(2^{-k}\xi)|^2\right)^{1/2} \leqslant C_{\varepsilon,\varepsilon'}B.$$

(iii) Prove that there exists a constant C_n such that for all $f \in L^2(\mathbb{R}^n)$ we have

$$\left\|\left(\int_0^\infty |f*\Psi_t|^2\frac{dt}{t}\right)^{1/2}\right\|_2 \leqslant C_n B\|f\|_2.$$

Hint i) Make use of the identity

$$\widehat{\Psi}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\Psi(x)đx = -\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\Psi(x-y)đx,$$

where $y = \frac{\xi}{2|\xi|^2}$ when $|\xi| \geqslant 1$. For $|\xi| \leqslant 1$ use the mean value property of Ψ to write $\widehat{\Psi}(\xi) = \int_{\mathbb{R}^n}(e^{-ix\cdot\xi} - 1)\Psi(x)đx$ and split the integral in the regions $|x| \leqslant 1$ and $|x| \geqslant 1$.

Exercise 5.11. [Gra14b, Exercise 6.2.8] Let $\widehat{\zeta}(\xi)$ be a smooth function on $\mathbb{R}^n$ supported in a compact set that does not contain the origin and let $\dot{\Delta}_j^\zeta f = \mathscr{F}^{-1}(\widehat{\zeta}(2^{-j}\xi)\widehat{f}(\xi))$. Show that the operator

$$f \to \sup_{N\in\mathbb{Z}}\left|\sum_{j<N}\dot{\Delta}_j^\zeta f\right|$$

is bounded on $L^p(\mathbb{R})$ when $1 < p < \infty$.

Hint Pick $\varphi \in \mathscr{S}$ satisfying $\sum_{j\in\mathbb{Z}}\widehat{\varphi}(2^{-j}\xi) = 1$ on $\mathbb{R}^n \setminus \{0\}$ with $\operatorname{supp}\widehat{\varphi} \subset \{\xi : 2^{-1} \leqslant |\xi| \leqslant 2\}$. Then, $\dot{\Delta}_k^\varphi\dot{\Delta}_j^\zeta = 0$ if $|j-k| < c_0$, and we have

$$\sum_{j<N}\dot{\Delta}_j^\zeta = \sum_{k<N+c_0}\dot{\Delta}_k^\varphi\sum_{j<N}\dot{\Delta}_j^\zeta = \sum_{k<N+c_0}\dot{\Delta}_k^\varphi\sum_{j}\dot{\Delta}_j^\zeta - \sum_{k<N+c_0}\dot{\Delta}_k^\varphi\sum_{j\geqslant N}\dot{\Delta}_j^\zeta,$$

which is a finite sum plus a term controlled by a multiple of the operator

$$f \to M\left(\sum_{j\in\mathbb{Z}}\dot{\Delta}_j^\zeta f\right),$$

where M is the Hardy-Littlewood maximal function.

Exercise 5.12. [Gra14b, Exercise 6.2.9] Let φ be given in Lemma 5.23. Let $\dot{\Delta}_j g = \left(\varphi(2^{-j}\xi)\widehat{g}\right)^\vee$. Prove that

$$\left\|\sum_{|j|<N}\dot{\Delta}_j g - g\right\|_p \to 0$$

as $N \to \infty$ for all $g \in \mathscr{S}(\mathbb{R}^n)$. Deduce that Schwartz functions whose Fourier transforms have compact supports that do not contain the origin are dense in $L^p(\mathbb{R}^n)$ for $1 < p < \infty$.

Hint Use the result of Exercise 5.11 and the dominated convergence theorem.

6

Riesz and Bessel potentials

In this chapter, we introduce the Riesz and Bessel potentials. Based on these potentials, we introduce the general Sobolev spaces, i.e., Bessel (Riesz) potential spaces.

§ 6.1 Riesz potentials and fractional integrals

The Laplacian satisfies the following identity for all $f \in \mathscr{S}(\mathbb{R}^n)$:

$$\widehat{-\Delta f}(\xi) = |\xi|^2 \hat{f}(\xi). \tag{6.1}$$

From this, we replace the exponent 2 in $|\xi|^2$ by a general exponent s and thus define (at least formally) the fractional power of the Laplacian by

$$(-\Delta)^{s/2} f = \left(|\xi|^s \hat{f}\right)^{\vee}. \tag{6.2}$$

Of special significance will be the negative powers s in the range $-n < s < 0$. In general, with a slight change in notation, we can define

Definition 6.1. Let $s > 0$. The ***Riesz potential*** of order s is the operator

$$I_s = (-\Delta)^{-s/2}. \tag{6.3}$$

For $0 < s < n$ and $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, I_s is actually given in the form

$$I_s f(x) = \frac{1}{\gamma(s)} \int_{\mathbb{R}^n} |x-y|^{-n+s} f(y) dy, \tag{6.4}$$

with

$$\gamma(s) = 2^{-\frac{n}{2}} 2^s \frac{\Gamma(s/2)}{\Gamma((n-s)/2)}.$$

Now, we state two further identities that can be obtained from (6.2) or (6.3) and that reflect the essential properties of the potentials I_s:

$$I_s(I_t f) = I_{s+t} f, \quad f \in \mathscr{S}, \ s, t > 0, \ s+t < n; \tag{6.5}$$

$$\Delta(I_s f) = I_s(\Delta f) = -I_{s-2} f, \quad f \in \mathscr{S}, \ n \geqslant 3, \ 2 \leqslant s \leqslant n. \tag{6.6}$$

The deduction of these two identities has no real difficulties, and these are best left to the interested reader to work out.

A simple consequence of (6.5) is the n-dimensional variant of the Beta function,[1]

$$\int_{\mathbb{R}^n} |x-y|^{-n+s}|y|^{-n+t}dy = \frac{\gamma(s)\gamma(t)}{\gamma(s+t)}|x|^{-n+(s+t)} \quad \text{in } \mathscr{S}', \tag{6.7}$$

with $s, t > 0$ and $s + t < n$. Indeed, for any $\varphi \in \mathscr{S}$, we have, by the definition of Riesz potentials and (6.5), that

$$\begin{aligned}
&\iint_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^{-n+s}|y|^{-n+t}dy\varphi(z-x)dx \\
=&\int_{\mathbb{R}^n}|y|^{-n+t}\int_{\mathbb{R}^n}|x-y|^{-n+s}\varphi(z-y-(x-y))dxdy \\
=&\int_{\mathbb{R}^n}|y|^{-n+t}\gamma(s)I_s\varphi(z-y)dy = \gamma(s)\gamma(t)I_t(I_s\varphi)(z) = \gamma(s)\gamma(t)I_{s+t}\varphi(z) \\
=&\frac{\gamma(s)\gamma(t)}{\gamma(s+t)}\int_{\mathbb{R}^n}|x|^{-n+(s+t)}\varphi(z-x)dx.
\end{aligned}$$

We have considered the Riesz potentials formally and the operation for Schwartz functions. However, 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

$$\|I_s f\|_q \leqslant A\|f\|_p? \tag{6.8}$$

There is a simple necessary condition that is merely a reflection of the homogeneity of the kernel $(\gamma(s))^{-1}|y|^{-n+s}$. In fact, we have

Proposition 6.2. *If inequality (6.8) holds for all $f \in \mathscr{S}$ and a finite constant A, then $1/q = 1/p - s/n$.*

Proof. Let us consider the dilation operator δ^ε, defined by $\delta^\varepsilon f(x) = f(\varepsilon x)$ for $\varepsilon > 0$. Then clearly, for $\varepsilon > 0$ and any $f \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\begin{aligned}
(I_s\delta^\varepsilon f)(x) =& \frac{1}{\gamma(s)}\int_{\mathbb{R}^n}|x-y|^{-n+s}f(\varepsilon y)dy \\
\overset{z=\varepsilon y}{=\!=}&\varepsilon^{-n}\frac{1}{\gamma(s)}\int_{\mathbb{R}^n}|x-\varepsilon^{-1}z|^{-n+s}f(z)dz \\
=&\varepsilon^{-s}I_sf(\varepsilon x).
\end{aligned} \tag{6.9}$$

Noticing that

$$\|\delta^\varepsilon f\|_p = \varepsilon^{-n/p}\|f\|_p, \quad \|\delta_{\varepsilon^{-1}}I_sf\|_q = \varepsilon^{n/q}\|I_sf\|_q. \tag{6.10}$$

[1]The Beta function, also called the Euler integral of the first kind, is a special function defined by $B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$ for $\operatorname{Re}\alpha > 0$ and $\operatorname{Re}\beta > 0$. It has the relation with Γ-function: $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$.

By (6.8), we obtain

$$\begin{aligned}\|I_s f\|_q =&\varepsilon^s \|\delta^{\varepsilon^{-1}} I_s \delta^\varepsilon f\|_q = \varepsilon^{s+n/q}\|I_s\delta^\varepsilon f\|_q \\ \leqslant& A\varepsilon^{s+n/q}\|\delta^\varepsilon f\|_p = A\varepsilon^{s+n/q-n/p}\|f\|_p.\end{aligned}$$

If $s+n/q-n/p > 0$, let $\varepsilon \to 0^+$; if $s+n/q-n/p < 0$, let $\varepsilon \to \infty$, we always have $\|I_s f\|_q = 0$ for any $f \in \mathscr{S}(\mathbb{R}^n)$. However, if $f \not\equiv 0$ is nonnegative, then $I_s f > 0$ everywhere and hence $\|I_s f\|_q > 0$; thus, we can conclude the desired relations

$$1/q = 1/p - s/n. \qquad \square$$

Now, we give the following Hardy-Littlewood-Sobolev theorem of fractional integration. The result was first considered in one dimension on the circle by Hardy and Littlewood and n-dimension by Sobolev.

Theorem 6.3 (Hardy-Littlewood-Sobolev theorem). *Let* $0 < s < n$, $1 \leqslant p < q < \infty$, $1/q = 1/p - s/n$.

(i) *If* $f \in L^p(\mathbb{R}^n)$, *then the integral* (6.4), *defining* $I_s f$, *converges absolutely for almost every* x.

(ii) *If, in addition,* $p > 1$, *then* $\|I_s f\|_q \leqslant A_{p,q}\|f\|_p$.

(iii) *If* $f \in L^1(\mathbb{R}^n)$, *then* $|\{x : |I_s f(x)| > \alpha\}| \leqslant (A\alpha^{-1}\|f\|_1)^q$, *for all* $\alpha > 0$. *That is, the mapping* I_s *is of weak type* $(1,q)$, *with* $1/q = 1 - s/n$.

Proof. We first prove parts (i) and (ii). 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 $E_j := B(x, 2^{-j}\delta) \setminus B(x, 2^{-(j+1)}\delta)$, $j = 0, 1, 2, \ldots$, we have

$$\begin{aligned}|L_\delta(x)| \leqslant& \left|\sum_{j=0}^\infty \int_{E_j} |x-y|^{-n+s} f(y)dy\right| \leqslant \sum_{j=0}^\infty \int_{E_j} |x-y|^{-n+s}|f(y)|dy \\ \leqslant& \sum_{j=0}^\infty \int_{E_j} (2^{-(j+1)}\delta)^{-n+s}|f(y)|dy \\ \leqslant& \sum_{j=0}^\infty \int_{B(x,2^{-j}\delta)} (2^{-(j+1)}\delta)^{-n+s}|f(y)|dy \\ =& \sum_{j=0}^\infty \frac{(2^{-(j+1)}\delta)^{-n+s}|B(x,2^{-j}\delta)|}{|B(x,2^{-j}\delta)|} \int_{B(x,2^{-j}\delta)} |f(y)|dy \\ =& \sum_{j=0}^\infty \frac{(2^{-(j+1)}\delta)^{-n+s} V_n (2^{-j}\delta)^n}{|B(x,2^{-j}\delta)|} \int_{B(x,2^{-j}\delta)} |f(y)|dy\end{aligned}$$

$$\leqslant V_n\delta^s 2^{n-s}\sum_{j=0}^{\infty}2^{-sj}Mf(x)=\frac{V_n\delta^s 2^n}{2^s-1}Mf(x).$$

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

$$\begin{aligned}|H_\delta(x)| \leqslant &\|f\|_p\left(\int_{\mathbb{R}^n\setminus B(x,\delta)}|x-y|^{(-n+s)p'}dy\right)^{1/p'}\\ =&\|f\|_p\left(\int_{S^{n-1}}\int_\delta^\infty r^{(-n+s)p'}r^{n-1}drd\sigma\right)^{1/p'}\\ =&\omega_{n-1}^{1/p'}\|f\|_p\left(\int_\delta^\infty r^{(-n+s)p'+n-1}dr\right)^{1/p'}\\ =&\left(\frac{\omega_{n-1}}{(n-s)p'-n}\right)^{1/p'}\delta^{n/p'-(n-s)}\|f\|_p=C(n,s,p)\delta^{s-n/p}\|f\|_p.\end{aligned}$$

By the above two inequalities, we have

$$|\gamma(s)I_sf(x)|\leqslant C(n,s)\delta^s Mf(x)+C(n,s,p)\delta^{s-n/p}\|f\|_p=:F(\delta).$$

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 obtain

$$|\gamma(s)I_sf(x)|\leqslant C(Mf(x))^{1-ps/n}\|f\|_p^{ps/n}.$$

Therefore, by part (i) of Theorem 2.6 for maximal functions, i.e., Mf is finite almost everywhere if $f\in L^p$ ($1\leqslant p\leqslant\infty$), it follows that $|I_sf(x)|$ is finite almost everywhere, which proves part (i) of the theorem.

By part (iii) of Theorem 2.6, we know $\|Mf\|_p\leqslant A_p\|f\|_p$ ($1<p\leqslant\infty$); thus,

$$\|I_sf\|_q\leqslant C\|Mf\|_p^{1-ps/n}\|f\|_p^{ps/n}=C\|f\|_p.$$

This gives the proof of part (ii).

Finally, we prove (iii). 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 2.6, we obtain

$$\begin{aligned}&|\{x:|I_sf(x)|>2(\gamma(s))^{-1}\alpha\}|\\ \leqslant&|\{x:|L_\delta(x)|>\alpha\}|+|\{x:|H_\delta(x)|>\alpha\}|\\ \leqslant&|\{x:|C\delta^sMf(x)|>\alpha\}|+0\\ \leqslant&\frac{C}{\delta^{-s}\alpha}\|f\|_1=C[\|f\|_1/\alpha]^{n/(n-s)}=C[\|f\|_1/\alpha]^q.\end{aligned}$$

This completes the proof of part (iii). ❑

§ 6.2 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 $|x| \to \infty$ worsens 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 6.4. Let $s > 0$. The ***Bessel potential*** of order s is the operator

$$J_s = (I - \Delta)^{-s/2}$$

whose action on functions f is given by

$$J_s f = (2\pi)^{n/2} \left(\widehat{G_s}\widehat{f}\right)^{\vee} = G_s * f,$$

where

$$G_s(x) = (2\pi)^{-n/2} \left(\langle\xi\rangle^{-s}\right)^{\vee}(x), \quad \langle\xi\rangle = (1 + |\xi|^2)^{1/2}.$$

Now we give some properties of $G_s(x)$ and show why this adjustment yields exponential decay for G_s at infinity.

Proposition 6.5. *Let $s > 0$.*

(i) $G_s(x) = \dfrac{1}{(4\pi)^{n/2}\Gamma(s/2)} \displaystyle\int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}$.

(ii) $G_s(x) > 0, \quad \forall x \in \mathbb{R}^n$*; and* $G_s \in L^1(\mathbb{R}^n)$*; precisely,* $\int_{\mathbb{R}^n} G_s(x)dx = 1$.

(iii) *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

$$\frac{1}{c(s,n)} \leqslant \frac{G_s(x)}{H_s(x)} \leqslant c(s,n), \quad \text{when } |x| \leqslant 2,$$

where H_s is a function satisfying

$$H_s(x) = \begin{cases} |x|^{s-n} + 1 + O(|x|^{s-n+2}), & 0 < s < n, \\ \ln \dfrac{2}{|x|} + 1 + O(|x|^2), & s = n, \\ 1 + O(|x|^{s-n}), & s > n, \end{cases}$$

as $|x| \to 0$.
(iv) $G_s \in L^{p'}(\mathbb{R}^n)$ *for any* $1 \leqslant p \leqslant \infty$ *and* $s > n/p$.

Proof. (i) 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

$$\langle \xi \rangle^{-s} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} e^{-t|\xi|^2} t^{s/2} \frac{dt}{t}.$$

Note that the above integral converges at both ends (as $|\xi| \to 0$, or ∞). Now, taking the inverse Fourier transform in ξ and using Theorem 3.3, we obtain

$$\begin{aligned} G_s(x) =&(2\pi)^{-n/2} \frac{1}{\Gamma(s/2)} \mathscr{F}_\xi^{-1} \int_0^\infty e^{-t} e^{-t|\xi|^2} t^{s/2} \frac{dt}{t} \\ =&(2\pi)^{-n/2} \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} \mathscr{F}_\xi^{-1} \left(e^{-t|\xi|^2} \right) t^{s/2} \frac{dt}{t} \\ =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}. \end{aligned}$$

(ii) We have easily[a] $\int_{\mathbb{R}^n} G_s(x) dx = (2\pi)^{n/2} \widehat{G_s}(0) = 1$. Thus, $G_s \in L^1(\mathbb{R}^n)$.

(iii) 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

$$\begin{aligned} G_s^1(x) =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^{|x|^2} e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}, \\ G_s^2(x) =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_{|x|^2}^4 e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}, \end{aligned}$$

$$G_s^3(x) = \frac{1}{(4\pi)^{n/2}\Gamma(s/2)}\int_4^\infty e^{-t}e^{-\frac{|x|^2}{4t}}t^{\frac{s-n}{2}}\frac{dt}{t}.$$

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

$$\begin{aligned}G_s^1(x) =&\frac{|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^1 e^{-t|x|^2}e^{-\frac{1}{4t}}t^{\frac{s-n}{2}}\frac{dt}{t}\\ =&\frac{|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^1 e^{-\frac{1}{4t}}t^{\frac{s-n}{2}}\frac{dt}{t} + \frac{O(|x|^{s-n+2})}{(4\pi)^{n/2}\Gamma(s/2)}\int_0^1 e^{-\frac{1}{4t}}t^{\frac{s-n}{2}}dt\\ =&\frac{2^{n-s-2}|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)}\int_{1/4}^\infty e^{-y}y^{\frac{n-s}{2}}\frac{dy}{y} + \frac{2^{n-s-4}O(|x|^{s-n+2})}{(4\pi)^{n/2}\Gamma(s/2)}\int_{1/4}^\infty e^{-y}y^{\frac{n-s}{2}}\frac{dy}{y^2}\\ =&c_{s,n}^1|x|^{s-n} + O(|x|^{s-n+2}), \quad \text{as } |x| \to 0.\end{aligned}$$

Since $0 \leqslant \frac{|x|^2}{4t} \leqslant \frac{1}{4}$ 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

$$G_s^2(x) \sim \int_{|x|^2}^4 t^{(s-n)/2}\frac{dt}{t} = \begin{cases}\dfrac{|x|^{s-n}}{n-s} - \dfrac{2^{s-n+1}}{n-s}, & s < n,\\ 2\ln\frac{2}{|x|}, & s = n,\\ \dfrac{2^{s-n+1}}{s-n}, & s > n.\end{cases}$$

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.

(iv) For $p = 1$ and so $p' = \infty$, by part (iii), we have $\|G_s\|_\infty \leqslant C$ for $s > n$.

Next, we assume that $1 < p \leqslant \infty$ and so $1 \leqslant p' < \infty$. Again by part (iii), 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

$$\begin{aligned}\int_{|x|\leqslant 2}(\ln\frac{2}{|x|})^q dx =&2\int_0^2(\ln 2/r)^q dr = 4\int_0^1(\ln 1/r)^q dr\\ =&4\int_0^\infty t^q e^{-t}dt = 4\Gamma(q+1)\end{aligned}$$

for $q > 0$ by changing the variable $r = e^{-t}$. 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 $\|G_s\|_{p'} \leqslant C$ for any $1 \leqslant p \leqslant \infty$ and $s > n/p$, which implies the desired result. ❏

[a]Or use (i) to show it. From part (i), 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

$$\begin{aligned}\int_{\mathbb{R}^n} G_s(x)dx =& \int_{\mathbb{R}^n} \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t} dx \\ =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} \int_{\mathbb{R}^n} e^{-\frac{|x|^2}{4t}} dx t^{\frac{s-n}{2}} \frac{dt}{t} \\ =& \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} (4\pi t)^{n/2} t^{\frac{s-n}{2}} \frac{dt}{t} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} t^{\frac{s}{2}-1} dt = 1.\end{aligned}$$

We also have a result analogous to that of Riesz potentials for the operator J_s.

Theorem 6.6. (i) *For all $0 < s < \infty$, the operator J_s maps $L^r(\mathbb{R}^n)$ into itself with norm 1 for all $1 \leqslant r \leqslant \infty$.*

(ii) *Let $0 < s < n$ and $1 < p < q < \infty$ satisfy $1/q = 1/p - s/n$. Then there exists a constant $C_{n,s,p} > 0$ such that for all $f \in L^p(\mathbb{R}^n)$, we have*

$$\|J_s f\|_q \leqslant C_{n,s,p} \|f\|_p.$$

(iii) *If $f \in L^1(\mathbb{R}^n)$, then $|\{x : |J_s f(x)| > \alpha\}| \leqslant (C_{n,s}\alpha^{-1}\|f\|_1)^q$, for all $\alpha > 0$. That is, the mapping J_s is of weak type $(1, q)$, with $1/q = 1 - s/n$.*

Proof. By Young's inequality, we have $\|J_s f\|_r = \|G_s * f\|_r \leqslant \|G_s\|_1 \|f\|_r = \|f\|_r$. This proves result (i).

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 can use the function $e^{-|y|/2} \in L^r$ for all $1 \leqslant r \leqslant \infty$, Young's inequality and Theorem 6.3 to complete the proofs of (ii) and (iii). ❏

§ 6.3 General Sobolev spaces H_p^s and $\dot{H}_p^s$

We start by weakening the notation of partial derivatives by the theory of distributions. The appropriate definition is stated in terms of the space

$\mathscr{S}(\mathbb{R}^n)$.

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

$$\int_{\mathbb{R}^n} f(x)\partial^\alpha \varphi(x)dx = (-1)^{|\alpha|}\int_{\mathbb{R}^n} g(x)\varphi(x)dx, \quad \forall \varphi \in \mathscr{S}. \tag{6.11}$$

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 partial derivatives exist, they are determined almost everywhere by the defining relation (6.11).

In this section, we study a quantitative way of measuring the 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 6.7. Let $k \in \mathbb{N}_0$ and $1 \leqslant p \leqslant \infty$. The ($L^p$-)***Sobolev space*** of order k (on $\mathbb{R}^n$) is defined by

$$W^{k,p}(\mathbb{R}^n) = \{f \in L^p(\mathbb{R}^n) : \partial^\alpha f \in L^p(\mathbb{R}^n) \text{ for all } |\alpha| \leqslant k\},$$

where $\partial^\alpha f$ must be understood in the sense of $\mathscr{S}'(\mathbb{R}^n)$, i.e., (6.11). Moreover, we define

$$\|f\|_{W^{k,p}(\mathbb{R}^n)} = \begin{cases} \left(\sum_{|\alpha|\leqslant k} \|\partial^\alpha f\|_p^p\right)^{1/p}, & \text{if } 1 \leqslant p < \infty, \\ \max_{|\alpha|\leqslant k} \|\partial^\alpha f\|_\infty, & \text{if } p = \infty. \end{cases}$$

where $\partial^{(0,\ldots,0)} f = f$.

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 $\varphi \in \mathscr{S}$. 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 $f_m \to f$ in $W^{k,p}(\mathbb{R}^n)$ and proves the completeness of this space.

Now, we generalize the Riesz and Bessel potentials to any $s \in \mathbb{R}$ by

$$I^s f = \mathscr{F}^{-1}(|\xi|^s \widehat{f}), \quad f \in \mathscr{S}'(\mathbb{R}^n),\ 0 \notin \operatorname{supp} \widehat{f}, \tag{6.12}$$

$$J^s f = \mathscr{F}^{-1}(\langle \xi \rangle^s \widehat{f}), \quad f \in \mathscr{S}'(\mathbb{R}^n). \tag{6.13}$$

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.

Observe that the condition $0 \notin \operatorname{supp} \widehat{f}$ in (6.12) induces that $\|I^s f\|_p$ does not satisfy the condition of the norms when $s \in \mathbb{N}$, since for $k > m \in \mathbb{N}$ we have $I^k P(x) = 0$ in $\mathscr{S}'$ for any $P \in \mathscr{P}_m$ where $\mathscr{P}_m$ denotes the set of all polynomials of degree less than or equal to m. Indeed, we have for any $\alpha \in \mathbb{N}_0^n$ with $|\alpha| = m < k$ and any $g \in \mathscr{S}$

$$\begin{aligned}\int_{\mathbb{R}^n} (I^k x^\alpha) g(x) dx =& \int_{\mathbb{R}^n} x^\alpha \widehat{|\xi|^k \check{g}}(x) dx \\ =& \int_{\mathbb{R}^n} e^{ix\cdot 0} i^{-|\alpha|} \widehat{\partial_\xi^\alpha (|\xi|^k \check{g})}(x) dx \\ =& (2\pi)^{n/2} i^{-|\alpha|} \left[\partial_\xi^\alpha (|\xi|^k \check{g})\right](0) = 0.\end{aligned}$$

It is not good to focus upon $\mathscr{S}'(\mathbb{R}^n)$ when we consider the homogeneous spaces. We need to work on the quotient space $\mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n)$, where $\mathscr{P}$ denotes the set of all polynomials. Generally speaking, it is slightly nasty to consider the quotient space; handling the representative is not so intuitive. Therefore, we seek to find an expression of the quotient $\mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n)$. From this standpoint, we give the following definition.

Definition 6.8. Define the ***Lizorkin function space*** $\dot{\mathscr{S}}(\mathbb{R}^n)$[a]

$$\dot{\mathscr{S}}(\mathbb{R}^n) = \left\{ f \in \mathscr{S}(\mathbb{R}^n) : \int_{\mathbb{R}^n} x^\alpha f(x) dx = 0,\ \forall \alpha \in \mathbb{N}_0^n \right\}, \tag{6.14}$$

with the topology induced by $\mathscr{S}(\mathbb{R}^n)$. The ***Lizorkin distribution space*** $\dot{\mathscr{S}}'(\mathbb{R}^n)$ is the topological dual space $\dot{\mathscr{S}}(\mathbb{R}^n)$.

[a] It also uses the symbols $\mathscr{S}_0$ or $\mathscr{S}_\infty$ in other references, e.g., [Jaw77; Saw18].

The main advantage of defining the class $\dot{\mathscr{S}}$ is that for given $f \in \dot{\mathscr{S}}$, the function given by $g = \mathscr{F}^{-1}[|\xi|^\alpha \widehat{f}]$ is in $\dot{\mathscr{S}}$. In fact, for $f \in \mathscr{S}$,

$$f \in \dot{\mathscr{S}} \iff (\partial^\alpha \widehat{f})(0) = 0, \ \forall \alpha \in \mathbb{N}_0^n.$$

We have the following fundamental theorem.

Theorem 6.9. *As a linear space, we have the following isomorphism*

$$\dot{\mathscr{S}}'(\mathbb{R}^n) = \mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n).$$

Proof. For each $u \in \mathscr{S}'(\mathbb{R}^n)$, let $J(u)$ be the restriction of u on the subspace $\dot{\mathscr{S}}(\mathbb{R}^n)$ of $\mathscr{S}(\mathbb{R}^n)$. Then J is a linear mapping from $\mathscr{S}'(\mathbb{R}^n)$ to $\dot{\mathscr{S}}'(\mathbb{R}^n)$.

First, we claim that the kernel of J is exactly $\mathscr{P}(\mathbb{R}^n)$, i.e., $\ker(J) = \{u \in \mathscr{S}' : J(u) = 0 \text{ in } \dot{\mathscr{S}}'\} = \mathscr{P}$. In fact, if $\langle u, \phi \rangle = 0$ for all $\phi \in \dot{\mathscr{S}}(\mathbb{R}^n)$, then $\langle \widehat{u}, \check{\phi} \rangle = 0$ for all $\phi \in \dot{\mathscr{S}}(\mathbb{R}^n)$, i.e., $\langle \widehat{u}, \psi \rangle = 0$ for all $\psi \in \mathscr{S}(\mathbb{R}^n)$ supported in $\mathbb{R}^n \setminus \{0\}$. It follows that $\widehat{u}$ is supported at the origin and thus u must be a polynomial by Corollary 3.43. This proves that the kernel of the mapping J is $\mathscr{P}(\mathbb{R}^n)$.

We also claim that the range of J is the entire $\dot{\mathscr{S}}'(\mathbb{R}^n)$. Indeed, given $v \in \dot{\mathscr{S}}'(\mathbb{R}^n)$, v is a linear functional on $\dot{\mathscr{S}}(\mathbb{R}^n)$, which is a subspace of the vector space $\mathscr{S}$, and $|\langle v, \varphi \rangle| \leqslant p(\varphi)$ for all $\varphi \in \dot{\mathscr{S}}$, where $p(\varphi)$ is equal to a constant times a finite sum of Schwartz seminorms of φ. By the Hahn-Banach theorem, v has an extension V on $\mathscr{S}$ such that $|\langle V, \Phi \rangle| \leqslant p(\Phi)$ for all $\Phi \in \mathscr{S}$. Then, $J(V) = v$, and this shows that J is surjective.

Combining these two facts, we conclude that there is an identification

$$\mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n) = \dot{\mathscr{S}}'(\mathbb{R}^n),$$

as claimed. ❑

In view of the identification in Theorem 6.9, we have that $u_j \to u$ in $\dot{\mathscr{S}}'$ iff u_j, u are elements of $\dot{\mathscr{S}}'$ and

$$\langle u_j, \phi \rangle \to \langle u, \phi \rangle$$

as $j \to \infty$ for all $\phi \in \dot{\mathscr{S}}$. Note that convergence in $\mathscr{S}$ implies convergence in $\dot{\mathscr{S}}$, and consequently, convergence in $\mathscr{S}'$ implies convergence in $\dot{\mathscr{S}}'$.

The Fourier transform of $\dot{\mathscr{S}}(\mathbb{R}^n)$ functions can be multiplied by $|\xi|^s$, $s \in \mathbb{R}$, and still be smooth and vanish to infinite order at zero.

Indeed, let $\phi \in \dot{\mathscr{S}}(\mathbb{R}^n)$. Then, we show that $\partial_j(|\xi|^s \widehat{\phi})(0)$ exists. Since every Taylor polynomial of $\widehat{\phi}$ at zero is identically equal to zero, it follows from Taylor's theorem[2] that $|\widehat{\phi}(\xi)| \leqslant C_M |\xi|^M$ for every $M \in \mathbb{N}_0$, whenever

[2] Let $f : U \to \mathbb{R}$ be a real-valued function defined on an open subset of $\mathbb{R}^n$. Suppose that $f \in \mathfrak{C}^{k+1}(U)$, let P be a point of U such that $B(P, \delta) \subset U$ for some $\delta > 0$. For any $h \in \mathbb{R}^n$ with $|h| < \delta$, there exists a real number $c = c_{P,h} \in [0, 1]$ such that

$$f(P+h) = \left(\sum_{i=0}^{k} \frac{1}{i!} H_i(f)(P)(h) \right) + \frac{1}{(k+1)!} H_{k+1}(f)(P+ch)(h),$$

ξ lies in a compact set. Consequently, if $M > 1 - s$,

$$\lim_{t \to 0} \frac{|te_j|^s \widehat{\phi}(te_j)}{t} = 0,$$

where e_j is the vector with 1 in the jth entry and zero elsewhere. This shows that all partial derivatives of $|\xi|^s \widehat{\phi}(\xi)$ at zero exist and are equal to zero.

By induction, we assume that $\partial^\alpha(|\xi|^s \widehat{\phi}(\xi))(0) = 0$, and we need to prove that

$$\partial_j \partial^\alpha (|\xi|^s \widehat{\phi}(\xi))(0)$$

also exists and equals zero. Applying Leibniz's rule, we express $\partial^\alpha(|\xi|^s \widehat{\phi}(\xi))$ as a finite sum of derivatives of $|\xi|^s$ times derivatives of $\widehat{\phi}(\xi)$. However, for each $|\beta| \leqslant |\alpha|$, we have $|\partial^\beta \widehat{\phi}(\xi)| \leqslant C_{M,\beta} |\xi|^M$ for all $M \in \mathbb{N}_0$ whenever $|\xi| \leqslant 1$. Picking $M > |\alpha| + 1 - s$ and using the fact that $|\partial^{\alpha-\beta}(|\xi|^s)| \leqslant C_\alpha |\xi|^{s-|\alpha|+|\beta|}$, we deduce that $\partial_j \partial^\alpha(|\xi|^s \widehat{\phi}(\xi))(0)$ also exists and equals zero.

We have now proved that $\mathscr{F}^{-1}(|\xi|^s \widehat{\phi}(\xi)) \in \dot{\mathscr{S}}$ for $\phi \in \dot{\mathscr{S}}$ and all $s \in \mathbb{R}$. This allows us to introduce the operation of multiplication by $|\xi|^s$ on the Fourier transforms of distributions modulo polynomials. For $s \in \mathbb{R}$ and $u \in \dot{\mathscr{S}}'(\mathbb{R}^n)$, we define another distribution $\mathscr{F}^{-1}(|\xi|^s \widehat{u}) \in \dot{\mathscr{S}}'(\mathbb{R}^n)$ by setting for all $\phi \in \dot{\mathscr{S}}(\mathbb{R}^n)$

$$\langle \mathscr{F}^{-1}(|\cdot|^s \widehat{u}), \phi \rangle = \langle u, \widehat{|\cdot|^s \check{\phi}} \rangle.$$

This definition is consistent with the corresponding operations on functions and makes sense since $\phi \in \dot{\mathscr{S}}$ implies that $\widehat{|\cdot|^s \check{\phi}}$ also lies in $\dot{\mathscr{S}}(\mathbb{R}^n)$, and thus, the action of u on this function is defined.

Moreover, recall $\langle \xi \rangle = (1 + |\xi|^2)^{1/2}$, since for any $s \in \mathbb{R}$ the function $\langle \xi \rangle^s$ is a smooth function satisfying

$$|\partial_\xi^\alpha \langle \xi \rangle^s| \leqslant C_{s,\alpha} (1 + |\xi|)^{s-|\alpha|} \quad \text{for all } \xi \in \mathbb{R}^n$$

for all $\alpha \in \mathbb{N}_0^n$ and some $C_{s,\alpha} > 0$. Thus, $\langle \xi \rangle^s \in \mathcal{C}^\infty_{\text{poly}}(\mathbb{R}^n)$. Hence, $\langle \xi \rangle^s \widehat{f}(\xi) \in \mathscr{S}(\mathbb{R}^n)$ for all $f \in \mathscr{S}(\mathbb{R}^n)$ from Proposition 3.11. By duality $\langle \xi \rangle^s \widehat{f} \in \mathscr{S}'(\mathbb{R}^n)$ for all $f \in \mathscr{S}'(\mathbb{R}^n)$. Therefore, $J^s : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is well-defined for any $s \in \mathbb{R}$.

Next, we shall extend the spaces $W^{k,p}(\mathbb{R}^n)$ to the case where the number k is real.

---

where $H_i(f)(P)(h) = \sum_{|\alpha|=i} \binom{i}{\alpha} (\partial^\alpha f)(P) h^\alpha$ for $i \in \mathbb{N}$ and $H_0(f)(P)(h) = f(P)$ with $\binom{i}{\alpha} = \frac{i!}{\alpha_1! \cdots \alpha_n!}$ for $\alpha \in \mathbb{N}_0^n$.

Definition 6.10. Let $s \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$. We write

$$\|f\|_{\dot{H}_p^s} = \|I^s f\|_p, \quad \|f\|_{H_p^s} = \|J^s f\|_p.$$

Then, the ***homogeneous Sobolev space*** $\dot{H}_p^s(\mathbb{R}^n)$ is defined by

$$\dot{H}_p^s(\mathbb{R}^n) = \left\{ f \in \dot{\mathscr{S}}'(\mathbb{R}^n) : \|f\|_{\dot{H}_p^s} < \infty \right\},$$

and the ***nonhomogeneous Sobolev space*** $H_p^s(\mathbb{R}^n)$ is defined by

$$H_p^s(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) : \|f\|_{H_p^s} < \infty \right\}.$$

If $p = 2$, we denote $\dot{H}_2^s(\mathbb{R}^n)$ by $\dot{H}^s(\mathbb{R}^n)$ and $H_2^s(\mathbb{R}^n)$ by $H^s(\mathbb{R}^n)$ for simplicity.

It is clear that space $H_p^s(\mathbb{R}^n)$ 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_p^s. Then, by the completeness of L^p, there exists a $g \in L^p$ such that

$$\|f_m - J^{-s} g\|_{H_p^s} = \|J^s f_m - g\|_p \to 0, \quad \text{as } m \to \infty.$$

Clearly, $J^{-s} g \in \mathscr{S}'$ and thus H_p^s is complete.

We give some elementary results about Sobolev spaces.

Theorem 6.11. *Let $s \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$; then, we have*

(i) *$\mathscr{S}$ is dense in H_p^s, $1 \leqslant p < \infty$.*

(ii) *$H_p^{s+\varepsilon} \hookrightarrow H_p^s$, $\forall \varepsilon > 0$.*

(iii) *$H_p^s \hookrightarrow L^\infty$, $\forall s > n/p$.*

(iv) *Suppose $1 < p < \infty$ and $s \geqslant 1$. Then $f \in H_p^s(\mathbb{R}^n)$ iff $f \in H_p^{s-1}(\mathbb{R}^n)$ and for each j, $\frac{\partial f}{\partial x_j} \in H_p^{s-1}(\mathbb{R}^n)$. Moreover, the two norms are equivalent:*

$$\|f\|_{H_p^s} \sim \|f\|_{H_p^{s-1}} + \sum_{j=1}^{n} \left\| \frac{\partial f}{\partial x_j} \right\|_{H_p^{s-1}}.$$

(v) *$H_p^k(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$, $1 < p < \infty$, $\forall k \in \mathbb{N}$.*

Proof. (i) Take $f \in H_p^s$, 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

$$\|f - J^{-s} g\|_{H_p^s} = \|J^s f - g\|_p$$

is smaller than any given positive number. Since $J^{-s} g \in \mathscr{S}$, $\mathscr{S}$ is dense in H_p^s.

(ii) Suppose that $f \in H_p^{s+\varepsilon}$. By part (i) in Theorem 6.6, we see that J_ε maps L^p into L^p with norm 1 for $\varepsilon > 0$. From this, we obtain the result since

$$\|f\|_{H_p^s} = \|J^s f\|_p = \|J^{-\varepsilon} J^{s+\varepsilon} f\|_p = \|J_\varepsilon J^{s+\varepsilon} f\|_p \leqslant \|J^{s+\varepsilon} f\|_p = \|f\|_{H_p^{s+\varepsilon}}.$$

(iii) By Young's inequality, the definition of the kernel $G_s(x)$ and part (iv) of Proposition 6.5, we obtain for $s > 0$

$$\begin{aligned}\|f\|_\infty =&\|\mathscr{F}^{-1}[\langle\xi\rangle^{-s} \langle\xi\rangle^s \widehat{f}]\|_\infty \\ =&(2\pi)^{-n/2}\|\mathscr{F}^{-1} \langle\xi\rangle^{-s} * J^s f\|_\infty \\ \leqslant&(2\pi)^{-n/2}\|\mathscr{F}^{-1} \langle\xi\rangle^{-s}\|_{p'}\|J^s f\|_p \\ =&\|G_s(x)\|_{p'}\|f\|_{H_p^s} \leqslant C\|f\|_{H_p^s}.\end{aligned}$$

(iv) From the Mikhlin multiplier theorem, we can obtain $\xi_j \langle\xi\rangle^{-1} \in \mathcal{M}_p$ for $1 < p < \infty$, and thus,

$$\begin{aligned}\left\|\frac{\partial f}{\partial x_j}\right\|_{H_p^{s-1}} =&\|\mathscr{F}^{-1}[\langle\xi\rangle^{(s-1)} (i\xi_j)\widehat{f}]\|_p \\ =&\|\mathscr{F}^{-1}[\langle\xi\rangle^{-1} \xi_j \langle\xi\rangle^s \widehat{f}]\|_p \\ =&(2\pi)^{-n/2}\|\mathscr{F}^{-1}[\langle\xi\rangle^{-1} \xi_j] * J^s f\|_p \\ \leqslant&C\|J^s f\|_p = C\|f\|_{H_p^s}.\end{aligned}$$

Combined with $\|f\|_{H_p^{s-1}} \leqslant \|f\|_{H_p^s}$, we obtain

$$\|f\|_{H_p^{s-1}} + \sum_{j=1}^n \left\|\frac{\partial f}{\partial x_j}\right\|_{H_p^{s-1}} \leqslant C\|f\|_{H_p^s}.$$

Now, we prove the converse inequality. We use the Mikhlin multiplier theorem once more and an auxiliary function $0 \leqslant \chi \in \mathcal{C}^\infty(\mathbb{R})$ with $\chi(x) = 1$ for $|x| > 2$ and $\chi(x) = 0$ for $|x| < 1$. We obtain

$$\langle\xi\rangle \left(1 + \sum_{j=1}^n \chi(\xi_j)|\xi_j|\right)^{-1} \in \mathcal{M}_p,\ \chi(\xi_j)|\xi_j|\xi_j^{-1} \in \mathcal{M}_p,\ 1 < p < \infty,$$

and then

$$\begin{aligned}\|f\|_{H_p^s} =&\|J^s f\|_p = \|\mathscr{F}^{-1}[\langle\xi\rangle \widehat{J^{s-1} f}]\|_p \\ \leqslant&C\|\mathscr{F}^{-1}[(1 + \sum_{j=1}^n \chi(\xi_j)|\xi_j|)\widehat{J^{s-1} f}]\|_p \\ \leqslant&C\|f\|_{H_p^{s-1}} + C\sum_{j=1}^n \left\|\mathscr{F}^{-1}(\chi(\xi_j)|\xi_j|\xi_j^{-1}\widehat{J^{s-1}\tfrac{\partial f}{\partial x_j}})\right\|_p \\ \leqslant&C\|f\|_{H_p^{s-1}} + C\sum_{j=1}^n \left\|\frac{\partial f}{\partial x_j}\right\|_{H_p^{s-1}}.\end{aligned}$$

Thus, we have obtained the desired result.

(v) It is obvious that $W^{0,p} = H_p^0 = L^p$ for $k = 0$. However, from part (iv), if $k \geqslant 1$, then $f \in H_p^k$ iff f and $\frac{\partial f}{\partial x_j} \in H_p^{k-1}$, $j = 1, ..., n$. Thus, we can extend the identity of $W^{k,p} = H_p^k$ from $k = 0$ to $k = 1, 2, \ldots$. ❑

Due to (ii) of Theorem 6.11, we will also use the notation $H^{-\infty} = \bigcup_s H^s$ and $H^\infty = \bigcap_s H^s$. The inclusion $\mathscr{S} \subset H^\infty \subset H^{-\infty} \subset \mathscr{S}'$ are immediate, they are all strict. It is not true that $\mathscr{S} = H^\infty$ by taking $f(x) = \langle x \rangle^{-2n}$, which satisfies $f \in H^\infty$ but $f \notin \mathscr{S}$; nor is it true that $H^{-\infty} = \mathscr{S}'$ since the control of the growth of f at infinity is not sufficient, or on the Fourier side the smoothness of $\hat{f}$ is not surfficient, i.e., $\hat{f}$ is not a function.

We continue with the Sobolev embedding theorem.

Theorem 6.12 (Sobolev embedding theorem). *Let $1 < p \leqslant p_1 < \infty$ and $s, s_1 \in \mathbb{R}$. Assume that $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$. The following conclusions hold*

$$H_p^s \hookrightarrow H_{p_1}^{s_1}, \quad \dot{H}_p^s \hookrightarrow \dot{H}_{p_1}^{s_1}.$$

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 (ii) of Theorem 6.6, we get

$$\|f\|_{H_{p_1}^{s_1}} = \|J^{s_1} f\|_{p_1} = \|J^{s_1 - s} J^s f\|_{p_1} = \|J_{s - s_1} J^s f\|_{p_1} \leqslant C \|J^s f\|_p = C \|f\|_{H_p^s}.$$

Similarly, we can show the homogeneous case. ❑

Theorem 6.13. *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 6.14. *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 $(L^p)' = L^{p'}$ if $1 \leqslant p < \infty$. ❑

Exercises

Exercise 6.1. For $0 \leqslant s < n$, define the fractional maximal function

$$M^s f(x) = \sup_{t>0} \frac{1}{(V_n t^n)^{\frac{n-s}{n}}} \int_{|y| \leqslant t} |f(x - y)| dy$$

where V_n is the volume of the unit ball in $\mathbb{R}^n$.

(i) Show that for some constant C we have

$$M^s f \leqslant C I_s f$$

for all $f \geqslant 0$ and conclude that M^s maps L^p into L^q whenever the Riesz potential I_s of order s does.

(ii) Let $s > 0$, $1 < p < \frac{n}{s}$, $1 \leqslant q \leqslant \infty$ such that $\frac{1}{r} = \frac{1}{p} - \frac{s}{n} + \frac{sp}{nq}$. Show that there is a constant $C > 0$ (depending on the previous parameters) such that for all positive functions f we have

$$\|I_s f\|_r \leqslant C \|M^{n/p} f\|_q^{sp/n} \|f\|_p^{1-sp/n}.$$

Hint For $f \neq 0$, write $I_s f = I_1 + I_2$, where

$$I_1 = \int_{|x-y|\leqslant\delta} f(y)|y|^{s-n} dy, \quad I_2 = \int_{|x-y|>\delta} f(y)|y|^{s-n} dy.$$

Show that $I_1 \leqslant C\delta^s M^0(f)$ and that $I_2(f) \leqslant C\delta^{s-n/p} M^{n/p} f$. Optimize over $\delta > 0$ to obtain $I_s f \leqslant C(M^{n/p} f)^{sp/n} (M^0 f)^{1-sp/n}$ from which the required conclusion follows easily.

Exercise 6.2. Find all $s \in \mathbb{R}$ such that the Dirac distribution δ_0 is in $H^s(\mathbb{R}^n)$.

Exercise 6.3. Let $1 < p < \infty$ and $s \in \mathbb{N}$.

(i) Suppose that $f \in H_p^s(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Prove that φf is also an element of $H_p^s(\mathbb{R}^n)$.

(ii) Let v be a function whose Fourier transform is a bounded and compactly supported function. Prove that if f is in $H^s(\mathbb{R}^n)$, then so is vf.

Exercise 6.4. Consider the equation

$$u_t + \Delta^2 u = 0 \quad \text{in} \quad (0, T) \times \mathbb{R}^n,$$

with the initial condition

$$u|_{t=0} = u_0 \quad \text{on } \mathbb{R}^n,$$

where $u_0 \in H^s(\mathbb{R}^n)$ for an $s > 0$. Prove that there exists a solution u belonging to the space $\mathcal{C}([0, T]; H^s(\mathbb{R}^n))$.

7

Hardy and BMO Spaces

In this chapter, we introduce the Hardy and BMO spaces, and the duality between them. We also introduce the Carleson measures and their relations with BMO functions.

§7.1 Hardy spaces

Hardy spaces are function spaces designed to be better suited to some application than L^1. We consider atomic Hardy spaces in this section.

Definition 7.1 (p-atom)**.** Let Q be a cube in $\mathbb{R}^n$, $1 < p \leqslant \infty$ and p' be its conjugate exponent. A Lebesgue measurable function $a : Q \to \mathbb{C}$ is called a ***p*-atom** on Q if

(i) $\operatorname{supp} a \subset Q$,
(ii) $\|a\|_p \leqslant |Q|^{-1/p'}$,
(iii) $\int_Q a\,dx = 0$.

We denote the collection of p-atoms on Q by $\mathcal{A}_Q^p$ and $\mathcal{A}^p = \bigcup_Q \mathcal{A}_Q^p$.

Remark 7.2. Note that (i) along with (ii) implies that $\|a\|_1 \leqslant 1$.

Definition 7.3 ($\mathcal{H}^{1,p}$)**.** Let $1 < p \leqslant \infty$ and $f \in L^1(\mathbb{R}^n)$. We say that $f \in \mathcal{H}^{1,p}$ if there exist p-atoms $\{a_i\}_{i\in\mathbb{N}} \subset \mathcal{A}^p$ and $(\lambda_i)_{i\in\mathbb{N}} \in \ell^1(\mathbb{N})$ such that

$$f = \sum_{i=1}^{\infty} \lambda_i a_i, \text{ a.e.,}$$

and define the norm

$$\|f\|_{\mathcal{H}^{1,p}} = \inf\left\{\sum |\lambda_i| : f = \sum \lambda_i a_i\right\},$$

where the infimum is taken over all possible representations of f.

Remark 7.4. From $(\lambda_i) \in \ell^1$ and $\|a_i\|_1 \leqslant 1$, it follows that $\sum \lambda_i a_i$ converges in $L^1(\mathbb{R}^n)$. For $f \in L^1(\mathbb{R}^n)$, then the equality $f = \sum \lambda_i a_i$ holds in $L^1(\mathbb{R}^n)$. We could certainly give a similar definition in the more general setting of $f \in \mathscr{S}'(\mathbb{R}^n)$. Then, we ask about the convergence of the series in the sense of $\mathscr{S}'(\mathbb{R}^n)$ in the definition. As $L^1(\mathbb{R}^n)$ embeds in $\mathscr{S}'(\mathbb{R}^n)$, it coincides with the L^1 function $\sum \lambda_i a_i$ after identification.

We have the following completeness and embedding relations, whose proofs are left to the reader.

Proposition 7.5. (i) *For* $1 < p \leqslant \infty$, $(\mathcal{H}^{1,p}, \|\cdot\|_{\mathcal{H}^{1,p}})$ *is a Banach space.*
(ii) *For* $1 < p < r < \infty$, *we have*

$$\mathcal{H}^{1,\infty} \subset \mathcal{H}^{1,r} \subset \mathcal{H}^{1,p} \subset L^1.$$

We begin with the construction of dyadic cubes on $\mathbb{R}^n$.

Definition 7.6 (Dyadic cubes)**.** Let $[0,1)^n$ be the reference cube and let $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^n$. Then define the ***dyadic cube*** of generation j with lower left corner $2^{-j}k$

$$Q_{j,k} = \left\{ x \in \mathbb{R}^n : 2^j x - k \in [0,1)^n \right\},$$

the set of generation j dyadic cubes

$$\mathcal{Q}_j = \left\{ Q_{j,k} : k \in \mathbb{Z}^n \right\},$$

and the set of all dyadic cubes

$$\mathcal{Q} = \bigcup_{j \in \mathbb{Z}} \mathcal{Q}_j = \left\{ Q_{j,k} : j \in \mathbb{Z} \text{ and } k \in \mathbb{Z}^n \right\}.$$

We define the length of a cube to be its side length $\ell(Q_{j,k}) = 2^{-j}$.

The crucial property of dyadic cubes is the nesting property: if two dyadic cubes overlap, then one must contain the other. This leads to a dyadic version of Vitali-type covering lemma (Lemma 2.5):

Lemma 7.7 (Dyadic Vitali-type covering lemma)**.** *Let* $Q_1, \cdots, Q_N$ *be a finite collection of dyadic cubes. Then there is a subcollection* $Q_{n_1}, \cdots, Q_{n_k}$ *of disjoint cubes such that*

$$Q_{n_1} \bigcup \cdots \bigcup Q_{n_k} = Q_1 \bigcup \cdots \bigcup Q_N.$$

Proof. Take the Q_{n_i} to be the maximal dyadic cubes in $Q_1, \cdots, Q_N$, i.e., the cubes that are not contained in any other cubes in this collection. The nesting property then ensures that they are disjoint and cover all of $Q_1, \cdots, Q_N$ between them. ❑

If we define the ***dyadic maximal function***

$$M_\Delta f(x) := \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| dy$$

where Q ranges over the dyadic cubes that contain x; then, the same argument as before gives the ***dyadic Hardy-Littlewood maximal inequality***

$$\|M_\Delta f\|_{L^{1,\infty}(\mathbb{R}^n)} \leqslant \|f\|_{L^1(\mathbb{R}^n)}, \tag{7.1}$$

(with no constant loss whatsoever!) which then leads via Marcinkiewicz interpolation to

$$\|M_\Delta f\|_{L^p(\mathbb{R}^n)} \leqslant C_p \|f\|_{L^p(\mathbb{R}^n)},$$

for $1 < p \leqslant \infty$.

If we consider a $Q_0 \in \mathcal{Q}$, let $\mathcal{D}(Q_0)$ be the collection of dyadic subcubes of Q_0, then we may define

$$M_\Delta f(x) = \sup_{Q \in \mathcal{D}(Q_0), Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| dy$$

for $x \in Q_0$ and $f \in L^1_{\text{loc}}(Q_0)$. We also have the weak type $(1,1)$ and type (p,p) estimates for $p \in (1, \infty]$.

Theorem 7.8 (Equivalence of $\mathcal{H}^{1,p}$ spaces)**.** *For* $1 < p < \infty$, $\mathcal{H}^{1,\infty} = \mathcal{H}^{1,p}$ *with equivalent norms.*

Proof. (i) We establish the Calderón-Zygmund decomposition of functions. More precisely, for a given p-atom a, we show that there exists a decomposition $a = b + g$ where $b \in \mathcal{H}^{1,p}$ with $\|b\|_{\mathcal{H}^{1,p}} \leqslant 1/2$ (due to $\|a\|_{\mathcal{H}^{1,p}} \leqslant 1$) and $g \in \mathcal{H}^{1,\infty}$ with $\|g\|_{\mathcal{H}^{1,\infty}} \leqslant C(n,p)$.

Let Q be a dyadic cube in $\mathbb{R}^n$ such that $a \in \mathcal{A}^p_Q$. Let $\mathcal{D}(Q)$ denote the dyadic subcubes of Q. We have

$$\operatorname*{Avg}_Q |a|^p = \frac{1}{|Q|} \int_Q |a|^p dx \leqslant \frac{1}{|Q|^p}.$$

Fix $\alpha > 0$ with $\alpha^p > \operatorname*{Avg}_Q |a|^p$ to be chosen later and let

$$E_\alpha = \left\{ x \in Q : [(M_\Delta |a|^p)(x)]^{1/p} > \alpha \right\}.$$

If $E_\alpha \neq \varnothing$, then for $x \in E_\alpha$, there exists a $Q_i \in \mathcal{D}(Q)$ such that $x \in Q_i$ and $\operatorname*{Avg}_{Q_i} |a|^p > \alpha^p$. Let $\mathcal{B}$ be the collection of all such cubes. Since $\mathcal{Q}$ is countable, $\mathcal{B}$ is also countable. Then, for every $Q_i \in \mathcal{B}$, we have $Q_i \subset E_\alpha$; therefore, $E_\alpha = \bigcup_{Q_i \in \mathcal{B}} Q_i$. The nesting property ensures that $E_\alpha = \bigcup_{i=1}^\infty Q_i$ where $Q_i \in \mathcal{D}(Q)$ is maximal for the property that $\operatorname*{Avg}_{Q_i} |a|^p > \alpha^p$. We also have $E_\alpha \neq Q$. Set

$$b_i = (a - \operatorname*{Avg}_{Q_i} a) \chi_{Q_i}$$

and $b = \sum\limits_{i=1}^\infty b_i$. Then, let $g = a - b$.

Let us see the properties of b_i. First, $\operatorname{supp} b_i \subset \overline{Q_i}$ and

$$\int_{Q_i} b_i dx = 0.$$

Due to $\int_{Q_i} |a|^p dx = |Q_i| \operatorname*{Avg}_{Q_i} |a|^p$ and $|\operatorname*{Avg}_{Q_i} a| \leqslant (\operatorname*{Avg}_{Q_i} |a|^p)^{1/p}$, we also have

$$\|b_i\|_p \leqslant \left(\int_{Q_i} |a|^p dx\right)^{1/p} + |\operatorname*{Avg}_{Q_i} a||Q_i|^{1/p} = |Q_i|^{-1/p'}\lambda_i,$$

where $\lambda_i = 2[\operatorname*{Avg}_{Q_i} |a|^p]^{1/p}|Q_i| > 2\alpha|Q_i| > 0$. Thus, $a_i = \frac{1}{\lambda_i} b_i$ is a p-atom.

By the Hölder inequality and the dyadic maximal inequality, we obtain

$$\begin{aligned}\sum_{i=1}^{\infty} \lambda_i =&2\sum_{i=1}^{\infty} \left[\int_{Q_i} |a|^p dx\right]^{1/p} |Q_i|^{1-1/p} \leqslant 2\left[\sum_{i=1}^{\infty} \int_{Q_i} |a|^p dx\right]^{1/p} \left[\sum_{i=1}^{\infty} |Q_i|\right]^{1/p'}\\ \leqslant&2\left[\int_Q |a|^p dx\right]^{1/p} |E_\alpha|^{1/p'} \leqslant 2\left[\int_Q |a|^p dx\right]^{1/p} \left(\frac{\|a\|_p^p}{\alpha^p}\right)^{1-1/p}\\ =&2\|a\|_p^p \alpha^{1-p} \leqslant 2|Q|^{-p/p'}\alpha^{1-p} = 2\left(\frac{1}{\alpha|Q|}\right)^{p-1}.\end{aligned}$$

Now, we choose α such that

$$2\left(\frac{1}{\alpha|Q|}\right)^{p-1} = \frac{1}{2},$$

that is,

$$\alpha = \frac{c_p}{|Q|},$$

with $c_p = 4^{1/(p-1)}$. Then, it follows that $\|b\|_{\mathcal{H}^{1,p}} \leqslant 1/2$.

Next, we consider g, i.e.,

$$g = \begin{cases} a & \text{in } Q \setminus E_\alpha, \\ \operatorname*{Avg}_{Q_i} a & \text{in } Q_i, \text{ for each } i. \end{cases}$$

In $Q \setminus E_\alpha$, since $|a|^p \leqslant M_\Delta |a|^p \leqslant \alpha^p$ a.e., we have $|g| \leqslant \alpha$ a.e. In Q_i, by maximality of Q_i and the Hölder inequality, it yields

$$|\operatorname*{Avg}_{Q_i} a| \leqslant \operatorname*{Avg}_{Q_i} |a| \leqslant 2^n \operatorname*{Avg}_{\widehat{Q_i}} |a| \leqslant 2^n (\operatorname*{Avg}_{\widehat{Q_i}} |a|^p)^{1/p} \leqslant 2^n \alpha,$$

where $\widehat{Q_i}$ is the parent cube of Q_i. Hence, $|g| \leqslant 2^n\alpha$. It follows that

$$\|g\|_\infty \leqslant 2^n \alpha = \frac{2^n c_p}{|Q|}.$$

We also have $\int_Q g = \int_Q a = 0$, so $\frac{1}{2^n c_p} g \in \mathcal{A}_Q^\infty$ which implies that $g \in \mathcal{H}^{1,\infty}$ with $\|g\|_{\mathcal{H}^{1,\infty}} \leqslant 2^n c_p$.

(ii) Fix $f_0 \in \mathcal{H}^{1,p}$ with $f_0 \neq 0$. We show that there exists a decomposition $f_0 = f_1 + g^0$ with

$$\|f_1\|_{\mathcal{H}^{1,p}} \leqslant \frac{2}{3}\|f_0\|_{\mathcal{H}^{1,p}}, \quad \|g^0\|_{\mathcal{H}^{1,\infty}} \leqslant \frac{4}{3}2^n c_p \|f_0\|_{\mathcal{H}^{1,p}}.$$

In fact, for every $\varepsilon > 0$, there exists an atomic decomposition $f_0 = \sum_{i=1}^{\infty} \lambda_i a_i$ such that

$$\sum_{i=1}^{\infty} |\lambda_i| \leqslant \|f_0\|_{\mathcal{H}^{1,p}} + \varepsilon.$$

Applying (i) to each a_i to find a decomposition $a_i = b_i + g_i$ with $\|b_i\|_{\mathcal{H}^{1,p}} \leqslant 1/2$. Then, $f_1 = \sum_{i=1}^{\infty} \lambda_i b_i$ exists since $\mathcal{H}^{1,p}$ is a Banach space, and

$$\|f_1\|_{\mathcal{H}^{1,p}} \leqslant \frac{1}{2}\sum_{i=1}^{\infty} |\lambda_i| \leqslant \frac{1}{2}(\|f_0\|_{\mathcal{H}^{1,p}} + \varepsilon).$$

Thus, we may choose $\varepsilon = \|f_0\|_{\mathcal{H}^{1,p}}/3$, then $\|f_1\|_{\mathcal{H}^{1,p}} \leqslant \frac{2}{3}\|f_0\|_{\mathcal{H}^{1,p}}$. Let $g^0 = \sum_{i=1}^{\infty} \lambda_i g_i$, where the sum converges in $\mathcal{H}^{1,\infty}$ because it is a Banach space and $\|g_j\|_{\mathcal{H}^{1,\infty}} \leqslant 2^n c_p$, we find

$$\|g^0\|_{\mathcal{H}^{1,\infty}} \leqslant 2^n c_p (\|f_0\|_{\mathcal{H}^{1,p}} + \varepsilon) = \frac{4}{3}2^n c_p \|f_0\|_{\mathcal{H}^{1,p}}.$$

(iii) We iterate

$$\begin{aligned} f_0 =& f_1 + g^0, \\ f_1 =& f_2 + g^1, \\ f_2 =& f_3 + g^2, \\ \vdots =& \ \vdots \end{aligned}$$

and so for each k,

$$f_0 = f_k + g^0 + g^1 + \cdots + g^{k-1}.$$

Of course, as $k \to \infty$, $f_k \to 0$ in $\mathcal{H}^{1,p}$ due to $\|f_k\|_{\mathcal{H}^{1,p}} \leqslant \left(\frac{2}{3}\right)^k \|f_0\|_{\mathcal{H}^{1,p}}$ and $g^0 + g^1 + \cdots + g^{k-1}$ converges to g in $\mathcal{H}^{1,\infty}$ with

$$\|g\|_{\mathcal{H}^{1,\infty}} \leqslant \frac{4}{3}2^n c_p \sum_{j=0}^{\infty} \left(\frac{2}{3}\right)^j \|f_0\|_{\mathcal{H}^{1,p}} = 2^{n+2p/(p-1)}\|f_0\|_{\mathcal{H}^{1,p}}.$$

By Proposition 7.5, convergence in $\mathcal{H}^{1,\infty}$ implies convergence in $\mathcal{H}^{1,p}$ and thus $f_0 = g$ and $f_0 \in \mathcal{H}^{1,\infty}$. ❑

This motivates the following definition.

Definition 7.9 (Hardy space $\mathcal{H}^1$)**.** We define ***Hardy space*** $\mathcal{H}^1$ to be any $\mathcal{H}^{1,p}$ for $1 < p \leqslant \infty$ with the corresponding norm.

Now, we state a characterization of this space.

Definition 7.10. Let $a, b > 0$. Let $\Phi \in \mathscr{S}(\mathbb{R}^n)$ and $f \in \mathscr{S}'(\mathbb{R}^n)$. We define the ***smooth maximal function of f with respect to Φ*** as

$$M(f;\Phi)(x) = \sup_{t>0} |(\Phi_t * f)(x)|.$$

Then, we recall the following result, and one can see [Gra14b, Theorem 2.1.4] and subsequent discussion therein.

Theorem 7.11. *For any $\Phi \in \mathscr{S}$ with $\int_{\mathbb{R}^n} \Phi(x)dx = 1$ and any bounded $f \in \mathscr{S}'(\mathbb{R}^n)$, the following quasinorms are equivalent*

$$\|f\|_{\mathcal{H}^1} \sim \|M(f;\Phi)\|_1,$$

with constants that depend only on Φ and n.

§7.2 BMO spaces

§7.2.1 Definition and basic properties of BMO

Functions of bounded mean oscillation were introduced by F. John and L. Nirenberg [JN61] in connection with differential equations.

Definition 7.12. The ***mean oscillation*** of $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ over a cube $Q \subset \mathbb{R}^n$ is defined as

$$\tilde{f}_Q = \frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f|dx,$$

where $\operatorname*{Avg}_Q f$ is the average value of f on the cube Q, i.e.,

$$\operatorname*{Avg}_Q f = \frac{1}{|Q|}\int_Q f(x)dx.$$

Definition 7.13 (BMO)**.** For f a complex-valued locally integrable function of $\mathbb{R}^n$, set

$$\|f\|_{\mathrm{BMO}} = \sup_Q \tilde{f}_Q = \sup_Q \operatorname*{Avg}_Q |f - \operatorname*{Avg}_Q f|,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$. The function f is of bounded mean oscillation if $\|f\|_{\mathrm{BMO}} < \infty$, the set

$$\mathrm{BMO}(\mathbb{R}^n) = \{f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) : \|f\|_{\mathrm{BMO}} < \infty\}$$

is called the ***function space of bounded mean oscillation*** or the BMO *space*.

Remark 7.14. 1) $\mathrm{BMO}(\mathbb{R}^n)$ is a linear space, that is, if $f, g \in \mathrm{BMO}(\mathbb{R}^n)$ and $\lambda \in \mathbb{C}$, then $f+g$ and λf are also in $\mathrm{BMO}(\mathbb{R}^n)$ and

$$\begin{aligned}\|f+g\|_{\mathrm{BMO}} \leqslant& \|f\|_{\mathrm{BMO}} + \|g\|_{\mathrm{BMO}},\\ \|\lambda f\|_{\mathrm{BMO}} =& |\lambda| \|f\|_{\mathrm{BMO}}.\end{aligned}$$

2) $\|\cdot\|_{\mathrm{BMO}}$ is not a norm. The problem is that if $\|f\|_{\mathrm{BMO}} = 0$, this does not imply that $f = 0$ but that f is a constant. Moreover, every constant function c satisfies $\|c\|_{\mathrm{BMO}} = 0$. Consequently, functions f and $f + c$ have the same BMO norms whenever c is a constant. In the sequel, we keep in mind that elements of BMO whose difference is a constant are identified. Although $\|\cdot\|_{\mathrm{BMO}}$ is only a seminorm, we occasionally refer to it as a norm when there is no possibility of confusion.

We give a list of basic properties of BMO.

Proposition 7.15. *The following properties of the space* $\mathrm{BMO}(\mathbb{R}^n)$ *are valid:*

(i) *If* $\|f\|_{\mathrm{BMO}} = 0$, *then* f *is a.e. equal to a constant.*

(ii) $L^\infty(\mathbb{R}^n) \hookrightarrow \mathrm{BMO}(\mathbb{R}^n)$ *and* $\|f\|_{\mathrm{BMO}} \leqslant 2\|f\|_\infty$.

(iii) *Suppose that there exists an* $A > 0$ *such that for all cubes* Q *in* $\mathbb{R}^n$ *there exists a constant* c_Q *such that*

$$\sup_Q \frac{1}{|Q|} \int_Q |f(x) - c_Q| dx \leqslant A. \tag{7.2}$$

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

(iv) *For all* $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, *we have*

$$\frac{1}{2}\|f\|_{\mathrm{BMO}} \leqslant \sup_Q \frac{1}{|Q|} \inf_{c_Q} \int_Q |f(x) - c_Q| dx \leqslant \|f\|_{\mathrm{BMO}}.$$

(v) *If* $f \in \mathrm{BMO}(\mathbb{R}^n)$, $h \in \mathbb{R}^n$ *and* $\tau^h f$ *is given by* $\tau^h f(x) = f(x-h)$, *then* $\tau^h f$ *is also in* $\mathrm{BMO}(\mathbb{R}^n)$ *and*

$$\|\tau^h f\|_{\mathrm{BMO}} = \|f\|_{\mathrm{BMO}}.$$

(vi) *If* $f \in \mathrm{BMO}(\mathbb{R}^n)$ *and* $\lambda > 0$, *then the function* $\delta^\lambda(f)$ *defined by* $\delta^\lambda f(x) = f(\lambda x)$ *is also in* $\mathrm{BMO}(\mathbb{R}^n)$ *and*

$$\|\delta^\lambda f\|_{\mathrm{BMO}} = \|f\|_{\mathrm{BMO}}.$$

(vii) *If* $f \in \mathrm{BMO}(\mathbb{R}^n)$, *then so is* $|f|$. *Similarly, if* f *and* g *are real-valued BMO functions, then so are* $\max(f,g)$ *and* $\min(f,g)$. *Moreover,*

$$\begin{aligned}\||f|\|_{\mathrm{BMO}} \leqslant& 2\|f\|_{\mathrm{BMO}},\\ \|\max(f,g)\|_{\mathrm{BMO}} \leqslant& \frac{3}{2}\left(\|f\|_{\mathrm{BMO}} + \|g\|_{\mathrm{BMO}}\right),\end{aligned}$$

$$\|\min(f,g)\|_{\mathrm{BMO}} \leqslant \frac{3}{2}\left(\|f\|_{\mathrm{BMO}} + \|g\|_{\mathrm{BMO}}\right).$$

(viii) *For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, define*

$$\|f\|_{\mathrm{BMO}_{balls}} = \sup_B \frac{1}{|B|}\int_B |f(x) - \operatorname*{Avg}_B f|dx, \tag{7.3}$$

where the supremum is taken over all balls B in $\mathbb{R}^n$. Then there are positive constants c_n and C_n such that

$$c_n\|f\|_{\mathrm{BMO}} \leqslant \|f\|_{\mathrm{BMO}_{balls}} \leqslant C_n\|f\|_{\mathrm{BMO}}.$$

(ix) *Let $f \in \mathrm{BMO}$ be real valued. Then we have the following approximation by truncation. For $N > 0$, let*

$$f_N(x) = \begin{cases} N, & f(x) > N, \\ f(x), & -N \leqslant f(x) \leqslant N, \\ -N, & f(x) < -N. \end{cases}$$

Then, $f_N \in L^\infty(\mathbb{R}^n)$, $\|f_N\|_{\mathrm{BMO}} \leqslant 2\|f\|_{\mathrm{BMO}}$ and $f_N \to f$ a.e. in $\mathbb{R}^n$.

(x) *Assume f is complex valued. Then $f \in \mathrm{BMO}$ iff $\operatorname{Re} f$, $\operatorname{Im} f \in \mathrm{BMO}$ and*

$$\|\operatorname{Re} f\|_{\mathrm{BMO}}, \|\operatorname{Im} f\|_{\mathrm{BMO}} \leqslant \|f\|_{\mathrm{BMO}} \leqslant \|\operatorname{Re} f\|_{\mathrm{BMO}} + \|\operatorname{Im} f\|_{\mathrm{BMO}}.$$

Proof. To prove (i), note that f has to be a.e. equal to its average c_N over every cube $[-N,N]^n$. Since $[-N,N]^n$ is contained in $[-N-1,N+1]^n$, it follows that $c_N = c_{N+1}$ for all N. This implies the required conclusion.

To prove (ii), observe that

$$\operatorname*{Avg}_Q |f - \operatorname*{Avg}_Q f| \leqslant \operatorname*{Avg}_Q \left(|f| + |\operatorname*{Avg}_Q f|\right) \leqslant 2\operatorname*{Avg}_Q |f| \leqslant 2\|f\|_\infty.$$

For (iii), note that

$$|f - \operatorname*{Avg}_Q f| \leqslant |f - c_Q| + |\operatorname*{Avg}_Q f - c_Q| \leqslant |f - c_Q| + \frac{1}{|Q|}\int_Q |f(t) - c_Q|dt.$$

Averaging over Q and using (7.2), we obtain that $\|f\|_{\mathrm{BMO}} \leqslant 2A$.

The lower inequality in (iv) follows from the last inequality while the upper inequality is trivial.

(v) follows from $\operatorname*{Avg}_Q \tau^h f = \operatorname*{Avg}_{Q-h} f$.

For (vi), note that $\operatorname*{Avg}_Q \delta^\lambda f = \operatorname*{Avg}_{\lambda Q} f$ and thus

$$\frac{1}{|Q|}\int_Q |f(\lambda x) - \operatorname*{Avg}_Q \delta^\lambda f|dx = \frac{1}{|\lambda Q|}\int_{\lambda Q} |f(x) - \operatorname*{Avg}_{\lambda Q} f|dx.$$

The first inequality in (vii) is a consequence of the fact that

$$\begin{aligned}\left| |f(x)| - \operatorname*{Avg}_Q |f| \right| &= \left| |f(x)| - \frac{1}{|Q|}\int_Q |f(t)|dt \right| \\ &= \left| \frac{1}{|Q|}\int_Q (|f(x)| - |f(t)|)dt \right| \\ &\leqslant \left| \frac{1}{|Q|}\int_Q (|f(x) - f(t)|)dt \right| \\ &\leqslant \left| \frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f|dt + \frac{1}{|Q|}\int_Q |\operatorname*{Avg}_Q f - f(t)|dt \right| \\ &\leqslant |f - \operatorname*{Avg}_Q f| + \operatorname*{Avg}_Q |f - \operatorname*{Avg}_Q f|.\end{aligned}$$

The second and the third inequalities in (vii) follow from the first inequality in (vii) and the facts that

$$\max(f,g) = \frac{f+g+|f-g|}{2}, \quad \min(f,g) = \frac{f+g-|f-g|}{2}.$$

We now turn to (viii). Given any cube Q in $\mathbb{R}^n$, let B be the smallest ball that contains it. Then $|B|/|Q| = 2^{-n}V_n\sqrt{n^n}$ due to $|Q| = (2r)^n$ and $|B| = V_n(\sqrt{n}r)^n$, and

$$\begin{aligned}\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_B f|dx &\leqslant \frac{|B|}{|Q|}\frac{1}{|B|}\int_B |f(x) - \operatorname*{Avg}_B f|dx \\ &\leqslant \frac{V_n\sqrt{n^n}}{2^n}\|f\|_{\mathrm{BMO_{balls}}}.\end{aligned}$$

It follows from (iii) that

$$\|f\|_{\mathrm{BMO}} \leqslant 2^{1-n}V_n\sqrt{n^n}\|f\|_{\mathrm{BMO_{balls}}}.$$

To obtain the reverse conclusion, given any ball B find the smallest cube Q that contains it, with $|B| = V_n r^n$ and $|Q| = (2r)^n$, and argue similarly using a version of (iii) for the space $\mathrm{BMO_{balls}}$.

For (ix), let Q be a cube and $x, y \in Q$. Then, $|f_N(x) - f_N(y)| \leqslant |f(x) - f(y)|$ and

$$f_N(x) - \operatorname*{Avg}_Q f_N = \frac{1}{|Q|}\int_Q (f_N(x) - f_N(y))dy.$$

Thus, it follows

$$\begin{aligned}&\frac{1}{|Q|}\int_Q |f_N(x) - \operatorname*{Avg}_Q f_N|dx \\ \leqslant&\frac{1}{|Q|^2}\int_Q\int_Q |f_N(x) - f_N(y)|dxdy \\ \leqslant&\frac{1}{|Q|^2}\int_Q\int_Q |f(x) - \operatorname*{Avg}_Q f + \operatorname*{Avg}_Q f - f(y)|dxdy\end{aligned}$$

$$\leqslant \frac{1}{|Q|^2}\int_Q\int_Q [|f(x)-\operatorname*{Avg}_Q f| + |\operatorname*{Avg}_Q f - f(y)|]dxdy$$
$$\leqslant 2\|f\|_{\mathrm{BMO}}.$$

Taking the supremum over all Q yields the desired result.

For (x), we leave it as an exercise. ❑

From Proposition 7.15 (ii), we know $L^\infty \hookrightarrow \mathrm{BMO}$. However, the converse is false; that is, $L^\infty(\mathbb{R}^n)$ is a proper subspace of $\mathrm{BMO}(\mathbb{R}^n)$. A simple example that already typifies some of the essential properties of BMO is given by the following.

Example 7.16. $\ln|x| \in \mathrm{BMO}(\mathbb{R}^n)$.

Solution. For every $x_0 \in \mathbb{R}^n$ and $R > 0$, we find a constant $C_{x_0,R}$ such that the average of $|\ln|x| - C_{x_0,R}|$ over the ball $\overline{B(x_0,R)} = \{x \in \mathbb{R}^n : |x-x_0| \leqslant R\}$ is uniformly bounded. The constant $C_{x_0,R} = \ln|x_0|$ if $|x_0| > 2R$ and $C_{x_0,R} = \ln R$ if $|x_0| \leqslant 2R$ has this property. Indeed, if $|x_0| > 2R$, then

$$\frac{1}{V_nR^n}\int_{|x-x_0|\leqslant R} |\ln|x| - C_{x_0,R}|\,dx$$
$$=\frac{1}{V_nR^n}\int_{|x-x_0|\leqslant R}\left|\ln\frac{|x|}{|x_0|}\right|dx \leqslant \max\left(\ln\frac{3}{2},\left|\ln\frac{1}{2}\right|\right) = \ln 2,$$

since $\frac{1}{2}|x_0| \leqslant |x| \leqslant \frac{3}{2}|x_0|$ when $|x-x_0| \leqslant R$ and $|x_0| > 2R$. Additionally, if $|x_0| \leqslant 2R$, then

$$\frac{1}{V_nR^n}\int_{|x-x_0|\leqslant R} |\ln|x| - C_{x_0,R}|\,dx$$
$$=\frac{1}{V_nR^n}\int_{|x-x_0|\leqslant R}\left|\ln\frac{|x|}{R}\right|dx \leqslant \frac{1}{V_nR^n}\int_{|x|\leqslant 3R}\left|\ln\frac{|x|}{R}\right|dx$$
$$=\frac{1}{V_n}\int_{|x|\leqslant 3}|\ln|x||\,dx = \frac{\omega_{n-1}}{V_n}\int_0^3 r^{n-1}|\ln r|dr$$
$$=n\int_0^1(-1)r^n\ln r\frac{dr}{r} + n\int_1^3 r^{n-1}\ln r dr \quad (\text{let } \ln r = -t)$$
$$\leqslant n\int_0^\infty te^{-nt}dt + n\ln 3\int_1^3 r^{n-1}dr \leqslant \frac{1}{n} + 3^n\ln 3.$$

Thus, $\ln|x|$ is in BMO in view of Proposition 7.15 (viii). ❑

Example 7.17. Let

$$f(x) = \begin{cases} \ln|x|, & x \leqslant 0,\\ -\ln|x|, & x > 0.\end{cases}$$

Since f is an odd function, we have $\operatorname*{Avg}_{[-a,a]} f = \frac{1}{2a}\int_{-a}^a f(x)dx = 0$ for every interval $[-a,a] \subset \mathbb{R}$. For $0 < a < 1$, we obtain

$$\|f\|_{\mathrm{BMO}} \geqslant \frac{1}{2a}\int_{-a}^a |f(x)|dx = -\frac{1}{a}\int_0^a \ln x dx$$

$$= \frac{1}{a}\left[-x\ln x|_0^a + \int_0^a dx \right] = 1 - \ln a \to \infty, \text{ as } a \to 0^+.$$

Thus, $f \notin \mathrm{BMO}$, even though $|f(x)| = |\ln|x|| \in \mathrm{BMO}$ by Example 7.16 and Proposition 7.15 (vii). Thus, $|f| \in \mathrm{BMO}$ ***does NOT imply that*** $f \in \mathrm{BMO}$. Since $f = |f| \operatorname{sgn} f$ with $|f| \in \mathrm{BMO}$ and $\operatorname{sgn} f \in L^\infty(\mathbb{R}^n) \subset \mathrm{BMO}$, this also shows that ***the product of two functions in BMO does not necessarily belong to BMO***.

It is interesting to observe that an abrupt cut-off of a BMO function may not give a function in the same space.

Example 7.18. The function $h(x) = \chi_{\{x>0\}} \ln \frac{1}{x} = \frac{1}{2}(f(x) - \ln|x|) \notin \mathrm{BMO}$ in view of Examples 7.16 and 7.17.

A useful related fact is the following, which describes the behavior of BMO functions at infinity.

Theorem 7.19. *Let $f \in \mathrm{BMO}$; then, $f(x)(1+|x|^{n+1})^{-1}$ is integrable on $\mathbb{R}^n$, and we have*

$$I = \int_{\mathbb{R}^n} \frac{|f(x) - \operatorname{Avg}_{Q_0} f|}{1+|x|^{n+1}} dx \leqslant C\|f\|_{\mathrm{BMO}},$$

where C is independent of f, and $Q_0 = Q(0,1)$ is the cube centered at the origin with side length 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

$$I_k = \int_{S_k} \frac{|f(x) - \operatorname{Avg}_{Q_0} f|}{1+|x|^{n+1}} dx, \quad k \in \mathbb{N}_0.$$

Then, we have

$$I = I_0 + \sum_{k=1}^{\infty} I_k.$$

Since

$$I_0 = \int_{Q_0} \frac{|f(x) - \operatorname{Avg}_{Q_0} f|}{1+|x|^{n+1}} dx \leqslant \int_{Q_0} |f(x) - \operatorname{Avg}_{Q_0} f| dx \leqslant |Q_0| \, \|f\|_{\mathrm{BMO}},$$

it suffices to prove $I_k \leqslant C_k \|f\|_{\mathrm{BMO}}$ and $\sum_k C_k < \infty$. For $x \in S_k = Q(0, 2^k) \setminus Q(0, 2^{k-1})$, 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,

$$I_k \leqslant 4^{n+1} 2^{-k(n+1)} \int_{Q_k} |f(x) - \operatorname{Avg}_{Q_0} f| dx$$

$$
\begin{aligned}
&\leqslant 4^{n+1}2^{-k(n+1)} \int_{Q_k} [|f(x) - \operatorname*{Avg}_{Q_k} f| + |\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f|]dx \\
&\leqslant 4^{n+1}2^{-k(n+1)} |Q_k| \, (\|f\|_{\text{BMO}} + |\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f|) \\
&= 4^{n+1}2^{-k(n+1)}2^{kn}(\|f\|_{\text{BMO}} + |\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f|).
\end{aligned}
$$

The second term can be controlled as follows:

$$
\begin{aligned}
|\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f| &\leqslant \sum_{i=1}^{k} |\operatorname*{Avg}_{Q_i} f - \operatorname*{Avg}_{Q_{i-1}} f| \\
&\leqslant \sum_{i=1}^{k} \frac{1}{|Q_{i-1}|} \int_{Q_{i-1}} |f(x) - \operatorname*{Avg}_{Q_i} f| dx \\
&\leqslant \sum_{i=1}^{k} \frac{2^n}{|Q_i|} \int_{Q_i} |f(x) - \operatorname*{Avg}_{Q_i} f| dx \\
&\leqslant k \cdot 2^n \|f\|_{\text{BMO}}.
\end{aligned} \tag{7.4}
$$

Therefore,

$$
I_k \leqslant 4^{n+1}2^{-k}(1 + k2^n)\|f\|_{\text{BMO}},
$$

where $C_k = Ck2^{-k}$ and $\sum_{k=1}^{\infty} C_k = 2C < \infty$ due to

$$
\begin{aligned}
\sum_{k=1}^{\infty} k2^{-k} &= 2\sum_{k=1}^{\infty} k2^{-k} - \sum_{k=1}^{\infty} k2^{-k} = \sum_{k=0}^{\infty} (k+1)2^{-k} - \sum_{k=1}^{\infty} k2^{-k} \\
&= 1 + \sum_{k=1}^{\infty} 2^{-k} = 2.
\end{aligned}
$$

This completes the proof. ❑

Let us now look at more basic properties of BMO functions. As in (7.4), if a cube Q_1 is contained in a cube Q_2, then

$$
\begin{aligned}
|\operatorname*{Avg}_{Q_1} f - \operatorname*{Avg}_{Q_2} f| &= \left| \frac{1}{|Q_1|} \int_{Q_1} f dx - \operatorname*{Avg}_{Q_2} f \right| \leqslant \frac{1}{|Q_1|} \int_{Q_1} |f - \operatorname*{Avg}_{Q_2} f| dx \\
&\leqslant \frac{1}{|Q_1|} \int_{Q_2} |f - \operatorname*{Avg}_{Q_2} f| dx \\
&\leqslant \frac{|Q_2|}{|Q_1|} \|f\|_{\text{BMO}}.
\end{aligned} \tag{7.5}
$$

The same estimate holds if sets Q_1 and Q_2 are balls.

A version of this inequality is the first statement in the following proposition. For simplicity, we denote by $\|f\|_{\text{BMO}}$ the expression given by $\|f\|_{\text{BMO}_{\text{balls}}}$ in (7.3) since these quantities are comparable. For a ball B and $a > 0$, aB denotes the ball that is concentric with B and whose radius is a times the radius of B.

Proposition 7.20. (i) *Let $f \in \mathrm{BMO}(\mathbb{R}^n)$. Given a ball B and a positive integer m, we have*

$$|\operatorname*{Avg}_{B} f - \operatorname*{Avg}_{2^m B} f| \leqslant 2^n m \|f\|_{\mathrm{BMO}}. \tag{7.6}$$

(ii) *For any $\delta > 0$, there is a constant $C_{n,\delta}$ such that for any ball $B(x_0, R)$ we have*

$$R^\delta \int_{\mathbb{R}^n} \frac{|f(x) - \operatorname*{Avg}_{B(x_0,R)} f|}{(R + |x - x_0|)^{n+\delta}} dx \leqslant C_{n,\delta} \|f\|_{\mathrm{BMO}}. \tag{7.7}$$

An analogous estimate holds for cubes with center x_0 and side length R.

(iii) *There exists a constant C_n such that for all $f \in \mathrm{BMO}(\mathbb{R}^n)$, we have*

$$\sup_{y \in \mathbb{R}^n} \sup_{t>0} \int_{\mathbb{R}^n} |f(x) - (P_t * f)(y)| P_t(x - y) dx \leqslant C_n \|f\|_{\mathrm{BMO}}. \tag{7.8}$$

Here

$$P_t(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \frac{t}{(t^2 + |x|^2)^{\frac{n+1}{2}}}$$

denotes the Poisson kernel.

(iv) *Conversely, there is a constant C'_n such that for all $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ for which*

$$\int_{\mathbb{R}^n} \frac{|f(x)|}{(1 + |x|)^{n+1}} dx < \infty,$$

we have $f * P_t$ is well-defined and

$$C'_n \|f\|_{\mathrm{BMO}} \leqslant \sup_{y \in \mathbb{R}^n} \sup_{t>0} \int_{\mathbb{R}^n} |f(x) - (P_t * f)(y)| P_t(x - y) dx. \tag{7.9}$$

Proof. (i) We have the desired result as in (7.4).

(ii) In the proof below, we take $B(x_0, R)$ to be the ball $B = B(0, 1)$ with radius 1 centered at the origin. Once this case is known, given a ball $B(x_0, R)$, we replace the function f by the function $f(Rx + x_0)$. When $B = B(0, 1)$, we have by (i)

$$\begin{aligned}
&\int_{\mathbb{R}^n} \frac{|f(x) - \operatorname*{Avg}_{B} f|}{(1 + |x|)^{n+\delta}} dx \\
&\leqslant \int_{B} \frac{|f(x) - \operatorname*{Avg}_{B} f|}{(1 + |x|)^{n+\delta}} dx \\
&\quad + \sum_{k=0}^{\infty} \int_{2^{k+1}B \setminus 2^k B} \frac{|f(x) - \operatorname*{Avg}_{2^{k+1}B} f| + |\operatorname*{Avg}_{2^{k+1}B} f - \operatorname*{Avg}_{B} f|}{(1 + |x|)^{n+\delta}} dx
\end{aligned}$$

$$\begin{aligned}
&\leqslant \int_B |f(x) - \operatorname*{Avg}_B f| dx \\
&\quad + \sum_{k=0}^{\infty} 2^{-k(n+\delta)} \int_{2^{k+1}B} (|f(x) - \operatorname*{Avg}_{2^{k+1}B} f| + |\operatorname*{Avg}_{2^{k+1}B} f - \operatorname*{Avg}_B f|) dx \\
&\leqslant V_n \|f\|_{\mathrm{BMO}} + \sum_{k=0}^{\infty} 2^{-k(n+\delta)} (1 + 2^n(k+1))(2^{k+1})^n V_n \|f\|_{\mathrm{BMO}} \\
&= C'_{n,\delta} \|f\|_{\mathrm{BMO}}.
\end{aligned}$$

(iii) The proof of (7.8) is a reprise of the argument given in (ii). Set $B_t = B(y,t)$. We first prove a version of (7.8) in which the expression $(P_t * f)(y)$ is replaced by $\operatorname*{Avg}_{B_t} f$. For fixed y, t, we have by (ii)

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \int_{\mathbb{R}^n} \frac{t|f(x) - \operatorname*{Avg}_{B_t} f|}{(t^2 + |x-y|^2)^{\frac{n+1}{2}}} dx \leqslant C''_n \|f\|_{\mathrm{BMO}}. \tag{7.10}$$

Moving the absolute value outside, this inequality implies

$$\begin{aligned}
&\int_{\mathbb{R}^n} |(P_t * f)(y) - \operatorname*{Avg}_{B_t} f| P_t(x-y) dx \\
&= |(P_t * f)(y) - \operatorname*{Avg}_{B_t} f| \\
&\leqslant \int_{\mathbb{R}^n} P_t(x-y) |f(x) - \operatorname*{Avg}_{B_t} f| dx \\
&\leqslant C''_n \|f\|_{\mathrm{BMO}}.
\end{aligned}$$

Combining this last inequality with (7.10) yields (7.8) with constant $C_n = 2C''_n$.

(iv) Conversely, let A be the expression on the right side of (7.9). For $|x-y| \leqslant t$, we have $P_t(x-y) \geqslant c_n t(2t^2)^{-(n+1)/2} = c'_n t^{-n}$, which gives

$$A \geqslant \int_{\mathbb{R}^n} |f(x) - (P_t * f)(y)| P_t(x-y) dx \geqslant \frac{c'_n}{t^n} \int_{|x-y| \leqslant t} |f(x) - (P_t * f)(y)| dx.$$

Proposition 7.15 (iii) now implies that

$$\|f\|_{\mathrm{BMO}} \leqslant 2A/(V_n c'_n).$$

This concludes the proof of the proposition. ❑

§7.2.2 John-Nirenberg inequality

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 7.21. Let $f(x) = \ln|x|$, $I = (0,b)$, and

$$E_\alpha = \{x \in I : |\ln x - \operatorname*{Avg}_I f| > \alpha\},$$

then we have

$$\begin{aligned} E_\alpha =& \{x \in I : \ln x - \operatorname*{Avg}_I f > \alpha\} \cup \{x \in I : \ln x - \operatorname*{Avg}_I f < -\alpha\} \\ =& \{x \in I : x > e^{\alpha + \operatorname*{Avg}_I f}\} \cup \{x \in I : x < e^{-\alpha + \operatorname*{Avg}_I f}\}. \end{aligned}$$

When α is sufficiently large, the first set is an empty set, and the second set is $\left(0, e^{-\alpha+\operatorname*{Avg}_I f}\right)$. Thus,

$$|E_\alpha| = e^{-\alpha + \operatorname*{Avg}_I f}.$$

Since

$$\operatorname*{Avg}_I f = \frac{1}{b}\int_0^b \ln x dx = \frac{1}{b}\left(x \ln x|_0^b - \int_0^b dx\right) = \ln b - 1,$$

we have,

$$|E_\alpha| = |I| e^{-\alpha-1}.$$

That is, the distribution function decays exponentially.

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. The John-Nirenberg inequality gives a similar exponential estimate for the distribution function of oscillation of an arbitrary BMO function. The proof that we present here is based on a recursive use of the Calderón-Zygmund decomposition of cubes.

Theorem 7.22 (John-Nirenberg inequality)**.** *For all* $f \in \mathrm{BMO}(\mathbb{R}^n)$ *such that* $\|f\|_{\mathrm{BMO}} \neq 0$, *for all cubes* Q, *and all* $\alpha > 0$, *we have*

$$|\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > \alpha\}| \leqslant e|Q|e^{-A\alpha/\|f\|_{\mathrm{BMO}}} \tag{7.11}$$

with $A = (2^n e)^{-1}$.

Proof. Since inequality (7.11) 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 for the function $f - \operatorname*{Avg}_Q f$ at height b in the cube Q (similar to Theorem 2.13 by replacing $\mathbb{R}^n$ with Q).

Step 1. We introduce the following selection criterion for a cube R:

$$\frac{1}{|R|}\int_R |f(x) - \operatorname*{Avg}_Q f| dx > b. \tag{7.12}$$

Since

$$\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f| dx \leqslant \|f\|_{\mathrm{BMO}} = 1 < b,$$

the cube Q does not satisfy the selection criterion (7.12). 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 (7.12). 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 (7.12). Continue this process indefinitely. We obtain a countable collection of dyadic 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) $b < \left|Q_j^{(1)}\right|^{-1} \int_{Q_j^{(1)}} |f(x) - \operatorname*{Avg}_{Q^{(0)}} f|dx \leqslant 2^n b$.

(C-1) $\left| \operatorname*{Avg}_{Q_j^{(1)}} f - \operatorname*{Avg}_{Q^{(0)}} f \right| \leqslant 2^n b$.

(D-1) $\sum_j \left|Q_j^{(1)}\right| \leqslant \frac{1}{b} \sum_j \int_{Q_j^{(1)}} |f(x) - \operatorname*{Avg}_{Q^{(0)}} f|dx \leqslant \frac{1}{b} \left|Q^{(0)}\right|$.

(E-1) $|f - \operatorname*{Avg}_{Q^{(0)}} f| \leqslant b$ a.e. on the set $Q^{(0)} \setminus \cup_j Q_j^{(1)}$.

We call the cubes $Q_j^{(1)}$ of the first generation. Note that (C-1) is due to the upper inequality in (B-1), and (D-1) follows from the lower inequality in (B-1) and the fact

$$\frac{1}{|Q|} \int_Q |f(x) - \operatorname*{Avg}_Q f|dx \leqslant \|f\|_{\mathrm{BMO}} = 1.$$

Step 2. We now fix a selected first-generation cube $Q_j^{(1)}$, and we introduce the following selection criterion for a cube R:

$$\frac{1}{|R|} \int_R |f(x) - \operatorname*{Avg}_{Q_j^{(1)}} f|dx > b. \tag{7.13}$$

Observe that $Q_j^{(1)}$ does not satisfy the selection criterion (7.13). We apply a similar Calderón-Zygmund decomposition for the function $f - \operatorname*{Avg}_{Q_j^{(1)}} f$ at height b in every cube $Q_j^{(1)}$. Subdivide $Q_j^{(1)}$ into 2^n 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 (7.13). Continue this process indefinitely. This process is repeated for any other cube $Q_j^{(1)}$ of the first generation. We obtain a collection of dyadic cubes $\{Q_l^{(2)}\}_l$ of the 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 superscript (k) to denote the generation of the selected cubes.

Step 3. For a fixed selected cube $Q_l^{(2)}$ of the second generation, introduce the selection criterion

$$\frac{1}{|R|}\int_R |f(x) - \operatorname*{Avg}_{Q_l^{(2)}} f|dx > b. \tag{7.14}$$

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.

Step 4. We iterate this procedure indefinitely to obtain a doubly indexed family of dyadic 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) $b < \left|Q_j^{(k)}\right|^{-1} \int_{Q_j^{(k)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f|dx \leqslant 2^n b$.

(C-k) $\left|\operatorname*{Avg}_{Q_j^{(k)}} f - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f\right| \leqslant 2^n b$.

(D-k) $\sum_j \left|Q_j^{(k)}\right| \leqslant \frac{1}{b}\sum_{j'} \left|Q_{j'}^{(k-1)}\right|$.

(E-k) $|f - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f| \leqslant b$ 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 almost 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

$$|f - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f|$$

over all these cubes is at most b.

It remains to be proven (D-k). We have by (B-k)

$$\begin{aligned}
\sum_j \left|Q_j^{(k)}\right| &< \frac{1}{b}\sum_j \int_{Q_j^{(k)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f|dx \\
&= \frac{1}{b}\sum_{j'} \sum_{j \text{ corresp. to } j'} \int_{Q_j^{(k)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f|dx \\
&\leqslant \frac{1}{b}\sum_{j'} \int_{Q_{j'}^{(k-1)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f|dx \\
&\leqslant \frac{1}{b}\sum_{j'} \left|Q_{j'}^{(k-1)}\right| \|f\|_{\mathrm{BMO}} = \frac{1}{b}\sum_{j'} \left|Q_{j'}^{(k-1)}\right|.
\end{aligned}$$

Having established (A-k)-(E-k), we turn to some consequences. Applying (D-k) successively $k-1$ times, we obtain

$$\sum_j \left|Q_j^{(k)}\right| \leqslant b^{-k}\left|Q^{(0)}\right|. \tag{7.15}$$

For any fixed j, since (C-1) $|\operatorname{Avg}_{Q_j^{(1)}} f - \operatorname{Avg}_{Q^{(0)}} f| \leqslant 2^n b$ and (E-2) $|f - \operatorname{Avg}_{Q_j^{(1)}} f| \leqslant b$ a.e. on $Q_j^{(1)} \setminus \cup_l Q_l^{(2)}$, we have

$$|f - \operatorname{Avg}_{Q^{(0)}} f| \leqslant 2^n b + b \quad \text{a.e. on } Q_j^{(1)} \setminus \cup_l Q_l^{(2)},$$

which, combined with (E-1), yields

$$|f - \operatorname{Avg}_{Q^{(0)}} f| \leqslant 2^n 2b \quad \text{a.e. on } Q^{(0)} \setminus \cup_l Q_l^{(2)}. \tag{7.16}$$

For every fixed l, we also have (E-3) $|f - \operatorname{Avg}_{Q_l^{(2)}} f| \leqslant b$ a.e. on $Q_l^{(2)} \setminus \cup_s Q_s^{(3)}$, which combined with (C-2) $|\operatorname{Avg}_{Q_l^{(2)}} f - \operatorname{Avg}_{Q_{l'}^{(1)}} f| \leqslant 2^n b$ and (C-1) $|\operatorname{Avg}_{Q_{l'}^{(1)}} f - \operatorname{Avg}_{Q^{(0)}} f| \leqslant 2^n b$ yields

$$|f - \operatorname{Avg}_{Q^{(0)}} f| \leqslant 2^n 3b \quad \text{a.e. on } Q_l^{(2)} \setminus \cup_s Q_s^{(3)}.$$

In view of (7.16), 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

$$|f - \operatorname{Avg}_{Q^{(0)}} f| \leqslant 2^n kb \quad \text{a.e. on } Q^{(0)} \setminus \cup_s Q_s^{(k)}.$$

This proves the almost everywhere inclusion

$$\{x \in Q : |f(x) - \operatorname{Avg}_Q f| > 2^n kb\} \overset{\text{a.e.}}{\subset} \cup_j Q_j^{(k)} \tag{7.17}$$

for all $k = 1, 2, 3, \cdots$. (This also holds when $k = 0$.) We fix an $\alpha > 0$. If

$$2^n kb < \alpha \leqslant 2^n(k+1)b$$

for some $k \geqslant 0$, then from (7.17) and (7.15), we have

$$\left|\left\{x \in Q : |f(x) - \operatorname{Avg}_Q f| > \alpha\right\}\right| \leqslant \left|\left\{x \in Q : |f(x) - \operatorname{Avg}_Q f| > 2^n kb\right\}\right|$$
$$\leqslant \sum_j \left|Q_j^{(k)}\right| \leqslant \frac{1}{b^k}\left|Q^{(0)}\right| = |Q| e^{-k\ln b} \leqslant |Q| b e^{-\alpha \ln b/(2^n b)},$$

since $-k \leqslant 1 - \frac{\alpha}{2^n b}$. Choosing $b = e > 1$ yields (7.11). ❑

The John-Nirenberg inequality tells us that logarithmic blowup, as for $f(x) = \ln|x|$, is the worst possible behavior for a general BMO function. In this sense the John-Nirenberg inequality is the best possible result we can hope for.

Having proven the important John-Nirenberg inequality (7.11), we are now able to deduce from it a few corollaries.

Corollary 7.23. *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*

$$\frac{1}{|Q|}\int_Q e^{\gamma\left|f(x)-\operatorname*{Avg}_Q f\right|/\|f\|_{\mathrm{BMO}}}dx \leqslant 1+\frac{2^n e^2\gamma}{1-2^n e\gamma}.$$

Proof. Let $h = \gamma|f(x) - \operatorname*{Avg}_Q f|/\|f\|_{\mathrm{BMO}}$ and $\gamma < A = (2^n e)^{-1}$. By (7.11), we have the distribution function

$$h_*(\alpha) = |\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > \alpha\|f\|_{\mathrm{BMO}}/\gamma\}| \leqslant e|Q|e^{-A\alpha/\gamma},$$

which yields, by Theorem 1.16 with $\varphi(t) = e^t - 1$ (so $\varphi(0) = 0$) and integration by parts, that

$$\begin{aligned}\frac{1}{|Q|}\int_Q e^h dx =&1+\frac{1}{|Q|}\int_Q (e^h-1)dx = 1-\frac{1}{|Q|}\int_0^\infty (e^\alpha-1)dh_*(\alpha)\\ =&1+\frac{1}{|Q|}\left[-(e^\alpha-1)h_*(\alpha)|_0^\infty+\int_0^\infty e^\alpha h_*(\alpha)d\alpha\right]\\ =&1+\frac{1}{|Q|}\int_0^\infty e^\alpha h_*(\alpha)d\alpha.\end{aligned}$$

Then, we obtain

$$\begin{aligned}&\frac{1}{|Q|}\int_Q e^{\gamma|f(x)-\operatorname*{Avg}_Q f|/\|f\|_{\mathrm{BMO}}}dx\\ \leqslant&1+\int_0^\infty e^\alpha e e^{-A\frac{\alpha}{\gamma}}d\alpha\\ =&1+e\int_0^\infty e^{\alpha(1-1/(2^n e\gamma))}d\alpha = 1+\frac{2^n e^2\gamma}{1-2^n e\gamma},\end{aligned}$$

thus, we complete the proof. ❑

For another important corollary, we define the following.

Definition 7.24 (BMO_p)**.** Let $1 < p < \infty$; for $f \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$, we define

$$\|f\|_{\mathrm{BMO}_p} = \sup_Q\left(\operatorname*{Avg}_Q|f-\operatorname*{Avg}_Q f|^p\right)^{1/p},$$

and

$$\mathrm{BMO}_p = \left\{f \in L^p_{\mathrm{loc}}(\mathbb{R}^n) : \|f\|_{\mathrm{BMO}_p} < \infty\right\}.$$

Corollary 7.25. *For all $1 < p < \infty$, there exists a finite constant $B_{p,n}$ such that*

$$\|f\|_{\mathrm{BMO}_p} \leqslant B_{p,n}\|f\|_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{7.18}$$

Proof. We have from Theorem 1.17 and the John-Nirenberg inequality

$$\begin{aligned}\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f|^p dx =& \frac{p}{|Q|}\int_0^\infty \alpha^{p-1}\left|\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > \alpha\}\right| d\alpha \\ \leqslant& \frac{p}{|Q|} e|Q| \int_0^\infty \alpha^{p-1} e^{-A\alpha/\|f\|_{\mathrm{BMO}}} d\alpha \\ =& p\Gamma(p)\frac{e}{A^p}\|f\|^p_{\mathrm{BMO}},\end{aligned}$$

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 7.25 can be reversed via Hölder's inequality, we obtain the following important L^p characterization of BMO norms.

Corollary 7.26. *For all $1 < p < \infty$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, we have*

$$\|f\|_{\mathrm{BMO}_p} \sim \|f\|_{\mathrm{BMO}}. \tag{7.19}$$

Proof. One direction follows from Corollary 7.25, and the other follows from the Hölder inequality. ❑

§7.3 Duality between $\mathcal{H}^1$ and BMO

The next result we give is a remarkable duality relationship between the Hardy space $\mathcal{H}^1$ and BMO. Specifically, we have that BMO is isomorphic to the dual space of $\mathcal{H}^1$ with equivalent norms. This means that every continuous linear functional on the Hardy space $\mathcal{H}^1$ can be realized as integration against a fixed BMO function, where integration in this context is an abstract operation, not necessarily given by an absolutely convergent integral. This relationship was first established by Fefferman and Stein in [FS72] but using a different characterization of $\mathcal{H}^1$.

Theorem 7.27 ($\mathcal{H}^1$-BMO duality)**.** *The dual of $\mathcal{H}^1$ is isomorphic to* BMO *with equivalent norms.*

Proof. We work with $\mathcal{H}^1 = \mathcal{H}^{1,2}$ and $\mathrm{BMO} = \mathrm{BMO}_2$ with corresponding norms $\|\cdot\|_{\mathcal{H}^{1,2}}$ and $\|\cdot\|_{\mathrm{BMO}_2}$, in view of Definition 7.9 and Corollary 7.26.

(i) Take $b \in L^\infty(\mathbb{R}^n)$ and $f \in \mathcal{H}^{1,2}$. Define

$$L_b(f) = \int_{\mathbb{R}^n} b(x)f(x)dx,$$

which is well-defined since $\mathcal{H}^{1,2} \subset L^1$. If $f = \sum_{i=1}^{\infty} \lambda_i a_i$, we can apply the dominated convergence theorem due to

$$\int_{\mathbb{R}^n} \sum_{i=1}^{\infty} |b(x)\lambda_i a_i(x)|dx \leqslant \|b\|_\infty \sum_{i=1}^{\infty} |\lambda_i|.$$

If $\operatorname{supp} a_i \subset Q_i$, then $\int_{Q_i} a_i dx = 0$ by definition. Thus, $\|a_i\|_1 \leqslant 1$ yields

$$|L_b(a_i)| = \left| \int_{\mathbb{R}^n} (b(x) - \operatorname*{Avg}_{Q_i} b) a_i(x)dx \right| \leqslant 2\|b\|_\infty,$$

which implies, by Hölder's inequality, that

$$\begin{aligned}
|L_b(f)| =& \left| \sum_{i=1}^{\infty} \int_{Q_i} (b(x) - \operatorname*{Avg}_{Q_i} b)\lambda_i a_i(x)dx \right| \\
\leqslant& \sum_{i=1}^{\infty} \left(\frac{1}{|Q_i|} \int_{Q_i} |b(x) - \operatorname*{Avg}_{Q_i} b|^2 dx \right)^{1/2} |\lambda_i| \left(|Q_i| \int_{Q_i} |a_i(x)|^2 dx \right)^{1/2} \\
\leqslant& \|b\|_{\mathrm{BMO}_2} \sum_{i=1}^{\infty} |\lambda_i|.
\end{aligned}$$

Taking an infimum over all possible λ_i, we have

$$|L_b(f)| \leqslant \|b\|_{\mathrm{BMO}_2} \|f\|_{\mathcal{H}^{1,2}}.$$

(ii) Now take $b \in \mathrm{BMO}_2$ to be real-valued without loss of generality (for complex-valued b, we may separate real and imaginary parts), and $f \in \operatorname{span} \mathcal{A}^2$. Let $(b_k)_{k=1}^{\infty}$ be the truncation of b given by (ix) in Proposition 7.15. Thus, $|b_k| \leqslant |b|$ a.e., $b_k \nearrow b$ a.e., and $\|b_k\|_{\mathrm{BMO}_2} \leqslant 2\|b\|_{\mathrm{BMO}_2}$. Suppose that $f = \sum_{i=1}^{m} \lambda_i a_i$ and $L_{b_k}(f) = \sum_{i=1}^{m} \lambda_i L_{b_k}(a_i)$. Since $b \in \mathrm{BMO}_2$, we have $b \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, which implies $b \in L^2(\operatorname{supp} a_i)$. Thus, $|b_k a_i| \leqslant |b a_i| \in L^1(\mathbb{R}^n)$ a.e. and by the dominated convergence theorem

$$\int_{\mathbb{R}^n} b_k a_i dx \to \int_{\mathbb{R}^n} b a_i dx,$$

i.e., $L_{b_k}(a_i) \to L_b(a_i)$ as $k \to \infty$. It follows from $|L_{b_k}(f)| \leqslant \|b_k\|_{\mathrm{BMO}_2} \|f\|_{\mathcal{H}^{1,2}} \leqslant 2\|b\|_{\mathrm{BMO}_2} \|f\|_{\mathcal{H}^{1,2}}$ that

$$|L_b(f)| \leqslant 2\|b\|_{\mathrm{BMO}_2} \|f\|_{\mathcal{H}^{1,2}}.$$

(iii) By the density of $\operatorname{span} \mathcal{A}^2$ in $\mathcal{H}^{1,2}$, we can extend L_b to the whole of $\mathcal{H}^{1,2}$. Let $\tilde{L}_b$ denote this extension. Thus, we have shown that whenever $b \in \mathrm{BMO}_2$ we have $\tilde{L}_b \in (\mathcal{H}^{1,2})'$. Let $T : \mathrm{BMO}_2 \to (\mathcal{H}^{1,2})'$ denote the map $b \mapsto \tilde{L}_b$ which is linear.

(iv) We show that T is injective. Let $b \in \mathrm{BMO}_2$ be such that

$\tilde{L}_b = 0$ and show that b is constant. Fix a cube Q and let $L^2_0(Q) = \left\{f \in L^2(Q) : \int_Q f dx = 0\right\}$. Note that $L^2_0(Q) \subset \mathcal{H}^{1,2}$. In fact, we take $f \in L^2_0(Q)$ and $\lambda = \|f\|_2 |Q|^{1/2}$, then

$$\left\|\frac{f}{\lambda}\right\|_2 \leqslant \frac{1}{|Q|^{1/2}},$$

so f/λ is a 2-atom. Thus, we get an expression $f = \lambda f/\lambda$, then by definition

$$\|f\|_{\mathcal{H}^{1,2}} \leqslant \lambda = \|f\|_2 |Q|^{1/2}.$$

Thus, from

$$0 = \tilde{L}_b(f) = L_b(f) = \int_Q bf dx, \quad \forall f \in L^2_0(Q),$$

it follows that $b|_Q$ is constant a.e., since we can take a special $f = b - \operatorname{Avg}_Q b \in L^2_0(Q)$ due to $b \in \mathrm{BMO}_2$ and then $0 = \int_Q |b - \operatorname{Avg}_Q b|^2 dx$. By exhaustion of $\mathbb{R}^n$ by increasing Q, we deduce that b is constant.

(v) Finally, we show that T is surjective. Let $L \in (\mathcal{H}^{1,2})'$ and fix a cube Q. Since $L^2_0(Q) \subset \mathcal{H}^{1,2}$, we have $L|_{L^2_0(Q)} \in (L^2_0(Q))' = L^2(Q)/\{\text{constants}\}$, and thus, by the Riesz representation theorem, there exists a $b_Q \in (L^2_0(Q))'$ such that for all $f \in L^2_0(Q)$,

$$L(f) = \int_Q b_Q f dx$$

and $\|b_Q\|_{L^2(Q)} = \sup_{\|f\|_{L^2_0(Q)} \leqslant 1} |L(f)| \leqslant \sup_{\|f\|_{L^2_0(Q)} \leqslant 1} \|L\|_{(\mathcal{H}^{1,2})'} \|f\|_{\mathcal{H}^{1,2}} \leqslant \|L\|_{(\mathcal{H}^{1,2})'} |Q|^{1/2}$.

Let Q and Q' denote two cubes with $Q \subset Q'$. Then, whenever $f \in L^2_0(Q)$, we have $f \in L^2_0(Q')$ and

$$L(f) = \int_Q b_Q f dx = \int_{Q'} b_{Q'} f dx,$$

so $b_Q - b_{Q'}$ is constant a.e. in Q as before. Define b as follows:

$$b(x) = \begin{cases} b_{[-1,1]^n}(x), & x \in [-1,1]^n, \\ b_{[-2^j,2^j]^n}(x) + c_j, & x \in [-2^j,2^j]^n \setminus [-2^{j-1},2^{j-1}]^n, \ j \geqslant 1, \end{cases}$$

where c_j is the constant such that $b_{[-2^j,2^j]^n} - b_{[-1,1]^n} = -c_j$ on $[-1,1]^n$.

We show that $b \in \mathrm{BMO}_2$, $\|b\|_{\mathrm{BMO}_2} \leqslant \|L\|_{(\mathcal{H}^{1,2})'}$ and $L = \tilde{L}_b$. Fix Q, and let $j \in \mathbb{N}$ such that $Q \subset [-2^j,2^j]^n$. Let k be such that $2 \leqslant k \leqslant j$. Then, $c_k - c_{k-1} = b_{[-2^k,2^k]^n} - b_{[-2^{k-1},2^{k-1}]^n}$ which is constant on $[-2^{k-1},2^{k-1}]^n$ and in particular on $[-2^{k-1},2^{k-1}]^n \setminus [-2^{k-2},2^{k-2}]^n$. Therefore, $b(x) = b_{[-2^j,2^j]^n}(x) + c_j$ on all of $[-2^j,2^j]^n$ and in particular on Q. Additionally, there exists a constant c such that $b_{[-2^j,2^j]^n} - b_Q = c$ on the cube Q and so $b = b_Q + c + c_j$ on Q. Then,

$$b - \operatorname*{Avg}_Q b = b_Q + c + c_j - \operatorname*{Avg}_Q b_Q - c - c_j = b_Q$$

since $\operatorname*{Avg}_Q b_Q = 0$. Therefore,

$$\int_Q |b - \operatorname*{Avg}_Q b|^2 dx = \int_Q |b_Q|^2 dx \leqslant \|L\|^2_{(\mathcal{H}^{1,2})'} |Q|.$$

The fact that $L = \tilde{L}_b$ follows from the fact that $L(a) = \tilde{L}_b(a)$ for all $a \in \mathcal{A}^2$. ❑

§ 7.4 Carleson measures

§ 7.4.1 Nontangential maximal functions and Carleson measures

Definition 7.28 (Cone)**.** Let $x \in \mathbb{R}^n$. We define the cone over x as follows:

$$\Gamma(x) = \{(y,t) \in \mathbb{R}^{n+1}_+ : |x - y| < t\}.$$

Definition 7.29 (Nontangential maximal function)**.** Let $F : \mathbb{R}^{n+1}_+ \to \mathbb{C}$ and define the nontangential maximal function of F:

$$M^* F(x) = \sup_{(y,t) \in \Gamma(x)} |F(y,t)| \in [0, \infty].$$

Remark 7.30. (i) We observe that if $M^* F(x) = 0$ for almost all $x \in \mathbb{R}^n$, then F is identically equal to zero on $\mathbb{R}^{n+1}_+$. To establish this claim, suppose that $|F(x_0, t_0)| > 0$ for some point $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}^+$. Then, for all z with $|z - x_0| < t_0$, we have $(x_0, t_0) \in \Gamma(z)$; hence, $M^* F(z) \geqslant |F(x_0, t_0)| > 0$. Thus, $M^* F > 0$ on the ball $B(x_0, t_0)$, which is a set of positive measures, a contradiction.

(ii) Given a Borel measure μ on $\mathbb{R}^{n+1}_+$, we can define the nontangential maximal function M^*_μ w.r.t. μ by replacing sup with ess sup. Note then that M^*_μ is defined μ-a.e.

Definition 7.31 (Tent)**.** Let $B = B(x_0, r) \subset \mathbb{R}^n$ be an open ball. We define the ***cylindrical tent*** over B to be the "cylindrical set"

$$T(B) = \{(x,t) \in \mathbb{R}^{n+1}_+ : x \in B,\ 0 < t \leqslant r\} = B \times (0, r].$$

Similarly, for a cube Q in $\mathbb{R}^n$, we define the tent over Q to be the cube

$$T(Q) = Q \times (0, \ell(Q)].$$

Definition 7.32 (Carleson measure)**.** A ***Carleson measure*** is a positive measure μ on $\mathbb{R}^{n+1}_{+}$ such that there exists a constant $C < \infty$ for which

$$\mu(T(B)) \leqslant C|B|$$

for all $B = B(x, r)$. We define the Carleson norm as

$$\|\mu\|_{\mathscr{C}} = \sup_{B} \frac{\mu(T(B))}{|B|}.$$

Remark 7.33. In the definition of the Carleson norm, B and $T(B)$ can be replaced by the cubes Q and $T(Q)$, respectively. One can easily verify that they are equivalent.

Example 7.34. The following measures are not Carleson measures.

(i) The Lebesgue measure $d\mu(x,t) = dxdt$ since no such constant C is possible for large balls.

(ii) $d\mu(x,t) = dx\frac{dt}{t}$ since $\mu(B \times (0,r]) = |B| \int_0^r \frac{dt}{t} = \infty$.

(iii) $d\mu(x,t) = \frac{dxdt}{t^\alpha}$ for $\alpha \in \mathbb{R}$. Note that

$$\mu(B \times (0,r]) = |B| \int_0^r \frac{dt}{t^\alpha} = \begin{cases} |B|\dfrac{r^{1-\alpha}}{1-\alpha}, & 1-\alpha > 0, \\ \infty, & \text{otherwise.} \end{cases}$$

Therefore, we only need to consider the case $\alpha < 1$, but in this case, we cannot obtain uniform control via a constant C.

Example 7.35. The following are examples of Carleson measures.

(i) $d\mu(x,t) = \chi_{[a,b]}(t)dx\frac{dt}{t}$ where $0 < a < b < \infty$. Then, the constant $C = \ln \frac{b}{a}$.

(ii) $d\mu(y,t) = \chi_{\Gamma(x)}(y)dy\frac{dt}{t}$. Then,

$$\mu(B \times (0,r]) \leqslant \int_0^r |B(x,t)|\frac{dt}{t} = \int_0^r t^n|B(0,1)|\frac{dt}{t} = \frac{r^n|B(0,1)|}{n} = \frac{|B|}{n}.$$

(iii) Let L be a line in $\mathbb{R}^2$. For measurable subsets $A \subset \mathbb{R}^2_+$, define $\mu(A)$ to be the linear Lebesgue measure of the set $L \cap A$. Then μ is a Carleson measure on $\mathbb{R}^2_+$. Indeed, the linear measure of the part of a line inside the box $T(B) = [x_0 - r, x_0 + r] \times (0, r]$ is at most equal to the diagonal of the box, i.e., $\sqrt{5}r$ where $B = [x_0 - r, x_0 + r]$, thus $\mu(T(B)) \leqslant \sqrt{5}r = \frac{\sqrt{5}}{2}|B|$.

Definition 7.36 (Carleson function)**.** The ***Carleson function*** of the measure μ is defined as

$$\mathscr{C}(\mu)(x) = \sup_{B \ni x} \frac{\mu(T(B))}{|B|} \in [0, \infty].$$

Observe that $\|\mathscr{C}(\mu)\|_\infty = \|\mu\|_{\mathscr{C}}$ for Carleson measure μ.

Theorem 7.37 (Carleson Lemma)**.** *There exists a dimensional constant C_n such that for all $\alpha > 0$, all measure μ on $\mathbb{R}^{n+1}_+$, and all μ-measurable functions $F : \mathbb{R}^{n+1}_+ \to \mathbb{C}$, the set $\Omega_\alpha = \{x \in \mathbb{R}^n : M^*F(x) > \alpha\}$ is open (thus M^*F is Lebesgue measurable) and we have*

$$\mu(\{(x,t) \in \mathbb{R}^{n+1}_+ : |F(x,t)| > \alpha\}) \leqslant C_n \int_{\{M^*F>\alpha\}} \mathscr{C}(\mu)(x)dx. \tag{7.20}$$

In particular, if μ is a Carleson measure, then

$$\mu(\{|F| > \alpha\}) \leqslant C_n \|\mu\|_{\mathscr{C}} |\{M^*F > \alpha\}|. \tag{7.21}$$

Proof. We first prove that for any μ-measurable function F, the set Ω_α is open, and consequently, M^*F is Lebesgue measurable. Indeed, if $x_0 \in \Omega_\alpha$, then for any $\varepsilon \in (0, M^*F(x_0) - \alpha)$ there exists a $(y_0, t_0) \in \Gamma(x_0) = \{(y,t) \in \mathbb{R}^n \times \mathbb{R}^+ : |y - x_0| < t\}$ such that $|F(y_0, t_0)| > M^*F(x_0) - \varepsilon > \alpha$. If d_0 is the distance from (y_0, t_0) to the sphere formed by the intersection of the hyperplane $t_0 + \mathbb{R}^n$ (i.e., $\mathbb{R}^n \times \{t_0\}$) with the boundary of the cone $\Gamma(x_0)$, then $|x_0 - y_0| = t_0 - d_0$. It follows the open ball $B(x_0, d_0) \subset \Omega_\alpha$ since for $z \in B(x_0, d_0)$ we have $|z - y_0| \leqslant |z - x_0| + |x_0 - y_0| < d_0 + t_0 - d_0 = t_0$, i.e., $(y_0, t_0) \in \Gamma(z)$; hence, $M^*F(z) \geqslant |F(y_0, t_0)| > \alpha$.

Let $\{Q_k\}$ be the Whitney decomposition (i.e., Lemma 2.15) of the set Ω_α. For each $x \in \Omega_\alpha$, set $\delta_\alpha(x) = \operatorname{dist}(x, \Omega_\alpha^c)$. Then, for $z \in Q_k$, we have

$$\begin{aligned}\delta_\alpha(z) \leqslant& \sqrt{n}\ell(Q_k) + \operatorname{dist}(Q_k, \Omega_\alpha^c)\\ \leqslant& \sqrt{n}\ell(Q_k) + 4\operatorname{diam}(Q_k) = 5\sqrt{n}\ell(Q_k)\end{aligned} \tag{7.22}$$

in view of Lemma 2.15 (iii). For each Q_k (centered at z_0), let B_k be the smallest ball that contains Q_k. Then B_k is of radius $\sqrt{n}\ell(Q_k)/2$ and centered at z_0. Combine this observation with (7.22) to obtain that for any $z \in Q_k$ and $y \in B(z, \delta_\alpha(z))$

$$\begin{aligned}|y - z_0| \leqslant& |y - z| + |z - z_0| \leqslant \delta_\alpha(z) + \sqrt{n}\ell(Q_k)/2\\ \leqslant& \frac{11}{2}\sqrt{n}\ell(Q_k) = 11\mathrm{rad}(B_k),\end{aligned}$$

namely,

$$z \in Q_k \Longrightarrow B(z, \delta_\alpha(z)) \subset 11B_k.$$

This implies that

$$\bigcup_{z\in\Omega_\alpha} T(B(z, \delta_\alpha(z))) \subset \bigcup_k T(11B_k). \tag{7.23}$$

Next, we claim that

$$\{|F| > \alpha\} \subset \bigcup_{z\in\Omega_\alpha} T(B(z, \delta_\alpha(z))). \tag{7.24}$$

Indeed, let $(x,t) \in \mathbb{R}^{n+1}_+$ be such that $|F(x,t)| > \alpha$. Then by the definition of M^*F, we have that $M^*F(y) > \alpha$ for all $y \in \mathbb{R}^n$ satisfying

$|x-y| < t$. Thus, $B(x,t) \subset \Omega_\alpha$ and so $\delta_\alpha(x) \geqslant t$. This gives that $(x,t) \in T(B(x,\delta_\alpha(x)))$, which proves (7.24).

Combining (7.23) and (7.24), we obtain

$$\{|F| > \alpha\} \subset \bigcup_k T(11B_k).$$

Applying the measure μ and using the definition of the Carleson function, we obtain

$$\begin{aligned}
\mu(\{|F| > \alpha\}) &\leqslant \sum_k \mu(T(11B_k)) \\
&\leqslant \sum_k |11B_k| \inf_{x\in 11B_k} \mathscr{C}(\mu)(x) \\
&\leqslant \sum_k |11B_k| \inf_{x\in Q_k} \mathscr{C}(\mu)(x) \quad (\because Q_k \subset 11B_k) \\
&\leqslant 11^n \sum_k \frac{|B_k|}{|Q_k|} \int_{Q_k} \mathscr{C}(\mu)(x)dx \\
&\leqslant (11\sqrt{n}/2)^n V_n \int_{\Omega_\alpha} \mathscr{C}(\mu)(x)dx.
\end{aligned}$$

This proves (7.20). It follows (7.21) in view of $\|\mathscr{C}(\mu)\|_\infty = \|\mu\|_{\mathscr{C}}$. ❑

Corollary 7.38. *For any Carleson measure μ and every μ-measurable function F on $\mathbb{R}^{n+1}_+$, we have*

$$\int_{\mathbb{R}^{n+1}_+} |F(x,t)|^p d\mu(x,t) \leqslant C_n \|\mu\|_{\mathscr{C}} \int_{\mathbb{R}^n} (M^*F(x))^p dx \tag{7.25}$$

for all $p \in [1,\infty)$.

Proof. From (7.21), applying Theorem 1.17 twice, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^{n+1}_+} |F(x,t)|^p d\mu(x,t) =& p\int_0^\infty \alpha^{p-1}\mu(\{|F| > \alpha\})d\alpha \\
\leqslant& C_n \|\mu\|_{\mathscr{C}} p \int_0^\infty \alpha^{p-1} |\{M^*F > \alpha\}| d\alpha \\
=& C_n \|\mu\|_{\mathscr{C}} \int_{\mathbb{R}^n} (M^*F(x))^p dx.
\end{aligned}$$

❑

A particular example of this situation arises when $F(x,t) = f * \Phi_t(x)$ for some nice integrable function Φ. Here and in the sequel, $\Phi_t(x) = t^{-n}\Phi(t^{-1}x)$. For instance, one may take Φ_t to be the Poisson kernel P_t.

Theorem 7.39. *Let Φ be a function on $\mathbb{R}^n$ that satisfies for some $0 < C, \delta < \infty$,*

$$|\Phi(x)| \leqslant \frac{C}{(1+|x|)^{n+\delta}}. \tag{7.26}$$

Let μ be a Carleson measure on $\mathbb{R}^{n+1}_+$. Then for every $1 < p < \infty$, there is a *constant* $C_{p,n}(\mu)$ *such that for all* $f \in L^p(\mathbb{R}^n)$ *we have*

$$\int_{\mathbb{R}^{n+1}_+} |(\Phi_t * f)(x)|^p d\mu(x,t) \leqslant C_{p,n}(\mu) \int_{\mathbb{R}^n} |f(x)|^p dx, \tag{7.27}$$

where $C_{p,n}(\mu) \leqslant C(p,n)\|\mu\|_{\mathscr{C}}$.

Conversely, suppose that Φ *is a nonnegative function that satisfies* (7.26) *and* $\int_{|x|\leqslant 1} \Phi(x)dx > 0$. *If* μ *is a measure on* $\mathbb{R}^{n+1}_+$ *such that for some* $1 < p < \infty$ *there is a constant* $C_{p,n}(\mu)$ *such that* (7.27) *holds for all* $f \in L^p(\mathbb{R}^n)$, *then* μ *is a Carleson measure with norm at most a multiple of* $C_{p,n}(\mu)$.

Proof. If μ is a Carleson measure, we may obtain (7.27) as a sequence of Corollary 7.38. Indeed, for $F(x,t) = (\Phi_t * f)(x)$, we have

$$\begin{aligned}
M^*F(x) =& \sup_{t>0} \sup_{\substack{y\in\mathbb{R}^n\\|y-x|<t}} |(\Phi_t * f)(y)| \\
\leqslant& \sup_{t>0} \sup_{\substack{y\in\mathbb{R}^n\\|y-x|<t}} \int_{\mathbb{R}^n} |\Phi_t(y-z)||f(z)|dz \\
=& \sup_{t>0} \sup_{\substack{y\in\mathbb{R}^n\\|y-x|<t}} \int_{\mathbb{R}^n} |\Phi_t(y-x+x-z)||f(z)|dz \\
\leqslant& \sup_{t>0} \Big(\sup_{\substack{y\in\mathbb{R}^n\\|y-x|<t}} |\Phi_t(y-x+\cdot)| * |f| \Big)(x) \\
=& \sup_{t>0} (\Psi_t * |f|)(x),
\end{aligned}$$

where

$$\Psi(x) := \sup_{|u|\leqslant 1} |\Phi(x-u)| \leqslant \begin{cases} C, & |x| \leqslant 1, \\ \dfrac{C}{|x|^{n+\delta}}, & |x| > 1, \end{cases}$$

by condition (7.26). Thus, it is clear that $\|\Psi\|_{L^1(\mathbb{R}^n)} \leqslant C(V_n + \omega_{n-1}/\delta)$. It follows from Theorem 2.10 that

$$M^*F(x) \leqslant C(n,\delta)M(|f|)(x).$$

Then, by Theorem 2.6, we obtain

$$\int_{\mathbb{R}^n} (M^*F(x))^p dx \leqslant C(n,\delta) \int_{\mathbb{R}^n} (M(|f|)(x))^p dx \leqslant C(n,\delta,p) \int_{\mathbb{R}^n} |f(x)|^p dx.$$

Therefore, from Corollary 7.38, (7.27) follows.

Conversely, if (7.27) holds, then we fix a ball $B = B(x_0, r)$ in $\mathbb{R}^n$ with center x_0 and radius $r > 0$. Then for $(x,t) \in T(B)$, we have

$$\begin{aligned}
(\Phi_t * \chi_{2B})(x) =& \int_{2B} \Phi_t(x-y)dy = \int_{x-2B} \Phi_t(y)dy \\
\geqslant& \int_{B(0,t)} \Phi_t(y)dy = \int_{B(0,1)} \Phi(y)dy = c_n > 0,
\end{aligned}$$

since $B(0,t) \subset x - 2B(x_0, r)$ whenever $t \leqslant r$. Therefore, we have from (7.27)

$$\begin{aligned}\mu(T(B)) = \int_{T(B)} d\mu(x,t) &\leqslant \frac{1}{c_n^p} \int_{T(B)} c_n^p d\mu(x,t) \\ &\leqslant \frac{1}{c_n^p} \int_{T(B)} |(\Phi_t * \chi_{2B})(x)|^p d\mu(x,t) \\ &\leqslant \frac{1}{c_n^p} \int_{\mathbb{R}_+^{n+1}} |(\Phi_t * \chi_{2B})(x)|^p d\mu(x,t) \\ &\leqslant \frac{C_{p,n}(\mu)}{c_n^p} \int_{\mathbb{R}^n} |\chi_{2B}(x)|^p dx \\ &= \frac{2^n C_{p,n}(\mu)}{c_n^p} |B|.\end{aligned}$$

This proves that μ is a Carleson measure with $\|\mu\|_{\mathscr{C}} \leqslant 2^n c_n^{-p} C_{n,p}(\mu)$. ❑

§7.4.2 BMO functions and Carleson measures

We now turn to an interesting connection between BMO functions and Carleson measures as follows.

Theorem 7.40. *Let $b \in \mathrm{BMO}(\mathbb{R}^n)$ and $\Psi \in L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \Psi(x)dx = 0$ satisfying*

$$|\Psi(x)| \leqslant A(1 + |x|)^{-n-\delta} \tag{7.28}$$

for some $0 < A, \delta < \infty$. Consider the dilation $\Psi_t = t^{-n}\Psi(t^{-1}x)$.

(i) *Suppose that*

$$\sup_{\xi \in \mathbb{R}^n} \sum_{j \in \mathbb{Z}} |\widehat{\Psi}(2^{-j}\xi)|^2 \leqslant B^2 < \infty \tag{7.29}$$

and let $\delta_{2^{-j}}(t)$ be the Dirac mass at the point $t = 2^{-j}$. Then there is a constant $C_{n,\delta}$ such that

$$d\mu(x,t) = \sum_{j \in \mathbb{Z}} |(\Psi_{2^{-j}} * b)(x)|^2 dx \delta_{2^{-j}}(t) dt$$

is a Carleson measure on $\mathbb{R}_+^{n+1}$ with norm at most $C_{n,\delta}(A + B)^2 \|b\|_{\mathrm{BMO}}^2$.

(ii) *Suppose that*

$$\sup_{\xi \in \mathbb{R}^n} \int_0^\infty |\widehat{\Psi}(t\xi)|^2 \frac{dt}{t} \leqslant B^2 < \infty. \tag{7.30}$$

Then the continuous version $d\nu(x,t)$ of $d\mu(x,t)$ defined by

$$d\nu(x,t) = |(\Psi_t * b)(x)|^2 dx \frac{dt}{t}$$

is a Carleson measure on $\mathbb{R}_+^{n+1}$ with norm at most $C_{n,\delta}(A +$

$B)^2\|b\|^2_{\mathrm{BMO}}$ for some constant $C_{n,\delta}$.

(iii) *Let $\delta, A > 0$. Suppose that $\{K_t\}_{t>0}$ are functions on $\mathbb{R}^n \times \mathbb{R}^n$ that satisfy*

$$|K_t(x,y)| \leqslant \frac{At^\delta}{(t+|x-y|)^{n+\delta}} \tag{7.31}$$

for all $t > 0$ and all $x, y \in \mathbb{R}^n$. Let R_t be the linear operator

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

which is well-defined for all $f \in \bigcup_{p\in[1,\infty]} L^p(\mathbb{R}^n)$. Suppose that $R_t(1) = 0$ for all $t > 0$ and that there is a constant $B > 0$ such that

$$\int_0^\infty \int_{\mathbb{R}^n} |R_t(f)(x)|^2 \frac{dxdt}{t} \leqslant B^2 \|f\|^2_{L^2(\mathbb{R}^n)} \tag{7.32}$$

for all $f \in L^2(\mathbb{R}^n)$. Then for all $b \in \mathrm{BMO}$, the measure

$$|R_t(b)(x)|^2 \frac{dxdt}{t}$$

is Carleson with norm at most a constant multiple of $(A+B)^2\|b\|^2_{\mathrm{BMO}}$.

Proof. (i) The measure μ is defined so that for every μ-integrable function F on $\mathbb{R}^{n+1}_+$, we have

$$\int_{\mathbb{R}^{n+1}_+} F(x,t) d\mu(x,t) = \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} |(\Psi_{2^{-j}} * b)(x)|^2 F(x, 2^{-j}) dx, \tag{7.33}$$

since $\int_{\mathbb{R}^+} \delta_{2^{-j}}(t) F(x,t) dt = F(x, 2^{-j})$.

For a cube $Q \subset \mathbb{R}^n$, let Q^* be the cube with the same center and orientation whose side length is $3\sqrt{n}\ell(Q)$, where $\ell(Q)$ is the side length of Q. Fix a cube $Q \subset \mathbb{R}^n$, take $F = \chi_{T(Q)}$, and split b as

$$b = (b - \operatorname*{Avg}_Q b)\chi_{Q^*} + (b - \operatorname*{Avg}_Q b)\chi_{(Q^*)^c} + \operatorname*{Avg}_Q b.$$

Since Ψ has a mean value of zero, $\Psi_{2^{-j}} * \operatorname*{Avg}_Q b = 0$. Then, (7.33) gives

$$\mu(T(Q)) = \sum_{2^{-j}\leqslant \ell(Q)} \int_Q |\Psi_{2^{-j}} * b(x)|^2 dx \leqslant 2\Sigma_1 + 2\Sigma_2,$$

where

$$\Sigma_1 = \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} |\Psi_{2^{-j}} * ((b - \operatorname*{Avg}_Q b)\chi_{Q^*})(x)|^2 dx,$$

$$\Sigma_2 = \sum_{2^{-j}\leqslant \ell(Q)} \int_Q |\Psi_{2^{-j}} * ((b - \operatorname*{Avg}_Q b)\chi_{(Q^*)^c})(x)|^2 dx.$$

Using Plancherel's theorem twice and (7.29), we obtain

$$\begin{aligned}
\Sigma_1 =&(2\pi)^n \int_{\mathbb{R}^n} \sum_{j\in\mathbb{Z}} |\widehat{\Psi}(2^{-j}\eta)|^2 |\widehat{((b-\operatorname{Avg}_Q b)\chi_{Q^*})}(\eta)|^2 d\eta \\
\leqslant&(2\pi)^n \sup_{\xi} \sum_{j\in\mathbb{Z}} |\widehat{\Psi}(2^{-j}\xi)|^2 \int_{\mathbb{R}^n} |\widehat{((b-\operatorname{Avg}_Q b)\chi_{Q^*})}(\eta)|^2 d\eta \\
\leqslant&C_n B^2 \int_{Q^*} |b(x)-\operatorname{Avg}_Q b|^2 dx \\
\leqslant&C_n B^2 \left[\int_{Q^*} |b(x)-\operatorname{Avg}_{Q^*} b|^2 dx + |Q^*| |\operatorname{Avg}_{Q^*} b - \operatorname{Avg}_Q b|^2\right] \\
\leqslant&C_n B^2 \left[\int_{Q^*} |b(x)-\operatorname{Avg}_{Q^*} b|^2 dx + \|b\|_{\mathrm{BMO}}^2 |Q|\right] \\
\leqslant&C_n B^2 \|b\|_{\mathrm{BMO}}^2 |Q|,
\end{aligned}$$

in view of (7.5) and Corollary 7.25. To estimate Σ_2, we use the size estimate (7.28) of the function Ψ to obtain

$$|(\Psi_{2^{-j}} * (b-\operatorname{Avg}_Q b)\chi_{(Q^*)^c})(x)| \leqslant \int_{(Q^*)^c} \frac{A 2^{-j\delta} |b(y)-\operatorname{Avg}_Q b|}{(2^{-j}+|x-y|)^{n+\delta}} dy. \tag{7.34}$$

Denote c_Q as the center of Q; then, for $x \in Q$ and $y \in (Q^*)^c$, we obtain

$$\begin{aligned}
&|y-x| \geqslant |y-c_Q| - |c_Q - x| \\
\geqslant&\frac{1}{2}|y-c_Q| + \frac{3\sqrt{n}}{4}\ell(Q) - |c_Q-x| \; (\because |y-c_Q| \geqslant \frac{1}{2}\ell(Q^*) = \frac{3\sqrt{n}}{2}\ell(Q)) \\
\geqslant&\frac{1}{2}|y-c_Q| + \frac{3\sqrt{n}}{4}\ell(Q) - \frac{\sqrt{n}}{2}\ell(Q) \\
=&\frac{1}{2}\left(|y-c_Q| + \frac{\sqrt{n}}{2}\ell(Q)\right) \\
\geqslant&\frac{1}{4}\left(|y-c_Q| + \ell(Q)\right).
\end{aligned}$$

Inserting this estimate in (7.34), integrating over Q, and summing over j with $2^{-j} \leqslant \ell(Q)$, we obtain

$$\begin{aligned}
\Sigma_2 \leqslant&C_n A^2 \sum_{j:2^{-j}\leqslant \ell(Q)} \int_Q \left(\int_{\mathbb{R}^n} \frac{2^{-j\delta}|b(y)-\operatorname{Avg}_Q b|}{(\ell(Q)+|c_Q-y|)^{n+\delta}} dy\right)^2 dx \\
\leqslant&C_n A^2 |Q| \left(\int_{\mathbb{R}^n} \frac{\ell(Q)^\delta |b(y)-\operatorname{Avg}_Q b|}{(\ell(Q)+|c_Q-y|)^{n+\delta}} dy\right)^2 \\
\leqslant&C_{n,\delta} A^2 |Q| \|b\|_{\mathrm{BMO}}^2
\end{aligned}$$

in view of (7.7). This proves that

$$\Sigma_1 + \Sigma_2 \leqslant C_{n,\delta}(A^2+B^2)|Q|\|b\|_{\mathrm{BMO}}^2,$$

which implies that $\mu(T(Q)) \leqslant C_{n,\delta}(A^2+B^2)|Q|\|b\|_{\mathrm{BMO}}^2$; thus, μ is a Carleson measure.

(ii) The proof can be obtained in a similar fashion as in (i).

(iii) This is a generalization of (ii) and is proven likewise. We sketch its proof. Write

$$b = (b - \operatorname*{Avg}_Q b)\chi_{Q^*} + (b - \operatorname*{Avg}_Q b)\chi_{(Q^*)^c} + \operatorname*{Avg}_Q b$$

and note that $R_t(\operatorname*{Avg}_Q b) = 0$ because $R_t(1) = 0$. We handle the term containing $R_t((b - \operatorname*{Avg}_Q b)\chi_{Q^*})$ by using an L^2 estimate over Q^* and condition (7.32), while for the term containing $R_t((b - \operatorname*{Avg}_Q b)\chi_{(Q^*)^c})$, we use an L^1 estimate and condition (7.31). In both cases, we obtain the required conclusion in a way analogous to that in (i). ❑

Exercises

Exercise 7.1. Prove Proposition 7.5. (Notice: Do NOT use Theorem 7.8!)

Exercise 7.2. [Gra14b, Exercise 2.1.7(a)] Let $1 < q \leqslant \infty$ and let $g \in L^q(\mathbb{R}^n)$ be a compactly supported function with integral zero. Show that $g \in \mathcal{H}^1(\mathbb{R}^n)$.

Hint Pick a function $\Phi \in \mathscr{D}$ supported in the unit ball with a nonvanishing integral and suppose that $\operatorname{supp} g \subset B(0,R)$. For $|x| \leqslant 2R$, we have that $M(g;\Phi)(x) \leqslant C_\Phi Mg(x)$, and since $Mg \in L^q$, it also lies in $L^1(B(0,2R))$. For $|x| > 2R$, write $(\Phi_t * g)(x) = \int_{\mathbb{R}^n}(\Phi_t(x-y) - \Phi_t(x))g(y)dy$ and use the mean value theorem to estimate this expression by

$$t^{-n-1}\|\nabla\Phi\|_\infty\|g\|_1 \leqslant |x|^{-n-1}C_\Phi\|g\|_q,$$

since $t \geqslant |x-y| \geqslant |x| - |y| \geqslant |x|/2$ whenever $|x| \geqslant 2R$ and $|y| \leqslant R$. Thus, $M(g,\Phi) \in L^1(\mathbb{R}^n)$ and then $g \in \mathcal{H}^1$ by Theorem 7.11.

Exercise 7.3. Prove (x) in Proposition 7.15.

Exercise 7.4. [Pey18, Exercise 6.8] Prove that $\mathrm{BMO}(\mathbb{R}^n)$ is complete.

Exercise 7.5. [Gra14b, Exercise 3.1.6] Let $a > 1$ and $f \in \mathrm{BMO}(\mathbb{R}^n)$. Show that there exist dimensional constants C_n, C_n' such that

(i) for all balls B_1 and B_2 in $\mathbb{R}^n$ with radius R whose centers are at distance aR we have

$$|\operatorname*{Avg}_{B_1} f - \operatorname*{Avg}_{B_2} f| \leqslant C_n' \ln(a+1)\|f\|_{\mathrm{BMO}}.$$

(ii) Conclude that

$$|\operatorname*{Avg}_{(a+1)B_1} f - \operatorname*{Avg}_{B_2} f| \leqslant C_n \ln(a+1)\|f\|_{\mathrm{BMO}}.$$

Hint (i) Replace $\operatorname{Avg}_{B_1} f$ by $\operatorname{Avg}_{2aB_1} f$ and $\operatorname{Avg}_{B_2} f$ by $\operatorname{Avg}_{aB_2} f$ and use the fact that $aB_2 \subset 2aB_1$ and use Proposition 7.20 (i).

Exercise 7.6. [Pey18, Exercise 6.5] Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ be real valued only. One assumes that there exists a constant C_0 such that the following hold:

- for all x, the function $y \mapsto f(x,y)$ lies in BMO with a norm less than or equal to C_0;
- for all y, the function $x \mapsto f(x,y)$ lies in BMO with a norm less than or equal to C_0.

(i) Show that if $Q = I \times J$ is a square in $\mathbb{R}^2$, one has

$$\frac{1}{|J|}\int_J \left| \frac{1}{|I|}\int_I f(x,y)dx - \frac{1}{|Q|}\int_Q f(x,t)dxdt \right| dy \leqslant C_0.$$

(ii) Deduce from this that $\|f\|_{\mathrm{BMO}(\mathbb{R}^2)} \leqslant 2C_0$.

(iii) Let P be a polynomial in two variables with complex coefficients. Show that $\ln|P|$ is in BMO and give an upper bound of its norm.

Exercise 7.7. [Pey18, Exercise 6.6] For $\alpha > 0$, set $g_\alpha(x) = \ln(x^2 + \alpha)$.

(i) Show that, for all a and b such that $0 \leqslant a \leqslant b$, one has

$$\int_a^b (\ln(1+b^2) - \ln(1+x^2))dx \leqslant 2(b-a).$$

Conclude that $g_1 \in \mathrm{BMO}$.

(ii) Show that, for all $\alpha > 0$, one has $\|g_\alpha\|_{\mathrm{BMO}} = \|g_1\|_{\mathrm{BMO}}$. Deduce from this fact that $\|g_\alpha\|_{\mathrm{BMO}} = 2\|\ln|x|\|_{\mathrm{BMO}}$.

(iii) Show that if P is a polynomial of degree n with complex coefficients (the coefficient is 1 for the term with the highest degree), then $\ln|P|$ is in BMO with a norm not exceeding $n\|\ln|x|\|_{\mathrm{BMO}}$.

Hint (ii) Use the monotonicity of the integral and Fatou's lemma. (iii) To factor P and use (ii).

Exercise 7.8. [Pey18, Exercise 6.11] Show that $(\ln|x|)^2$ is not in $\mathrm{BMO}(\mathbb{R})$.

Hint This function does not fulfil the John-Nirenberg inequality.

8

Standard Kernels and $T(1)$ Theorem

We study singular integrals whose kernels do not necessarily commute with translations. Such operators appear in many places in harmonic analysis and PDEs. For instance, a large class of pseudodifferential operators fall under the scope of this theory.

This broader point of view does not necessarily bring additional complications in the development of the subject except in the study of L^2 boundedness, where Fourier transform techniques are lacking. The L^2 boundedness of convolution operators is easily understood via a careful examination of the Fourier transform of the kernel, but for nonconvolution operators different tools are required in this study. The main result of this chapter is the derivation of a set of necessary and sufficient conditions for non-convolution singular integrals to be L^2 bounded. This result is referred to as the $T(1)$ theorem and owes its name to a condition expressed in terms of the action of operator T on function 1.

§ 8.1 General background and the role of BMO

We begin by recalling the notion of the adjoint and transpose operator. One may choose to work with either a real or a complex inner product on pairs of functions. For f, g complex-valued functions with integrable products, we denote the real inner product by

$$\langle f, g\rangle = \int_{\mathbb{R}^n} f(x)g(x)dx.$$

This notation is suitable when we think of f as a distribution acting on a test function g. We also have the complex inner product

$$(f, g) = \int_{\mathbb{R}^n} f(x)\overline{g(x)}dx,$$

which is an appropriate notation when we think of f and g as elements of a Hilbert space over the complex numbers. Now suppose that T is a linear operator bounded on L^p. Then the adjoint operator T^* of T is uniquely defined via the identity

$$(Tf, g) = (f, T^*g)$$

for all $f \in L^p$ and $g \in L^{p'}$. The transpose operator T^t of T is uniquely defined via the identity

$$\langle Tf, g\rangle = \langle f, T^t g\rangle = \langle T^t g, f\rangle$$

for all $f \in L^p$ and $g \in L^{p'}$. We easily have the following relationship between the transpose and the adjoint of a linear operator T:

$$T^* f = \overline{T^t \bar{f}},$$

indicating that they have almost interchangeable use. In fact, it holds

$$(g, T^* f) = (Tg, f) = \langle Tg, \bar{f}\rangle = \langle g, T^t \bar{f}\rangle = (g, \overline{T^t \bar{f}}),$$

which also implies that

$$\overline{T^* f} = T^t \bar{f}, \quad T^t f = \overline{T^* \bar{f}}.$$

Because of these, in many cases, it is convenient to avoid complex conjugates and work with the transpose operator for simplicity. Observe that if a linear operator T has kernel $K(x, y)$, that is,

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

then the kernel of T^t is $K^t(x, y) = K(y, x)$ and that of T^* is $K^*(x, y) = \overline{K(y, x)}$. Indeed, we have for $f, g \in \mathscr{S}(\mathbb{R}^n)$

$$\begin{aligned}\langle T^t g, f\rangle =& \langle Tf, g\rangle = \iint K(x, y) f(y) dy\, g(x) dx = \iint K(x, y) g(x) dx\, f(y) dy\\ =& \iint K(y, x) g(y) dy\, f(x) dx = \left\langle \int K(y, \cdot) g(y) dy, f \right\rangle,\end{aligned}$$

and

$$\begin{aligned}(f, T^* g) =& (Tf, g) = \iint K(x, y) f(y) dy\, \overline{g(x)} dx = \iint K(x, y) \overline{g(x)} dx\, f(y) dy\\ =& \iint K(y, x) \overline{g(y)} dy\, f(x) dx = \left(f, \int \overline{K(y, x)} g(y) dy \right).\end{aligned}$$

An operator is called ***self-adjoint*** if $T = T^*$ and ***self-transpose*** if $T = T^t$. For example, the operator iH, where H is the Hilbert transform, is self-adjoint but not self-transpose, and the operator with kernel $i(x+y)^{-1}$ is self-transpose but not self-adjoint.

§ 8.1.1 Standard kernels

The singular integrals we study in this chapter have kernels that satisfy size and regularity properties similar to those of Calderón-Zygmund operators of the classical convolution type. We introduce the relevant background.

Definition 8.1 (Standard kernels)**.** The function $K(x,y)$ defined on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x,x) : x \in \mathbb{R}^n\}$ that satisfies for some $A > 0$ ***the standard size condition***

$$|K(x,y)| \leqslant \frac{A}{|x-y|^n} \tag{8.1}$$

and for some $\delta > 0$ ***the (Hölder) regularity conditions***

$$|K(x,y) - K(x',y)| \leqslant \frac{A|x-x'|^\delta}{(|x-y|+|x'-y|)^{n+\delta}}, \tag{8.2}$$

whenever $|x-x'| \leqslant \frac{1}{2}\max(|x-y|,|x'-y|)$ and

$$|K(x,y) - K(x,y')| \leqslant \frac{A|y-y'|^\delta}{(|x-y|+|x-y'|)^{n+\delta}}, \tag{8.3}$$

whenever $|y-y'| \leqslant \frac{1}{2}\max(|x-y|,|x-y'|)$, is called a ***standard kernel*** (or ***Calderón-Zygmund kernel***) with constants δ, A. The class of all standard kernels with constants δ, A is denoted by $\mathrm{SK}(\delta, A)$.

Given a kernel $K(x,y)$ in $\mathrm{SK}(\delta,A)$, we observe that the functions $K(y,x)$ and $\overline{K(y,x)}$ are also in $\mathrm{SK}(\delta,A)$. These kernels have special names. The function $K^t(x,y) = K(y,x)$ is called the ***transpose kernel*** of K, while the function $K^*(x,y) = \overline{K(y,x)}$ is called the ***adjoint kernel*** of K.

Remark 8.2. (i) Observe that if $|x-x'| \leqslant \frac{1}{2}\max(|x-y|,|x'-y|)$, then

$$\begin{aligned}
2\min(|x-y|,|x'-y|) =& |x-y| + |x'-y| - \big||x-y| - |x'-y|\big| \\
\geqslant& |x-y| + |x'-y| - |x-x'| \\
\geqslant& |x-y| + |x'-y| - \frac{1}{2}\max(|x-y|,|x'-y|) \\
=& \max(|x-y|,|x'-y|) + \min(|x-y|,|x'-y|) \\
&- \frac{1}{2}\max(|x-y|,|x'-y|),
\end{aligned}$$

which yields

$$\max(|x-y|,|x'-y|) \leqslant 2\min(|x-y|,|x'-y|),$$

and

$$\begin{aligned}
\frac{1}{2}|x-y| \leqslant& \min(|x-y|,|x'-y|) \leqslant |x'-y| \\
\leqslant& \max(|x-y|,|x'-y|) \leqslant 2|x-y|,
\end{aligned}$$

i.e., the numbers $|x-y|$ and $|x'-y|$ are comparable. Likewise, if the roles of x and y are interchanged.

(ii) If (8.1) holds, we assume

$$|\nabla_x K(x,y)| + |\nabla_y K(x,y)| \leqslant \frac{A}{|x-y|^{n+1}}, \quad \forall x \neq y,$$

then $K \in \mathrm{SK}(1, 4^{n+1}A)$. In fact, if $|x - x'| \leqslant \frac{1}{2}\max(|x-y|, |x'-y|)$, by the mean value theorem, we obtain, for some $\theta = \sigma x + (1-\sigma)x'$ and $\sigma \in [0,1]$, that

$$|K(x,y) - K(x',y)| \leqslant |\nabla_x K(\theta, y)||x - x'| \leqslant \frac{A|x-x'|}{|\theta - y|^{n+1}},$$

and from

$$\begin{aligned}
|\theta - y| =& |\sigma x + (1-\sigma)x' - y| \\
=& \frac{1}{2}[|\sigma(x-x') + x' - y| + |(1-\sigma)(x'-x) + x - y|] \\
\geqslant& \frac{1}{2}[|x'-y| - \sigma|x-x'| + |x-y| - (1-\sigma)|x-x'|] \\
=& \frac{1}{2}[|x-y| + |x'-y| - |x-x'|] \\
\geqslant& \frac{1}{2}\left[|x-y| + |x'-y| - \frac{1}{2}\max(|x-y|, |x'-y|)\right] \\
\geqslant& \frac{1}{4}(|x-y| + |x'-y|),
\end{aligned} \tag{8.4}$$

it follows that

$$|K(x,y) - K(x',y)| \leqslant \frac{4^{n+1}A|x-x'|}{(|x-y| + |x'-y|)^{n+1}},$$

i.e., (8.1). Similarly, we also have (8.2) whenever $|y - y'| \leqslant \frac{1}{2}\max(|x - y|, |x - y'|)$. Thus, $K \in \mathrm{SK}(1, 4^{n+1}A)$.

Example 8.3. The function $K(x,y) = |x-y|^{-n}$ defined away from the diagonal of $\mathbb{R}^n \times \mathbb{R}^n$ is in $\mathrm{SK}(1, n4^{n+1})$. Indeed, for $|x - x'| \leqslant \frac{1}{2}\max(|x - y|, |x'-y|)$, the mean value theorem gives

$$\left||x-y|^{-n} - |x'-y|^{-n}\right| \leqslant \frac{n|x-x'|}{|\theta - y|^{n+1}}$$

for some $\theta = \sigma x + (1-\sigma)x'$ and $\sigma \in [0,1]$. From (8.4), it follows (8.2) with $A = n4^{n+1}$ and $\delta = 1$. Likewise, (8.3) holds.

We are interested in standard kernels $K \in \mathrm{SK}(\delta, A)$ for which there are tempered distributions $W \in \mathscr{S}'(\mathbb{R}^n \times \mathbb{R}^n)$ that coincide with K on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x,x) : x \in \mathbb{R}^n\}$. This means that the convergent integral representation

$$\langle W, F \rangle = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)F(x,y)dxdy \tag{8.5}$$

is valid whenever the Schwartz function $F \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^n)$ is supported away from the diagonal $\{(x,x) : x \in \mathbb{R}^n\}$. Note that the integral in (8.5) is well-defined and absolutely convergent whenever F is a Schwartz function whose support does not intersect the set $\{(x,x) : x \in \mathbb{R}^n\}$. Additionally, observe that there may be several distributions W coinciding with a fixed function $K(x,y)$. In fact, if W is such a distribution, then so is $W + \delta_{x=y}$, where $\delta_{x=y}$ denotes the Lebesgue measure on the diagonal of $\mathbb{R}^{2n}$ (i.e., some kind of Dirac distribution).

We recall a result about the Schwartz kernel (cf. [Hö3, pp.128-130]) as follows.

Theorem 8.4 (Schwartz kernel theorem). *Every $K \in \mathscr{D}'(X_1 \times X_2)$ defines according to*

$$\langle \mathcal{K}\phi, \psi \rangle = \langle K, \psi \otimes \phi \rangle, \quad \psi \in \mathscr{D}(X_1),\ \phi \in \mathscr{D}(X_2), \tag{8.6}$$

a linear map $\mathcal{K}$ from $\mathscr{D}(X_2)$ to $\mathscr{D}'(X_1)$, which is continuous in the sense that $\mathcal{K}\phi_j \to 0$ in $\mathscr{D}'(X_1)$ if $\phi_j \to 0$ in $\mathscr{D}(X_2)$. Conversely, for every such linear map $\mathcal{K}$, there is one and only one distribution K such that (8.6) is valid. One calls K the kernel of $\mathcal{K}$.

Remark 8.5. The same theorem also holds when $\mathscr{D}$, $\mathscr{D}'$, X_1 and X_2 replaced by $\mathscr{S}$, $\mathscr{S}'$, $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.[a] The Schwartz kernel theorem is a philosophically useful fact, establishing a 1-1 correspondence between the 'most general' operators in the present context and distributional integral kernels, also called ***Schwartz kernels***.

[a]One can see the proof in http://math.mit.edu/~eyjaffe/Short Notes/Distribution Theory/.

For continuous linear operators

$$T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n),$$

it follows that there is a distribution $W \in \mathscr{S}'(\mathbb{R}^{2n})$ satisfying

$$\langle Tf, \phi \rangle = \langle W, \phi \otimes f \rangle \tag{8.7}$$

for $f, \phi \in \mathscr{S}(\mathbb{R}^n)$, where $(\phi \otimes f)(x, y) = \phi(x) f(y)$ for all $x, y \in \mathbb{R}^n$, and there exist constants C, N, M such that for all $f, g \in \mathscr{S}(\mathbb{R}^n)$, we have

$$|\langle Tf, g \rangle| = |\langle W, g \otimes f \rangle| \leqslant C \Bigg[\sum_{|\alpha|, |\beta| \leqslant N} |g|_{\alpha,\beta} \Bigg] \Bigg[\sum_{|\alpha|, |\beta| \leqslant N} |f|_{\alpha,\beta} \Bigg]. \tag{8.8}$$

Here $|\phi|_{\alpha,\beta} = \sup_{x \in \mathbb{R}^n} |\partial_x^\alpha (x^\beta \phi)(x)|$ are the seminorms for the topology in $\mathscr{S}$. A distribution W that satisfies (8.7) and (8.8) is called a ***Schwartz kernel*** or the ***distributional kernel*** of T.

Here, we study continuous linear operators $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ whose distributional kernels coincide with standard kernels $K(x, y)$ on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x, x) : x \in \mathbb{R}^n\}$. This means that (8.7) admits the absolutely convergent integral representation

$$\langle Tf, \phi \rangle = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x, y) f(y) \phi(x) dx dy \tag{8.9}$$

whenever f and ϕ are Schwartz functions whose supports do not intersect.

We make some remarks concerning duality in this context. Given a continuous linear operator $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ with distributional kernel W, we can define another distribution W^t as follows:

$$\langle W^t, F \rangle = \langle W, F^t \rangle,$$

where $F^t(x,y) = F(y,x)$. This implies that for all $f, \phi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\langle W, \phi \otimes f \rangle = \langle W^t, f \otimes \phi \rangle.$$

It is a simple fact that the transpose operator T^t of T, which satisfies

$$\langle T\phi, f \rangle = \langle T^t f, \phi \rangle \tag{8.10}$$

for all $f, \phi \in \mathscr{S}(\mathbb{R}^n)$, is the unique continuous linear operator from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ whose Schwartz kernel is the distribution W^t, that is, we have

$$\langle T^t f, \phi \rangle = \langle T\phi, f \rangle = \langle W, f \otimes \phi \rangle = \langle W^t, \phi \otimes f \rangle. \tag{8.11}$$

We now observe that a large class of standard kernels admits extensions to tempered distributions W on $\mathbb{R}^{2n}$.

Example 8.6. Suppose that $K(x,y)$ satisfies (8.1) and (8.2) and is anti-symmetric in the sense that

$$K(x,y) = -K(y,x)$$

for all $x \neq y$ in $\mathbb{R}^n$. Then K also satisfies (8.3), and moreover, there is a distribution W on $\mathbb{R}^{2n}$ that extends K on $\mathbb{R}^n \times \mathbb{R}^n$.

Indeed, define

$$\langle W, F \rangle = \lim_{\varepsilon \to 0} \iint_{|x-y|>\varepsilon} K(x,y) F(x,y) dydx \tag{8.12}$$

for all $F \in \mathscr{S}(\mathbb{R}^{2n})$. In view of anti-symmetry, we may write

$$\iint_{|x-y|>\varepsilon} K(x,y)F(x,y)dydx = \frac{1}{2} \iint_{|x-y|>\varepsilon} K(x,y)[F(x,y) - F(y,x)]dydx.$$

By the mean value theorem, it holds for some $\theta, \sigma \in [0,1]$ and $F \in \mathscr{S}(\mathbb{R}^{2n})$

$$\begin{aligned} & |F(x,y) - F(y,x)| \\ \leqslant & |F(x,y) - F(y,y)| + |F(y,y) - F(y,x)| \\ = & [|\nabla_x F(\theta x + (1-\theta)y, y)| + |\nabla_y F(y, \sigma y + (1-\sigma)x)|]|x-y| \\ \leqslant & \frac{C|x-y|}{(1+|x|^2+|y|^2)^{n+1}}. \end{aligned}$$

Then, by (8.1), we have

$$\begin{aligned} |\langle W, F \rangle| \leqslant & \lim_{\varepsilon \to 0} C \iint_{|x-y|>\varepsilon} \frac{A}{|x-y|^{n-1}} \frac{1}{(1+|x|^2+|y|^2)^{n+1}} dydx \\ \leqslant & CA \lim_{\varepsilon \to 0} \iint_{\varepsilon<|x-y|\leqslant 1} \frac{1}{|x-y|^{n-1}} \frac{1}{(1+|x|^2)^{n+1}} dydx \end{aligned}$$

$$+ CA \iint_{|x-y|>1} \frac{1}{(1+|x|^2+|y|^2)^{n+1}} dydx$$
$$\leqslant CA.$$

Thus, the limit in (8.12) exists and gives a tempered distribution on $\mathbb{R}^{2n}$. We can then define an operator $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ with kernel W via

$$\begin{aligned}\langle Tf, \phi\rangle =& \lim_{\varepsilon\to 0} \iint_{|x-y|>\varepsilon} K(x,y) f(y)\phi(x) dydx \\ =& \frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)[f(y)\phi(x) - f(x)\phi(y)] dydx,\end{aligned}$$

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

Example 8.7. Let A be a real-valued Lipschitz function on $\mathbb{R}$. This means that it satisfies the estimate $|A(x) - A(y)| \leqslant L|x-y|$ for some $L < \infty$ and all $x, y \in \mathbb{R}$. For $x, y \in \mathbb{R}$, $x \neq y$, let

$$K_A(x,y) = \frac{1}{x - y + i(A(x) - A(y))}. \tag{8.13}$$

When $|y-y'| \leqslant \frac{1}{2}\max(|x-y|, |x-y'|)$, which yields $\frac{1}{2}|x-y'| \leqslant |x-y| \leqslant 2|x-y'|$ by Remark 8.2, it holds

$$\begin{aligned}&|K_A(x,y) - K_A(x,y')| \\ =& \left|\frac{1}{x-y+i(A(x)-A(y))} - \frac{1}{x-y'+i(A(x)-A(y'))}\right| \\ =& \left|\frac{y-y'+i(A(y)-A(y'))}{[x-y+i(A(x)-A(y))][x-y'+i(A(x)-A(y'))]}\right| \\ =& \frac{[(y-y')^2+(A(y)-A(y'))^2]^{1/2}}{[(x-y)^2+(A(x)-A(y))^2]^{1/2}[(x-y')^2+(A(x)-A(y'))^2]^{1/2}} \\ \leqslant& \frac{[1+L^2]^{1/2}|y-y'|}{|x-y||x-y'|} \leqslant \frac{4(1+L)|y-y'|}{|x-y|^2+|x-y'|^2} \\ \leqslant& \frac{8(1+L)|y-y'|}{(|x-y|+|x-y'|)^2}.\end{aligned}$$

Since K_A is anti-symmetric, it follows that K_A satisfies (8.2), and then $K_A \in \mathrm{SK}(1, 8(1+L))$.

Example 8.8. Let function A be as in the previous example. For each integer $m \geqslant 1$ and $x, y \in \mathbb{R}$, we set

$$K_m(x,y) = \left(\frac{A(x)-A(y)}{x-y}\right)^m \frac{1}{x-y}. \tag{8.14}$$

Clearly, K_m is an anti-symmetric function. To see that each K_m is a standard kernel, notice that when $|y-y'| \leqslant \frac{1}{2}\max(|x-y|, |x-y'|)$ we have

$$\begin{aligned}&|K_m(x,y) - K_m(x,y')| \\ =& \left|\left(\frac{A(x)-A(y)}{x-y}\right)^m \frac{1}{x-y} - \left(\frac{A(x)-A(y')}{x-y'}\right)^m \frac{1}{x-y'}\right|\end{aligned}$$

$$
\begin{aligned}
&\leqslant \left| \left(\frac{A(x)-A(y)}{x-y} \right)^m - \left(\frac{A(x)-A(y')}{x-y'} \right)^m \right| \frac{1}{|x-y|} \\
&\quad + \left| \frac{A(x)-A(y')}{x-y'} \right|^m \left| \frac{1}{x-y} - \frac{1}{x-y'} \right| \\
&\leqslant \left| \frac{A(x)-A(y)}{x-y} - \frac{A(x)-A(y')}{x-y'} \right| \frac{1}{|x-y|} \\
&\quad \cdot \sum_{j=0}^{m-1} \left| \frac{A(x)-A(y)}{x-y} \right|^{m-1-j} \left| \frac{A(x)-A(y')}{x-y'} \right|^j \\
&\quad + \left| \frac{A(x)-A(y')}{x-y'} \right|^m \frac{|y-y'|}{|x-y||x-y'|} \\
&\leqslant mL^{m-1} \frac{1}{|x-y|} \frac{|(x-y')(A(x)-A(y)) - (x-y)(A(x)-A(y'))|}{|x-y||x-y'|} \\
&\quad + L^m \frac{|y-y'|}{|x-y||x-y'|} \\
&= mL^{m-1} \frac{1}{|x-y|} \frac{|(y-y')(A(x)-A(y)) + (x-y)(A(y')-A(y))|}{|x-y||x-y'|} \\
&\quad + L^m \frac{|y-y'|}{|x-y||x-y'|} \\
&\leqslant 2mL^m \frac{|y-y'|}{|x-y||x-y'|} + L^m \frac{|y-y'|}{|x-y||x-y'|} \\
&\leqslant 8(2m+1)L^m \frac{|y-y'|}{(|x-y|+|x-y'|)^2}.
\end{aligned}
$$

It follows that $K_m \in \mathrm{SK}(1, 8(2m+1)L^m)$. The linear operator with kernel $(\pi i)^{-1}K_m$ is called the ***m-th Calderón commutator***.

§ 8.1.2 Operators associated with standard kernels

Having introduced standard kernels, we are in a position to define linear operators associated with them.

Definition 8.9. Let $0 < \delta, A < \infty$, and $K \in \mathrm{SK}(\delta, A)$. A continuous linear operator T from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ is said to be ***associated with*** K if it satisfies

$$
Tf(x) = \int_{\mathbb{R}^n} K(x,y) f(y) dy \tag{8.15}
$$

for all $f \in \mathscr{D}$ and $x \notin \operatorname{supp} f$. If T is associated with K, then the Schwartz kernel W of T coincides with K on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x,x) : x \in \mathbb{R}^n\}$.

If T is associated with K and satisfies

$$
\|T\phi\|_2 \leqslant B\|\phi\|_2 \tag{8.16}
$$

for all $\phi \in \mathscr{S}(\mathbb{R}^n)$, then T is called a ***Calderón-Zygmund operator*** associated with the standard kernel K. Such operators T admit a bounded extension on $L^2(\mathbb{R}^n)$, i.e., given any $f \in L^2(\mathbb{R}^n)$ one can define Tf as the unique L^2 limit of the Cauchy sequence $\{T\phi_k\}_k$, where $\phi_k \in \mathscr{S}(\mathbb{R}^n)$ and $\phi_k \to f$ in L^2. In this case, we keep the same notation for the L^2 extension of T.

In the sequel, we denote by $\mathrm{CZO}(\delta, A, B)$ the class of all Calderón-Zygmund operators associated with standard kernels in $\mathrm{SK}(\delta, A)$ that admit L^2-bounded extensions with norm at most B.

Next, we discuss the important fact that once an operator T admits an extension that is L^2 bounded, then (8.15) holds for all f that are bounded and compactly supported whenever $x \notin \operatorname{supp} f$.

Proposition 8.10. *Let $T \in \mathrm{CZO}(\delta, A, B)$ be associated with $K \in \mathrm{SK}(\delta, A)$. Then for every bounded and compactly supported function f and ϕ that satisfy*

$$\operatorname{dist}(\operatorname{supp} \phi, \operatorname{supp} f) > 0, \tag{8.17}$$

then we have the (absolutely convergent) integral representation

$$\int_{\mathbb{R}^n} Tf(x)\phi(x)dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f(y)\phi(x)dydx. \tag{8.18}$$

Moreover, given any bounded function f with compact support, there is a set of measure zero $E(f)$ such that $x_0 \notin E(f) \bigcup \operatorname{supp} f$ we have the (absolutely convergent) integral representation

$$Tf(x_0) = \int_{\mathbb{R}^n} K(x_0,y)f(y)dy. \tag{8.19}$$

Proof. We first prove (8.18). Given f and ϕ bounded functions with compact support, we select $f_j, \phi_j \in \mathscr{D}$ such that ϕ_j are uniformly bounded and supported in a small neighborhood of the support of ϕ, $\phi_j \to \phi$ in L^2 and a.e., $f_j \to f$ in L^2 and a.e., and

$$\operatorname{dist}(\operatorname{supp} \phi_j, \operatorname{supp} f_j) \geqslant \frac{1}{2}\operatorname{dist}(\operatorname{supp} \phi, \operatorname{supp} f) = c > 0 \tag{8.20}$$

for all $j \in \mathbb{Z}^+$. In view of (8.9), identity (8.18) is valid for the functions f_j and ϕ_j in place of f and ϕ, i.e.,

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f_j(y)\phi_j(x)dydx = \int_{\mathbb{R}^n} Tf_j(x)\phi_j(x)dx. \tag{8.21}$$

By the boundedness of T, it follows that Tf_j converges to Tf in L^2 and thus as $j \to \infty$ we have

$$\left|\int_{\mathbb{R}^n} Tf_j(x)\phi_j(x)dx - \int_{\mathbb{R}^n} Tf(x)\phi(x)dx\right|$$

$$\begin{aligned}&\leqslant \int_{\mathbb{R}^n} |Tf_j(x) - Tf(x)|\,|\phi_j(x)|dx + \int_{\mathbb{R}^n} |Tf(x)|\,|\phi_j(x) - \phi(x)|\,dx\\ &\leqslant \|Tf_j - Tf\|_2\|\phi_j\|_2 + \|Tf\|_2\|\phi_j - \phi\|_2 \to 0,\end{aligned}$$

i.e.,

$$\int_{\mathbb{R}^n} Tf_j(x)\phi_j(x)dx \to \int_{\mathbb{R}^n} Tf(x)\phi(x)dx. \tag{8.22}$$

From (8.20), it follows that as $j \to \infty$

$$\begin{aligned}&\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f_j(y)\phi_j(x)dydx - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f(y)\phi(x)dydx\right|\\ \leqslant&\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)[f_j(y) - f(y)]\phi_j(x)dydx\right|\\ &+\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f(y)[\phi_j(x) - \phi(x)]dydx\right|\\ \leqslant& Ac^{-n}\left(\|f_j - f\|_1\|\phi_j\|_1 + \|f\|_1\|\phi_j - \phi\|_1\right)\\ \leqslant& CAc^{-n}\left(\|f_j - f\|_2\|\phi_j\|_2 + \|f\|_2\|\phi_j - \phi\|_2\right) \to 0,\end{aligned}$$

which proves the validity of (8.18). Note that the double integral on the right of (8.18) is absolutely convergent and bounded by $A(2c)^{-n}\|f\|_1\|\phi\|_1$ in view of (8.20).

To prove (8.19), we fix a compactly supported and bounded function f, and we pick f_j as before. Then, $Tf_j \to Tf$ in L^2, and thus, a subsequence Tf_{jl} converges pointwise on $\mathbb{R}^n \setminus E(f)$ by Riesz's theorem, for some measurable set $E(f)$ with $|E(f)| = 0$. Given $x_0 \notin E(f) \bigcup \operatorname{supp} f$, we have

$$Tf_{jl}(x_0) = \int_{\mathbb{R}^n} K(x_0,y)f_{jl}(y)dy$$

and letting $l \to \infty$, we obtain (8.19) since $Tf_{jl}(x_0) \to Tf(x_0)$ and

$$\begin{aligned}&\left|\int_{\mathbb{R}^n} K(x_0,y)f_{jl}(y)dy - \int_{\mathbb{R}^n} K(x_0,y)f(y)dy\right|\\ \leqslant& Ac^{-n}\|f_{jl} - f\|_1 \leqslant CAc^{-n}\|f_{jl} - f\|_2 \to 0,\end{aligned}$$

as $l \to \infty$. Thus, (8.19) holds. ❑

We now define truncated kernels and operators.

Definition 8.11. Given a kernel $K \in \mathrm{SK}(\delta, A)$ and $\varepsilon > 0$, we define the ***truncated kernel***

$$K^{(\varepsilon)}(x,y) = K(x,y)\chi_{|x-y|>\varepsilon}.$$

Given a continuous linear operator T from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ and $\varepsilon > 0$, we define the ***truncated operator*** $T^{(\varepsilon)}$ by

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

and the ***maximal singular operator*** associated with T as follows:

$$T^{(*)}f(x) = \sup_{\varepsilon>0} \left| T^{(\varepsilon)} f(x) \right|.$$

Note that both $T^{(\varepsilon)}$ and $T^{(*)}$ are well-defined for $f \in \bigcup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$ by an application of Hölder's inequality.

We investigate a certain connection between the boundedness of T and the boundedness of the family $\{T^{(\varepsilon)}\}_{\varepsilon>0}$ uniformly in $\varepsilon > 0$.

Proposition 8.12. *Let $T \in \mathrm{CZO}(\delta, A, B)$ be associated with $K \in \mathrm{SK}(\delta, A)$. For $\varepsilon > 0$, let $T^{(\varepsilon)}$ be the truncated operators obtained from T. Assume that there exists a constant $B' < \infty$ such that*

$$\sup_{\varepsilon>0} \|T^{(\varepsilon)}\|_{L^2\to L^2} \leqslant B'. \tag{8.23}$$

Then, there exists a linear operator T_0 defined on $L^2(\mathbb{R}^n)$ such that

(i) *For some subsequence $\varepsilon_j \searrow 0$, we have*

$$\int_{\mathbb{R}^n} T^{(\varepsilon_j)} f(x)g(x)dx \to \int_{\mathbb{R}^n} T_0 f(x)g(x)dx \tag{8.24}$$

as $j \to \infty$ for all $f, g \in L^2(\mathbb{R}^n)$.

(ii) *The distributional kernel of T_0 coincides with K on*

$$\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x, x) : x \in \mathbb{R}^n\}.$$

(iii) *T_0 is bounded on $L^2(\mathbb{R}^n)$ with norm*

$$\|T_0\|_{L^2\to L^2} \leqslant B'.$$

(iv) *There exists a measurable function b on $\mathbb{R}^n$ with $\|b\|_\infty \leqslant B + B'$ such that*

$$T_0 f - Tf = bf,$$

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

Proof. (i) Since $L^2(\mathbb{R}^n)$ is separable, let $\{f_k\}_{k=1}^\infty$ be a dense countable subset of $L^2(\mathbb{R}^n)$. By (8.23), the functions $T^{(\varepsilon)} f_k$ lie in multiple of the unit closed ball of $(L^2)^*$, which is weak* compact by the Banach-Alaoglu theorem, and then has a weak* converging subsequence due to the separability of $L^2(\mathbb{R}^n)$. Hence for each f_k, we find a sequence $\{\varepsilon_j^k\}_{j=1}^\infty$ such that for each $g \in L^2(\mathbb{R}^n)$, we have

$$\lim_{j\to\infty} \int_{\mathbb{R}^n} T^{(\varepsilon_j^k)} f_k(x)g(x)dx = \int_{\mathbb{R}^n} T_0^{f_k}(x)g(x)dx, \tag{8.25}$$

for some function $T_0^{f_k} \in L^2(\mathbb{R}^n)$. Moreover, each $\{\varepsilon_j^k\}_{j=1}^\infty$ can be chosen to be a subsequence of $\{\varepsilon_j^{k-1}\}_{j=1}^\infty$, $k \geqslant 2$. Then, the diagonal sequence

$\{\varepsilon_j^j\}_{j=1}^\infty = \{\varepsilon_j\}_{j=1}^\infty$ satisfies

$$\lim_{j\to\infty}\int_{\mathbb{R}^n} T^{(\varepsilon_j)} f_k(x)g(x)dx = \int_{\mathbb{R}^n} T_0^{f_k}(x)g(x)dx \tag{8.26}$$

for each k and $g \in L^2$. Since $\{f_k\}_{k=1}^\infty$ is dense in $L^2(\mathbb{R}^n)$, a standard $\varepsilon/4$ argument gives that the sequence of complex numbers

$$\begin{aligned} N(j) &:= \int_{\mathbb{R}^n} T^{(\varepsilon_j)} f(x)g(x)dx \\ &= \int_{\mathbb{R}^n} T^{(\varepsilon_j)}[f(x)-f_k(x)]g(x)dx + \int_{\mathbb{R}^n} T^{(\varepsilon_j)} f_k(x)g(x)dx \end{aligned}$$

is Cauchy, and thus, it converges. Indeed, for sufficiently large $j, l > k$, it holds

$$\begin{aligned} |N(j)-N(l)| \leqslant & \left|\int_{\mathbb{R}^n} T^{(\varepsilon_j)}[f(x)-f_k(x)]g(x)dx\right| \\ & + \left|\int_{\mathbb{R}^n} T^{(\varepsilon_l)}[f(x)-f_k(x)]g(x)dx\right| \\ & + \left|\int_{\mathbb{R}^n} T^{(\varepsilon_j)} f_k(x)g(x)dx - \int_{\mathbb{R}^n} T_0^{f_k}(x)g(x)dx\right| \\ & + \left|\int_{\mathbb{R}^n} T_0^{f_k}(x)g(x)dx - \int_{\mathbb{R}^n} T^{(\varepsilon_l)} f_k(x)g(x)dx\right| \\ \leqslant & 2B'\|f-f_k\|_2\|g\|_2 + \varepsilon/4 + \varepsilon/4 \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Now L^2 is complete in the weak* topology since the unit ball of L^2 in the weak* topology is compact and metrizable; therefore, for each $f \in L^2(\mathbb{R}^n)$, there is a function $T_0 f$ such that (8.24) holds for all $f, g \in L^2(\mathbb{R}^n)$ as $j \to \infty$. It is easy to see that T_0 is a linear operator with the property $T_0 f_k = T_0^{f_k}$ for each $k = 1, 2, \ldots$. This proves (i).

(ii) Let $j \to \infty$ in the integral representation

$$\int_{\mathbb{R}^n} T^{(\varepsilon_j)} f(x)g(x)dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K^{(\varepsilon_j)}(x,y)f(y)dyg(x)dx,$$

the l.h.s. tends to $\int_{\mathbb{R}^n} T_0 f(x)g(x)dx$ and the r.h.s. tends to $\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f(y)dyg(x)dx$ whenever f, g are Schwartz functions with disjoint supports. This gives the result.

(iii) From (8.24) and (8.23), it follows that

$$\|T_0 f\|_2 \leqslant \sup_{\|g\|_2\leqslant 1} \limsup_{j\to\infty} \left|\int_{\mathbb{R}^n} T^{(\varepsilon_j)} f(x)g(x)dx\right| \leqslant B'\|f\|_2.$$

(iv) We first observe that if g is a bounded function with compact support and Q is an open cube in $\mathbb{R}^n$, we have

$$(T^{(\varepsilon_j)} - T)(g\chi_Q)(x) = \chi_Q(x)(T^{(\varepsilon_j)} - T)g(x), \tag{8.27}$$

for a.e. $x \notin \partial Q$ whenever ε_j is small enough (depending on x). Indeed, since $g\chi_Q$ is bounded and has compact support, by the integral representation formula (8.19), there is a null set $E(g\chi_Q)$ such that for

$x \notin \overline{Q} \cup E(g\chi_Q)$ and for $\varepsilon_j <$ dist $(x, \operatorname{supp} g\chi_Q)$, the l.h.s. in (8.27) , i.e.,

$$(T^{(\varepsilon_j)} - T)(g\chi_Q)(x) = -\int_{|x-y|\leqslant\varepsilon_j} K(x,y)(g\chi_Q)(y)dy,$$

is zero due to $y \notin \operatorname{supp} g\chi_Q$ for $|x-y| \leqslant \varepsilon_j$, which implies (8.27) for a.e. $x \in \overline{Q}^c$. Moreover, since $g\chi_{Q^c}$ is also bounded and compactly supported, there is a null set $E(g\chi_{Q^c})$ such that for $x \notin Q^c \cup E(g\chi_{Q^c})$ and $\varepsilon_j <$ dist $(x, \operatorname{supp} g\chi_{Q^c})$, we have $y \notin \operatorname{supp} g\chi_{Q^c}$, and thus $(T^{(\varepsilon_j)} - T)(g\chi_{Q^c})(x) = 0$ which implies for a.e. $x \in Q$

$$(T^{(\varepsilon_j)} - T)(g\chi_Q)(x) = (T^{(\varepsilon_j)} - T)g(x).$$

Hence, (8.27) holds for a.e. $x \notin \partial Q$.

Taking weak limits in (8.27) as $\varepsilon_j \to 0$, we obtain that

$$(T_0 - T)(g\chi_Q) = \chi_Q(T_0 - T)g, \quad \text{a.e.,} \tag{8.28}$$

for all open cubes Q in $\mathbb{R}^n$. This means that for any g bounded function with compact support and open cube Q in $\mathbb{R}^n$, there is a set of measure zero $E_{Q,g}$ such that (8.28) holds on $\mathbb{R}^n \setminus E_{Q,g}$. Consider the countable family $\mathscr{F}$ of all cubes in $\mathbb{R}^n$ with corners in $\mathbb{Q}^n$ and set $E_g = \bigcup_{Q\in\mathscr{F}} E_{Q,g}$. Then $|E_g| = 0$ and by linearity we obtain

$$(T_0 - T)(gh) = h(T_0 - T)g, \quad \text{on } \mathbb{R}^n \setminus E_g$$

whenever h is a finite linear combination of the characteristic functions of cubes in $\mathscr{F}$, which is a dense subspace of L^2. Via a simple density argument, using the fact that $T_0 - T$ is L^2 bounded, we obtain that for all $f \in L^2$ and g bounded with compact support, there is a null set $E_{f,g}$ such that

$$(T_0 - T)(gf) = f(T_0 - T)g, \quad \text{on } \mathbb{R}^n \setminus E_{f,g}. \tag{8.29}$$

Now, if $B(0,j)$ is the open ball with center 0 and radius j, when $j \leqslant j'$, we have

$$(T_0 - T)\chi_{B(0,j)} = (T_0 - T)(\chi_{B(0,j)}\chi_{B(0,j')}) = \chi_{B(0,j)}(T_0 - T)(\chi_{B(0,j')}), \quad \text{a.e.}$$

Therefore, the functions $(T_0 - T)\chi_{B(0,j)}$ satisfy the "consistency" property

$$(T_0 - T)\chi_{B(0,j)} = (T_0 - T)\chi_{B(0,j')}, \quad \text{a.e. on } B(0,j)$$

when $j \leqslant j'$. It follows that there exists a well-defined measurable function b such that

$$b\chi_{B(0,j)} = (T_0 - T)\chi_{B(0,j)}, \quad \text{a.e.}$$

Applying (8.29) with $f \in L^2$ and $g = \chi_{B(0,j)}$, we obtain

$$(T_0 - T)(f\chi_{B(0,j)}) = f(T_0 - T)\chi_{B(0,j)} = fb, \quad \text{a.e. on } B(0,j). \tag{8.30}$$

Since the norm of $T - T_0$ on L^2 is at most $B + B'$, we obtain from (8.30) that

$$B+B' \geqslant \sup_{j\geqslant 1} \sup_{\substack{0\neq f\in L^2 \\ \operatorname{supp} f\subset B(0,j)}} \frac{\|(T_0-T)(f\chi_{B(0,j)})\|_2}{\|f\|_2}$$

$$= \sup_{\substack{0\neq f\in L^2 \\ \operatorname{supp} f \text{compact}}} \frac{\|fb\|_2}{\|f\|_2} = \|b\|_\infty.$$

The fact that $b\in L^\infty$ together with (8.30) easily yields

$$(T_0-T)f = bf \quad \text{a.e.}$$

for all $f\in L^2$. This identifies T_0-T and concludes the proof of (iv). ❑

We give a special name to operators of this form.

Definition 8.13. Suppose that for a given $T\in \mathrm{CZO}(\delta,A,B)$, there is a sequence $\varepsilon_j \searrow 0$ as $j\to\infty$ such that for all $f\in L^2(\mathbb{R}^n)$,

$$T^{(\varepsilon_j)}f \to Tf$$

weakly in L^2. Then, T is called a ***Calderón-Zygmund singular integral operator***. Thus, Calderón-Zygmund singular integral operators are special kinds of Calderón-Zygmund operators. The subclass of $\mathrm{CZO}(\delta,A,B)$, consisting of all Calderón-Zygmund singular integral operators, is denoted by $\mathrm{CZSIO}(\delta,A,B)$.

§ 8.1.3 Calderón-Zygmund operators acting on bounded functions

We are now interested in defining the action of a Calderón-Zygmund operator T on bounded and smooth functions. To achieve this, we first need to define the space of special test functions $\mathscr{D}_0$.

Definition 8.14. We define $\mathscr{D}_0(\mathbb{R}^n)$ to be the space of all smooth functions with compact support and integral zero, i.e.,

$$\mathscr{D}_0(\mathbb{R}^n) = \left\{\phi\in\mathscr{D}(\mathbb{R}^n) : \int_{\mathbb{R}^n}\phi(x)dx = 0\right\}.$$

We equip $\mathscr{D}_0(\mathbb{R}^n)$ with the same topology as the space $\mathscr{D}(\mathbb{R}^n)$. This means that a linear functional $u\in\mathscr{D}'(\mathbb{R}^n)$ is continuous if for any compact set $K\subset\mathbb{R}^n$ there is a constant C_K and an integer M such that

$$|\langle u,\phi\rangle| \leqslant C_K \sum_{|\alpha|\leqslant M} \|\partial^\alpha\phi\|_\infty$$

for all ϕ smooth functions supported in K. The dual space of $\mathscr{D}_0(\mathbb{R}^n)$ under this topology is denoted by $\mathscr{D}_0'(\mathbb{R}^n)$. This is a space of distributions larger than $\mathscr{D}'(\mathbb{R}^n)$.

Example 8.15. $\mathrm{BMO}(\mathbb{R}^n) \subset \mathscr{D}_0'(\mathbb{R}^n)$. Indeed, given $b \in \mathrm{BMO}(\mathbb{R}^n)$, for any compact set K, there is the smallest cube $Q \supset K$ and a constant C_K such that

$$\left| \int_{\mathbb{R}^n} b(x)\phi(x)dx \right| = \left| \int_Q (b(x) - \operatorname*{Avg}_Q b)\phi(x)dx \right|$$
$$\leqslant |Q| \|b\|_{\mathrm{BMO}} \|\phi\|_\infty \leqslant C_K \|\phi\|_\infty$$

for any $\phi \in \mathscr{D}_0(K)$. Moreover, observe that the preceding integral remains unchanged if the BMO function b is replaced by $b + c$, where c is a constant, since $\int \phi(x)dx = 0$.

Definition 8.16. Let T be a continuous linear operator from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ that satisfies (8.7) for some distribution W that coincides with a standard kernel $K \in \mathrm{SK}(\delta, A)$. Given f bounded and smooth, we define an element Tf of $\mathscr{D}_0'(\mathbb{R}^n)$ as follows: For a given $\phi \in \mathscr{D}_0(\mathbb{R}^n)$, select $\eta \in \mathscr{D}$ with $0 \leqslant \eta \leqslant 1$ and equal to 1 in a neighborhood of $\operatorname{supp} \phi$. Since T maps $\mathscr{S}$ to $\mathscr{S}'$, the expression $T(f\eta)$ is a tempered distribution, and its action on ϕ is well-defined. We define ***the action of Tf on ϕ*** via the identity

$$\langle Tf, \phi \rangle = \langle T(f\eta), \phi \rangle + \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} K(x,y)\phi(x)dx \right] f(y)(1 - \eta(y))dy, \tag{8.31}$$

provided we make sense of the double integral as an absolutely convergent integral.

To show $Tf \in \mathscr{D}_0'(\mathbb{R}^n)$, we pick $x_0 \in \operatorname{supp} \phi$ and split the y-integral in (8.31) into the sum of integrals over the regions $I_0 = \{y \in \mathbb{R}^n : |x - x_0| > \frac{1}{2}|x_0 - y|\}$ and $I_\infty = \{y \in \mathbb{R}^n : |x - x_0| \leqslant \frac{1}{2}|x_0 - y|\}$.

By the choice of η, we must necessarily have $c = \operatorname{dist}(\operatorname{supp}(1 - \eta), \operatorname{supp} \phi) > 0$, and hence the part of the double integral in (8.31) is absolutely convergent in view of (8.1) when y is restricted to I_0, i.e.,

$$\int_{I_0} \int_{\mathbb{R}^n} |K(x,y)\phi(x)f(y)(1 - \eta(y))| dxdy$$
$$\leqslant \int_{I_0} \int_{\mathbb{R}^n} \frac{A}{|x - y|^n} |\phi(x)||f(y)|(1 - \eta(y))dxdy$$
$$\leqslant \frac{CA}{c^n} \int_{I_0} \int_{\mathbb{R}^n} |\phi(x)| dxdy < \infty,$$

since $\overline{I_0}$ is compact due to $x, x_0 \in \operatorname{supp} \phi$.

For $y \in I_\infty$, we use the zero mean value property of ϕ to write the expression inside the square brackets in (8.31) as

$$\int_{\mathbb{R}^n} (K(x,y) - K(x_0, y))\phi(x)dx.$$

With the aid of (8.2), we have

$$\begin{aligned}
&\int_{I_\infty}\int_{\mathbb{R}^n} |(K(x,y)-K(x_0,y))\phi(x)f(y)(1-\eta(y))|dxdy\\
=&\iint_{|y-x_0|\geqslant 2|x-x_0|} |K(x,y)-K(x_0,y)||\phi(x)||f(y)|(1-\eta(y))|dydx\\
\leqslant&\iint_{|y-x_0|\geqslant 2|x-x_0|} \frac{A|x-x_0|^\delta}{(|x-y|+|x_0-y|)^{n+\delta}}|\phi(x)||f(y)|dydx\\
\leqslant&\|f\|_\infty \int_{\mathbb{R}^n} A|x-x_0|^\delta \int_{|y-x_0|\geqslant 2|x-x_0|} |x_0-y|^{-n-\delta}dy|\phi(x)|dx\\
\leqslant&\|f\|_\infty \int_{\mathbb{R}^n} A|x-x_0|^\delta \omega_{n-1}\int_{2|x-x_0|}^\infty r^{-n-\delta}r^{n-1}dr|\phi(x)|dx\\
\leqslant&\|f\|_\infty \int_{\mathbb{R}^n} A|x-x_0|^\delta \frac{\omega_{n-1}}{\delta}(2|x-x_0|)^{-\delta}|\phi(x)|dx\\
=&A\frac{\omega_{n-1}}{\delta 2^\delta}\|\phi\|_1\|f\|_\infty<\infty.
\end{aligned}$$

Thus, the double integral in (8.31) is absolutely integrable for $f \in L^\infty \cap \mathcal{C}^\infty$. Hence, this yields $Tf \in \mathscr{D}_0'$ when $f \in L^\infty \cap \mathcal{C}^\infty$ and certainly (8.31) is independent of x_0, but leaves two points open. First, we need to show that this definition is independent of η, and second, that whenever f is a Schwartz function, the distribution Tf defined in (8.31) coincides with the original element of $\mathscr{S}'(\mathbb{R}^n)$ given in Definition 8.9.

We first show that the definition of Tf is independent of the choice of the function η. Indeed, if ζ is another function satisfying $0 \leqslant \zeta \leqslant 1$ that is also equal to 1 in a neighborhood of the $\operatorname{supp}\phi$, then $f(\eta-\zeta)$ and ϕ have disjoint supports, and by (8.9), we have the absolutely convergent integral realization

$$\langle T(f(\eta-\zeta)),\phi\rangle = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x,y)f(y)(\eta-\zeta)(y)dy\phi(x)dx.$$

It follows

$$\begin{aligned}
&\langle T(f\zeta),\phi\rangle + \int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} K(x,y)\phi(x)dx\right] f(y)(1-\zeta(y))dy\\
=&\langle T(f\eta),\phi\rangle + \int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n} K(x,y)\phi(x)dx\right] f(y)(1-\eta(y))dy.
\end{aligned}$$

Next, if f is a Schwartz function, then both ηf and $(1-\eta)f$ are Schwartz functions. By the linearity of T, we have

$$\langle Tf,\phi\rangle = \langle T(\eta f),\phi\rangle + \langle T((1-\eta)f),\phi\rangle,$$

and by (8.9), the second expression in (8.31) can be written as the double absolutely convergent integral, since ϕ and $(1-\eta)f$ have disjoint supports. Thus, the distribution Tf defined in (8.31) coincides with the original element of $\mathscr{S}'(\mathbb{R}^n)$ given in Definition 8.9.

Remark 8.17. When T has a bounded extension that maps L^2 to itself, we may define Tf for all $f \in L^\infty(\mathbb{R}^n)$, which is not necessarily smooth. Simply observe that under this assumption, the expression $T(f\eta)$ is a well-defined L^2 function and thus

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

is given by an absolutely convergent integral for all $\phi \in \mathscr{D}_0$.

Finally, observe that although $\langle Tf, \phi \rangle$ is defined for $f \in L^\infty$ and $\phi \in \mathscr{D}_0$, this definition is valid for all square integrable functions ϕ with compact support and integral zero; indeed, the smoothness of ϕ was never an issue in the definition of $\langle Tf, \phi \rangle$. In summary, if T is a Calderón-Zygmund operator and $f \in L^\infty(\mathbb{R}^n)$, then Tf has a well-defined ***action*** $\langle Tf, \phi \rangle$ on square integrable functions ϕ with compact support and integral zero. This action satisfies

$$|\langle Tf, \phi \rangle| \leqslant \|T(f\eta)\|_2 \|\phi\|_2 + C_{n,\delta} A \|\phi\|_1 \|f\|_\infty < \infty. \tag{8.32}$$

In the next section, we show that in this case, Tf is in fact an element of BMO.

§ 8.2 Consequences of L^2 boundedness

Calderón-Zygmund singular integral operators admit L^2-bounded extensions. As in the case of convolution operators, L^2 boundedness has several consequences. In this section, we are concerned with the consequences of the L^2 boundedness of Calderón-Zygmund singular integral operators. Throughout the entire discussion, we assume that $K(x, y)$ is a kernel defined away from the diagonal in $\mathbb{R}^{2n}$ that satisfies the standard size and regularity conditions (8.1), (8.2) and (8.3). These conditions may be relaxed.

§ 8.2.1 Weak type $(1, 1)$ and L^p boundedness of singular integrals

We now prove that operators in $\mathrm{CZO}(\delta, A, B)$ have bounded extensions from L^1 to $L^{1,\infty}$.

Theorem 8.18. *Assume that $K(x, y)$ is in $\mathrm{SK}(\delta, A)$ and let T be an element of $\mathrm{CZO}(\delta, A, B)$ associated with the kernel K. Then T has a bounded extension that maps $L^1(\mathbb{R}^n)$ to $L^{1,\infty}(\mathbb{R}^n)$ with norm*

$$\|T\|_{L^1 \to L^{1,\infty}} \leqslant C_n(A + B),$$

and also maps $L^p(\mathbb{R}^n)$ to itself for $1 < p < \infty$ with norm

$$\|T\|_{L^p \to L^p} \leqslant C_n \max(p, (p-1)^{-1/p})(A + B),$$

where C_n is a dimensional constant.

Proof. Fix $\alpha > 0$ and let $f \in L^1(\mathbb{R}^n)$. Since T may not be defined on general integrable functions, we work with the class $\mathscr{F}_0$ of finite linear combination of characteristic functions of dyadic cubes. The class $\mathscr{F}_0$ is dense in L^1 and also contained in L^2, on which the operator is already defined. Once we obtain a weak type $(1,1)$ estimate for $\mathscr{F}_0$, by density, this extends to the entire L^1.

We apply the Calderón-Zygmund decomposition to $f \in \mathscr{F}_0$ at height $\gamma\alpha$, where γ is a positive constant to be chosen later. Write $f = g + b$, where $b = \sum_j b_j$ and conditions (i)-(iii) of Theorem 2.17 are satisfied with the constant α replaced by $\gamma\alpha$. Due to $f \in \mathscr{F}_0$, the sum $b = \sum_j b_j$ extends over a finite set of indices. Moreover, each bad function b_j is bounded by the boundedness of f and is also compactly supported by construction. Thus, Tb_j is an L^2 function, and for almost all $x \notin \operatorname{supp} b_j$ we have the integral representation

$$Tb_j(x) = \int_{Q_j} K(x,y) b_j(y) dy$$

in view of Proposition 8.10.

We denote by $\ell(Q)$ the side length of a cube Q. Let Q_j^* be the unique cube with sides parallel to the axes having the same center as Q_j and having side length

$$\ell(Q_j^*) = 2\sqrt{n}\ell(Q_j).$$

We have

$$\begin{aligned}
(Tf)_*(\alpha) \leqslant& (Tg)_*(\alpha/2) + (Tb)_*(\alpha/2) \\
=& \left|\left\{x \in \mathbb{R}^n : |Tg(x)| > \frac{\alpha}{2}\right\}\right| + \left|\left\{x \in \mathbb{R}^n : |Tb(x)| > \frac{\alpha}{2}\right\}\right| \\
\leqslant& \left(\frac{2}{\alpha}\right)^2 \|Tg\|_2^2 + |\bigcup_j Q_j^*| + \left|\left\{x \notin \bigcup_j Q_j^* : |Tb(x)| > \frac{\alpha}{2}\right\}\right| \\
\leqslant& \frac{2^2}{\alpha^2} B^2 \|g\|_2^2 + \sum_j |Q_j^*| + \frac{2}{\alpha} \int_{(\bigcup_j Q_j^*)^c} |Tb(x)| dx \\
\leqslant& \frac{2^2}{\alpha^2} B^2 \|f\|_1 2^n \gamma\alpha + (2\sqrt{n})^n \frac{\|f\|_1}{\gamma\alpha} + \frac{2}{\alpha} \sum_j \int_{(\bigcup_j Q_j^*)^c} |Tb_j(x)| dx \\
\leqslant& \left(\frac{(2^{n+1}B\gamma)^2}{2^n\gamma} + \frac{(2\sqrt{n})^n}{\gamma}\right) \frac{\|f\|_1}{\alpha} + \frac{2}{\alpha} \sum_j \int_{(\bigcup_j Q_j^*)^c} |Tb_j(x)| dx.
\end{aligned}$$

It suffices to show that the last sum is bounded by some constant multiple of $\|f\|_1$. Let y_j be the center of the cube Q_j. For $x \in (Q_j^*)^c$, we have

$$|x - y_j| \geqslant \frac{1}{2}\ell(Q_j^*) = \sqrt{n}\ell(Q_j).$$

However, if $y \in Q_j$, we have $|y - y_j| \leqslant \sqrt{n}\ell(Q_j)/2$; thus, $|y - y_j| \leqslant \frac{1}{2}|x - y_j|$ since the diameter of a cube is equal to $\sqrt{n}$ times its side length. We now estimate the last sum as follows.

$$\begin{aligned}
&\sum_j \int_{(\bigcup_j Q_j^*)^c} |Tb_j(x)|dx = \sum_j \int_{(\bigcup_j Q_j^*)^c} \left| \int_{Q_j} b_j(y)K(x,y)dy \right| dx \\
=&\sum_j \int_{(\bigcup_j Q_j^*)^c} \left| \int_{Q_j} b_j(y)(K(x,y) - K(x,y_j))dy \right| dx \\
\leqslant&\sum_j \int_{Q_j} |b_j(y)| \int_{(\bigcup_j Q_j^*)^c} |K(x,y) - K(x,y_j)|dxdy \\
\leqslant&\sum_j \int_{Q_j} |b_j(y)| \int_{|x-y_j|\geqslant 2|y-y_j|} |K(x,y) - K(x,y_j)|dxdy \\
\leqslant&\sum_j \int_{Q_j} |b_j(y)| \int_{|x-y_j|\geqslant 2|y-y_j|} \frac{A|y-y_j|^\delta}{(|x-y|+|x-y_j|)^{n+\delta}}dxdy \\
\leqslant&\sum_j \int_{Q_j} |b_j(y)| A\omega_{n-1}|y-y_j|^\delta \int_{2|y-y_j|}^\infty r^{-n-\delta}r^{n-1}drdy \\
\leqslant&C_nA\sum_j \int_{Q_j} |b_j(y)|dy \\
\leqslant&C_nA\sum_j 2^{n+1}\gamma\alpha|Q_j| \\
\leqslant&C_nA2^{n+1}\|f\|_1.
\end{aligned}$$

Combining these facts and choosing $\gamma = B^{-1}$, we deduce the claimed inequality for $f \in \mathscr{F}_0$. By density, we obtain that T has a bounded extension from L^1 to $L^{1,\infty}$ with bound at most $C_n(A+B)$. The L^p result for $1 < p < 2$ follows from interpolation, while the fact that the constant blows up like $(p-1)^{-1/p}$ as $p \to 1$ can be deduced from the result of Exercise 1.9. The result for $2 < p < \infty$ follows from duality; one uses here that the dual operator T^t has a kernel $K^t(x,y) = K(y,x)$ that satisfies the same estimates as K, and by the result just proved, it is also bounded on L^p for $1 < p < 2$ with norm at most $C_n(A+B)$. Thus T must be bounded on L^p for $2 < p < \infty$ with norm at most a constant multiple of $A+B$. ❑

Consequently, for operators $T \in \mathrm{CZO}(\delta, A, B)$ and $f \in L^p$, $1 \leqslant p < \infty$, the expression Tf makes sense as L^p (or $L^{1,\infty}$ when $p = 1$) functions. The following result addresses whether these functions can be expressed as integrals.

Proposition 8.19. *Let $T \in \mathrm{CZO}(\delta, A, B)$ be an operator associated with a kernel $K \in \mathrm{SK}(\delta, A)$. Then for $g \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, the following absolutely convergent integral representation is valid*

$$Tg(x) = \int_{\mathbb{R}^n} K(x,y)g(y)dy \tag{8.33}$$

for almost all $x \in \mathbb{R}^n \setminus \operatorname{supp} g$, provided that $\operatorname{supp} g \subsetneq \mathbb{R}^n$.

Proof. Set $g_k(x) = g(x)\chi_{|g(x)|\leqslant k}\chi_{|x|\leqslant k}$. These are L^p functions with compact support contained in the support of g. Additionally, the g_k converge to g in L^p as $k \to \infty$. In view of Proposition 8.10, for every k we have

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

for almost all $x \in \mathbb{R}^n \setminus \operatorname{supp} g$. Since T maps L^p to L^p (or to weak L^1 when $p = 1$), it follows that Tg_k converges to Tg in $L^{p,\infty}$ and hence in measure by Proposition 1.20. From Riesz's theorem, a subsequence of Tg_k a.e. converges to Tg. On the other hand, for $x \in \mathbb{R}^n \setminus \operatorname{supp} g$ we have

$$\int_{\mathbb{R}^n} K(x,y)g_k(y)dy \to \int_{\mathbb{R}^n} K(x,y)g(y)dy$$

when $k \to \infty$, since the absolute value of the difference is bounded by $B'\|g_k - g\|_p$, which tends to zero. The constant B' is the $L^{p'}$ norm of the function $|x-y|^{-n}$ on the support of g, and one has $|x-y| \geqslant c > 0$ for all $y \in \operatorname{supp} g$ because $\mathbb{R}^n \setminus \operatorname{supp} g$ is open, and thus $B' < \infty$. Therefore, $Tg_k(x)$ converges a.e. to both sides of the identity (8.33) for $x \notin \operatorname{supp} g$. This concludes the proof of this identity. ❑

§ 8.2.2 Boundedness of maximal singular integrals

We pose the question whether there is a result concerning the maximal singular integral operator $T^{(*)}$ analogous to Theorem 8.18. We note that given $f \in L^p(\mathbb{R}^n)$ for some $1 \leqslant p < \infty$, the expression $T^{(*)}f(x)$ is well-defined for all $x \in \mathbb{R}^n$. This is a simple consequence of condition (8.1) and the Hölder inequality.

Theorem 8.20. *Let K be in $\mathrm{SK}(\delta, A)$ and $T \in \mathrm{CZO}(\delta, A, B)$ be associated with K. Let $r \in (0,1)$. Then, there is a constant $C(n,r)$ such that* **Cotlar's inequality**

$$|T^{(*)}f(x)| \leqslant C(n,r)\left[(M(|Tf|^r)(x))^{1/r} + (A+B)Mf(x)\right] \tag{8.34}$$

is valid for all functions in $\bigcup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$. Additionally, there exist dimen-

sional constants C_n, C'_n such that

$$\|T^{(*)}f\|_{L^{1,\infty}} \leqslant C'_n(A+B)\|f\|_1, \tag{8.35}$$

$$\|T^{(*)}f\|_p \leqslant C_n(A+B)\max(p,(p-1)^{-1/p})\|f\|_p, \tag{8.36}$$

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

Proof. We fix $r \in (0,1)$ and $f \in L^p(\mathbb{R}^n)$ for some $p \in [1,\infty)$. To prove (8.34), we also fix $\varepsilon > 0$ and set $f_0^{\varepsilon,x} = f\chi_{B(x,\varepsilon)}$ and $f_\infty^{\varepsilon,x} = f\chi_{B(x,\varepsilon)^c}$. Since $x \notin \operatorname{supp} f_\infty^{\varepsilon,x}$, using Proposition 8.19, we can write

$$Tf_\infty^{\varepsilon,x}(x) = \int_{\mathbb{R}^n} K(x,y)f_\infty^{\varepsilon,x}(y)dy = \int_{|x-y|\geqslant\varepsilon} K(x,y)f(y)dy = T^{(\varepsilon)}f(x).$$

In view of (8.2), for $z \in B(x,\varepsilon/2)$, we have $|z-x| \leqslant |x-y|/2$ whenever $|x-y| \geqslant \varepsilon$ and thus

$$\begin{aligned}|Tf_\infty^{\varepsilon,x}(x) - Tf_\infty^{\varepsilon,x}(z)| =& \left|\int_{|x-y|\geqslant\varepsilon}(K(z,y)-K(x,y))f(y)dy\right| \\ \leqslant& |x-z|^\delta \int_{|x-y|\geqslant\varepsilon} \frac{A|f(y)|}{(|x-y|+|y-z|)^{n+\delta}}dy \\ \leqslant& \left(\frac{\varepsilon}{2}\right)^\delta \int_{|x-y|\geqslant\varepsilon} \frac{A|f(y)|}{(|x-y|+\varepsilon/2)^{n+\delta}}dy \\ \leqslant& C_{n,\delta}AMf(x),\end{aligned}$$

where the last estimate is a consequence of Theorem 2.10 (with $\varphi = \frac{\chi_{|x|\geqslant 1}}{(|x|+1)^{n+\delta}}$). We conclude that for all $z \in B(x,\varepsilon/2)$, we have

$$\begin{aligned}|T^{(\varepsilon)}f(x)| =& |Tf_\infty^{\varepsilon,x}(x)| \\ \leqslant& |Tf_\infty^{\varepsilon,x}(x) - Tf_\infty^{\varepsilon,x}(z)| + |Tf_\infty^{\varepsilon,x}(z)| \\ \leqslant& C_{n,\delta}AMf(x) + |Tf_0^{\varepsilon,x}(z)| + |Tf(z)|.\end{aligned} \tag{8.37}$$

For $r \in (0,1)$, it follows from (8.37) that for $z \in B(x,\varepsilon/2)$, we have

$$|T^{(\varepsilon)}f(x)|^r \leqslant C_{n,\delta}^r A^r (Mf(x))^r + |Tf_0^{\varepsilon,x}(z)|^r + |Tf(z)|^r. \tag{8.38}$$

Integrating over $z \in B(x,\varepsilon/2)$, dividing by $|B(x,\varepsilon/2)|$, and raising to the power $1/r$, we obtain

$$\begin{aligned}|T^{(\varepsilon)}f(x)| \leqslant 3^{1/r}\Big[C_{n,\delta}A(Mf(x)) + \left(\frac{1}{|B(x,\varepsilon/2)|}\int_{B(x,\varepsilon/2)}|Tf_0^{\varepsilon,x}(z)|^r dz\right)^{1/r} \\ + (M(|Tf|^r)(x))^{1/r}\Big].\end{aligned}$$

For the middle term, by Theorems 1.17 and 8.18, we have

$$\begin{aligned}&\int_{B(x,\varepsilon/2)} |Tf_0^{\varepsilon,x}(z)|^r dz \\ =& \int_0^\infty r\alpha^{r-1}\left|\left\{z \in B(x,\varepsilon/2) : |Tf_0^{\varepsilon,x}(z)| > \alpha\right\}\right| d\alpha\end{aligned}$$

$$
\begin{aligned}
&\leqslant \int_0^\infty r\alpha^{r-1}\min\left(|B(x,\varepsilon/2)|, C\|f_0^{\varepsilon,x}\|_1/\alpha\right)d\alpha\\
&= \int_0^{\frac{C\|f_0^{\varepsilon,x}\|_1}{|B(x,\varepsilon/2)|}} r\alpha^{r-1}|B(x,\varepsilon/2)|d\alpha + \int_{\frac{C\|f_0^{\varepsilon,x}\|_1}{|B(x,\varepsilon/2)|}}^\infty r\alpha^{r-1}C\|f_0^{\varepsilon,x}\|_1/\alpha d\alpha\\
&=|B(x,\varepsilon/2)|\left(\frac{C\|f_0^{\varepsilon,x}\|_1}{|B(x,\varepsilon/2)|}\right)^r + C\|f_0^{\varepsilon,x}\|_1\frac{r}{1-r}\left(\frac{C\|f_0^{\varepsilon,x}\|_1}{|B(x,\varepsilon/2)|}\right)^{r-1}\\
&=(C\|f_0^{\varepsilon,x}\|_1)^r|B(x,\varepsilon/2)|^{1-r} + \frac{r}{1-r}(C\|f_0^{\varepsilon,x}\|_1)^r|B(x,\varepsilon/2)|^{1-r}\\
&=\frac{1}{1-r}(C\|f_0^{\varepsilon,x}\|_1)^r|B(x,\varepsilon/2)|^{1-r},
\end{aligned}
$$

where $C = C_n(A+B)$, and thus

$$
\begin{aligned}
&\left(\frac{1}{|B(x,\varepsilon/2)|}\int_{B(x,\varepsilon/2)}|Tf_0^{\varepsilon,x}(z)|^r dz\right)^{1/r}\\
\leqslant&\left(\frac{1}{1-r}(C\|f_0^{\varepsilon,x}\|_1)^r|B(x,\varepsilon/2)|^{-r}\right)^{1/r}\\
=&C_{n,r}(A+B)\frac{1}{|B(x,\varepsilon/2)|}\int_{B(x,\varepsilon)}|f(y)|dy\\
\leqslant&C'_{n,r}(A+B)Mf(x).
\end{aligned}
$$

This proves (8.34).

We now use (8.34) to show that $T^{(*)}$ is L^p bounded and of weak type $(1,1)$. To obtain the weak type $(1,1)$ estimate for $T^{(*)}$, we need to use that the Hardy-Littlewood maximal operator maps $L^{p,\infty}$ to $L^{p,\infty}$ for all $p\in(1,\infty)$ (see Exercise 2.2). We also use the trivial fact that for all $0<p,q<\infty$, we have

$$\||f|^q\|_{L^{p,\infty}} = \|f\|^q_{L^{pq,\infty}}.$$

Take any $r<1$ in (8.34). Then, we have

$$
\begin{aligned}
\|(M(|Tf|^r))^{1/r}\|_{L^{1,\infty}} =&\|M(|Tf|^r)\|^{1/r}_{L^{1/r,\infty}}\\
\leqslant&C_{n,r}\||Tf|^r\|^{1/r}_{L^{1/r,\infty}} = C_{n,r}\|Tf\|_{L^{1,\infty}}\\
\leqslant&C'_{n,r}(A+B)\|f\|_1,
\end{aligned}
$$

where we used the weak type $(1,1)$ bound for T in the last estimate.

To obtain the L^p boundedness of $T^{(*)}$ for $p\in(1,\infty)$, we use the same argument as before. We fix $r=1/2$. Recall that the maximal function is bounded on L^{2p} with norm at most $2(3^n/(2p-1))^{1/2p}\leqslant C_n$ by Theorem 2.6. We have

$$
\begin{aligned}
\|(M(|Tf|^{1/2}))^2\|_p =&\|M(|Tf|^{1/2})\|^2_{2p}\leqslant C_n\||Tf|^{1/2}\|^2_{2p}\leqslant C_n\|Tf\|_p\\
\leqslant&C_n\max(1/(p-1)^{1/p},p)(A+B)\|f\|_p,
\end{aligned}
$$

where we use the L^p boundedness of T in the last estimate. ❑

From the above proof, we have the following corollary by noticing

$B = \|T\|_{L^2\to L^2}$.

Corollary 8.21. *Let K be in $\mathrm{SK}(\delta, A)$ and $T \in \mathrm{CZO}(\delta, A, B)$ be associated with K. Then there exists a dimensional constant C_n such that*

$$\sup_{\varepsilon>0} \|T^{(\varepsilon)}\|_{L^2\to L^2} \leqslant C_n \left(A + \|T\|_{L^2\to L^2}\right).$$

§ 8.2.3 $\mathcal{H}^1 \to L^1$ and $L^\infty \to \mathrm{BMO}$ boundedness of singular integrals

We discuss some endpoint results concerning operators in $\mathrm{CZO}(\delta, A, B)$.

Theorem 8.22. *Let $T \in \mathrm{CZO}(\delta, A, B)$. Then, T has an extension that maps $\mathcal{H}^1(\mathbb{R}^n)$ to $L^1(\mathbb{R}^n)$. Precisely, there is a constant $C_{n,\delta}$ such that*

$$\|T\|_{\mathcal{H}^1\to L^1} \leqslant C_{n,\delta}(A + \|T\|_{L^2\to L^2}). \tag{8.39}$$

Proof. Recall $B = \|T\|_{L^2\to L^2}$. We start by examining the action of T on L^2 atoms for $\mathcal{H}^1$. Let $f = a$ be such an atom supported in a cube Q. Let c_Q be the center of Q and let $Q^* = 2\sqrt{n}Q$. We write

$$\int_{\mathbb{R}^n} |Ta(x)|dx = \int_{Q^*} |Ta(x)|dx + \int_{(Q^*)^c} |Ta(x)|dx \tag{8.40}$$

and we estimate each term separately. By the Cauchy-Schwarz inequality, $T : L^2 \to L^2$, and property (ii) of atoms in Definition 7.1, we have

$$\begin{aligned}\int_{Q^*} |Ta(x)|dx \leqslant & |Q^*|^{1/2} \left(\int_{Q^*} |Ta(x)|^2 dx\right)^{1/2} \\ \leqslant & B|Q^*|^{1/2} \left(\int_{Q^*} |a(x)|^2 dx\right)^{1/2} \\ \leqslant & B|Q^*|^{1/2}|Q|^{-1/2} \\ = & C_n B.\end{aligned}$$

Now, observe that if $x \notin Q^*$ and $y \in Q$, then

$$|y - c_Q| \leqslant \frac{1}{2}|x - c_Q|;$$

hence, $x - y$ stays away from zero due to $|x - y| \geqslant \frac{1}{2}(\ell(Q^*) - \ell(Q)) = \frac{1}{2}(2\sqrt{n} - 1)\ell(Q)$, and $Ta(x)$ can be expressed as a convergent integral by Proposition 8.19

$$Ta(x) = \int_{\mathbb{R}^n} K(x, y)a(y)dy.$$

We have, due to $\operatorname{supp} a \subset Q$, $\int_Q a(y)dy = 0$, Fubini's theorem, (8.3), Remark 7.2, that

$$\begin{aligned}\int_{(Q^*)^c}|Ta(x)|dx &= \int_{(Q^*)^c}\left|\int_{\mathbb{R}^n}K(x,y)a(y)dy\right|dx\\ &= \int_{(Q^*)^c}\left|\int_{\mathbb{R}^n}(K(x,y)-K(x,c_Q))a(y)dy\right|dx\\ &\leqslant \int_Q\int_{(Q^*)^c}|K(x,y)-K(x,c_Q)|dx|a(y)|dy\\ &\leqslant \int_Q\int_{(Q^*)^c}\frac{A|y-C_Q|^\delta}{(|x-y|+|x-c_Q|)^{n+\delta}}dx|a(y)|dy\\ &\leqslant \int_Q\int_{|x-c_Q|\geqslant 2|y-c_Q|}\frac{A|y-C_Q|^\delta}{|x-c_Q|^{n+\delta}}dx|a(y)|dy\\ &= C'_{n,\delta}A\int_Q|a(y)|dy\\ &\leqslant C'_{n,\delta}A.\end{aligned}$$

Thus, we obtain for L^2 atoms of $\mathcal{H}^1$

$$\|Ta\|_1 \leqslant C_{n,\delta}(A+B). \tag{8.41}$$

To pass to general functions in $\mathcal{H}^1$, we use Definition 7.3 to write an $f \in \mathcal{H}^1$ as

$$f = \sum_{j=1}^{\infty}\lambda_j a_j,$$

where the series converges in $\mathcal{H}^1$, a_j are L^2 atoms for $\mathcal{H}^1$, and

$$\|f\|_{\mathcal{H}^1} = \inf\sum_{j=1}^{\infty}|\lambda_j|. \tag{8.42}$$

Since T maps L^1 to $L^{1,\infty}$, by Theorem 8.18, Tf is already a well-defined $L^{1,\infty}$ function. We plan to prove that

$$Tf = \sum_{j=1}^{\infty}\lambda_j Ta_j \quad \text{a.e.}, \tag{8.43}$$

where the series converges in L^1 and defines an a.e. integrable function. Once (8.43) is established, the required conclusion (8.39) follows easily by taking L^1 norms in (8.43) and using (8.41) and (8.42).

To prove (8.43), we use that T is of weak type $(1,1)$. For a given $\mu > 0$, we have, by Proposition 7.5 and (8.41),

$$\begin{aligned}&\left|\left\{\left|Tf-\sum_{j=1}^{\infty}\lambda_j Ta_j\right|>\mu\right\}\right|\\ \leqslant&\left|\left\{\left|Tf-\sum_{j=1}^{N}\lambda_j Ta_j\right|>\mu/2\right\}\right|+\left|\left\{\left|\sum_{j=N+1}^{\infty}\lambda_j Ta_j\right|>\mu/2\right\}\right|\end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{2}{\mu}\|T\|_{L^1\to L^{1,\infty}} \left\| f - \sum_{j=1}^{N} \lambda_j a_j \right\|_1 + \frac{2}{\mu}\left\| \sum_{j=N+1}^{\infty} \lambda_j T a_j \right\|_1 \\
&\leqslant \frac{2}{\mu}\|T\|_{L^1\to L^{1,\infty}} \left\| f - \sum_{j=1}^{N} \lambda_j a_j \right\|_{\mathcal{H}^1} + \frac{2}{\mu} C_{n,\delta}(A+B) \sum_{j=N+1}^{\infty} |\lambda_j|.
\end{aligned}
$$

Since $\sum_{j=1}^{N} \lambda_j a_j$ converges to f in $\mathcal{H}^1$ and $\sum_{j=1}^{\infty} |\lambda_j| < \infty$, both terms in the sum converge to zero as $N \to \infty$. We conclude that

$$
\left| \left\{ \left| Tf - \sum_{j=1}^{\infty} \lambda_j T a_j \right| > \mu \right\} \right| = 0
$$

for all $\mu > 0$, which implies (8.43). ❑

Theorem 8.23. *Let $T \in \mathrm{CZO}(\delta, A, B)$. Then for any bounded function f, the distribution Tf can be identified with a* BMO *function that satisfies*

$$
\|Tf\|_{\mathrm{BMO}} \leqslant C_{n,\delta}(A+B)\|f\|_{\infty}, \tag{8.44}
$$

where $C_{n,\delta}$ is a constant.

Proof. Let $L^2_{0,c}$ be the space of all square integrable functions with compact support and integral zero on $\mathbb{R}^n$. This space is contained in $\mathcal{H}^1(\mathbb{R}^n)$ (cf. Exercise 7.2) and contains the set of finite sums of L^2 atoms for $\mathcal{H}^1$, which is dense in $\mathcal{H}^1$; thus, $L^2_{0,c}$ is dense in $\mathcal{H}^1$. Recall that for $f \in L^\infty$, Tf has a well-defined action $\langle Tf, \varphi\rangle$ on functions $\varphi \in L^2_{0,c}$ and (8.32) holds, i.e., for $\eta \in \mathscr{D}$

$$
|\langle Tf, \varphi\rangle| \leqslant \|T(f\eta)\|_2 \|\varphi\|_2 + C_{n,\delta} A \|\varphi\|_1 \|f\|_\infty < \infty. \tag{8.45}
$$

Suppose we have proved the identity

$$
\langle Tf, \varphi\rangle = \int_{\mathbb{R}^n} T^t\varphi(x) f(x)\,dx, \tag{8.46}
$$

for all bounded functions f and all $\varphi \in L^2_{0,c}$. Since such a φ is in $\mathcal{H}^1$, Theorem 8.22 yields that $T^t\varphi \in L^1$. Consequently, the integral in (8.46) converges absolutely. Assuming (8.46) and using Theorem 8.22 we obtain that

$$
|\langle Tf, \varphi\rangle| \leqslant \|T^t\varphi\|_1 \|f\|_\infty \leqslant C_{n,\delta}(A+B)\|\varphi\|_{\mathcal{H}^1}\|f\|_\infty.
$$

We conclude that $L(\varphi) = \langle Tf, \varphi\rangle$ is a bounded linear functional on $L^2_{0,c} \subset \mathcal{H}^1$ with norm at most $C_{n,\delta}(A+B)\|f\|_\infty$. By Theorem 7.27, there exists a BMO function b_f that satisfies

$$
\begin{aligned}
\|b_f\|_{\mathrm{BMO}} &\leqslant C_n' \|L\|_{\mathcal{H}^1\to\mathbb{C}} = C_n' \sup \frac{|L(\varphi)|}{\|\varphi\|_{\mathcal{H}^1}} = C_n' \sup \frac{|\langle Tf, \varphi\rangle|}{\|\varphi\|_{\mathcal{H}^1}} \\
&\leqslant C_n' C_{n,\delta}(A+B)\|f\|_\infty
\end{aligned}
$$

such that the linear functional L has the form L_{b_f}. In other words, the distribution Tf can be identified with a BMO function that satisfies (8.44) with the constant $C_n' C_{n,\delta}$, i.e.,

$$\|Tf\|_{\mathrm{BMO}} = \|b_f\|_{\mathrm{BMO}} \leqslant C_n' C_{n,\delta}(A+B)\|f\|_\infty.$$

We return to the proof of identity (8.46). Pick a smooth function with compact support η that satisfies $\eta \in [0,1]$ and is equal to 1 in a neighborhood of the support of φ. We write the r.h.s. of (8.46) as

$$\int_{\mathbb{R}^n} T^t(\varphi)\eta f dx + \int_{\mathbb{R}^n} T^t(\varphi)(1-\eta) f dx = \langle T(\eta f), \varphi\rangle + \int_{\mathbb{R}^n} T^t(\varphi)(1-\eta) f dx.$$

In view of Definition 8.16, to prove (8.46) it will suffice to show that

$$\int_{\mathbb{R}^n} T^t(\varphi)(1-\eta) f dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (K(x,y)-K(x_0,y))\varphi(x)dx(1-\eta(y))f(y)dy,$$

where $x_0 \in \operatorname{supp}\varphi$. In the outer integral above, we have $y \notin \operatorname{supp}\varphi$ and the inner integral above is absolutely convergent and equal to

$$\int_{\mathbb{R}^n} (K(x,y)-K(x_0,y))\varphi(x)dx = \int_{\mathbb{R}^n} K^t(y,x)\varphi(x)dx = T^t\varphi(y),$$

by Proposition 8.10, since $y \notin \operatorname{supp}\varphi$. Thus, (8.46) is valid. ❑

§8.3 The $T(1)$ theorem

We now turn to one of the main results of this chapter, the so-called $T(1)$ theorem. This theorem gives necessary and sufficient conditions for linear operators T with standard kernels to be bounded on $L^2(\mathbb{R}^n)$. In this section we obtain several such equivalent conditions. The name of theorem $T(1)$ is due to the fact that one of the equivalent ways to characterize boundedness is expressed in terms of properties of the distribution $T(1)$, which was introduced in Definition 8.16.

§8.3.1 Preliminaries and statement of the theorem

We begin with some preliminary facts and definitions.

Definition 8.24. A ***normalized bump*** is a smooth function φ supported in the ball $B(0,10)$ that satisfies

$$|(\partial_x^\alpha \varphi)(x)| \leqslant 1$$

for all multi-indices $|\alpha| \leqslant 2\left[\frac{n}{2}\right]+2$, where $[x]$ denotes the integer part of x.

Observe that every smooth function supported inside the ball $B(0,10)$ is a constant multiple of a normalized bump. Additionally, note that if a normalized bump is supported in a compact subset of $B(0,10)$, then small translations of it are also normalized bumps.

Given a function f on $\mathbb{R}^n$, $R>0$, and $x_0\in\mathbb{R}^n$, we use the notation f_R to denote the function $f_R(x)=R^{-n}f(R^{-1}x)$ and $\tau^{x_0}f$ to denote the function $\tau^{x_0}f(x)=f(x-x_0)$. Thus,

$$\tau^{x_0}f_R(y)=f_R(y-x_0)=R^{-n}f(R^{-1}(y-x_0)).$$

Set $N=\left[\frac{n}{2}\right]+1$. Using that all derivatives up to order $2N$ of normalized bumps are bounded by 1, we easily deduce that for all $x_0\in\mathbb{R}^n$, all $R>0$, and all normalized bumps φ we have the estimate

$$\begin{aligned}
&R^n\int_{\mathbb{R}^n}\left|\widehat{\tau^{x_0}\varphi_R}(\xi)\right|d\xi\\
&=R^n\int_{\mathbb{R}^n}\left|e^{-ix_0\cdot\xi}\widehat{\varphi_R}(\xi)\right|d\xi\\
&=\int_{\mathbb{R}^n}|(\widehat{\varphi})_{R^{-1}}(\xi)|\,d\xi\\
&=\int_{\mathbb{R}^n}|\widehat{\varphi}(\xi)|\,d\xi\\
&=\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}e^{-iy\cdot\xi}\varphi(y)\,đy\right|d\xi\\
&=\int_{\mathbb{R}^n}\left|\int_{B(0,10)}e^{-iy\cdot\xi}(I-\Delta)^N\varphi(y)\,đy\right|\frac{d\xi}{(1+|\xi|^2)^N}\\
&\leqslant C_n,
\end{aligned}\tag{8.47}$$

since $|(\partial_x^\alpha\varphi)(x)|\leqslant 1$ for all multi-indices α with $|\alpha|\leqslant 2N$, and C_n is independent of the bump φ. Here $I-\Delta$ denotes the operator

$$(I-\Delta)\varphi=\varphi-\sum_{j=1}^n\frac{\partial^2\varphi}{\partial x_j^2}.$$

Definition 8.25. We say that a continuous linear operator

$$T:\mathscr{S}(\mathbb{R}^n)\to\mathscr{S}'(\mathbb{R}^n)$$

satisfies the ***weak boundedness property*** (WBP) if there is a constant C such that for all normalized bumps f and g and for all $x_0\in\mathbb{R}^n$ and $R>0$ we have

$$|\langle T(\tau^{x_0}f_R),\tau^{x_0}g_R\rangle|\leqslant CR^{-n}.\tag{8.48}$$

The smallest constant C in (8.48) is denoted by $\|T\|_{WB}$.

Note that

$$\|\tau^{x_0}f_R\|_2=\|f_R\|_2=\left(\int_{\mathbb{R}^n}|R^{-n}f(x/R)|^2R^nd(x/R)\right)^{1/2}=\|f\|_2R^{-n/2}$$

and thus if T has a bounded extension from $L^2(\mathbb{R}^n)$ to itself, then T satisfies the weak boundedness property with bound

$$\begin{aligned}\|T\|_{WB} =& \sup R^n |\langle T(\tau^{x_0} f_R), \tau^{x_0} g_R\rangle| \\ \leqslant& \|T\|_{L^2\to L^2} \sup R^n \|\tau^{x_0} f_R\|_2 \|\tau^{x_0} g_R\|_2 \\ =& \|T\|_{L^2\to L^2} \sup \|f\|_2\|g\|_2 \\ \leqslant& \|T\|_{L^2\to L^2} \int_{B(0,10)} dx \\ =& 10^n V_n \|T\|_{L^2\to L^2},\end{aligned}$$

where the supremum is taken over all normalized bumps f and g.

We now state one of the main theorems in this chapter.

Theorem 8.26. *Let T be a continuous linear operator from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ whose Schwartz kernel coincides with a standard kernel $K \in \mathrm{SK}(\delta, A)$ for some $0 < A < \infty$ and $0 < \delta \leqslant 1$. Let $K^{(\varepsilon)}$ and $T^{(\varepsilon)}$ be the usual truncated kernel and operator for $\varepsilon > 0$. Assume that there exists a sequence $\varepsilon_j \searrow 0$ such that for all $\varphi, \psi \in \mathscr{S}(\mathbb{R}^n)$, we have*

$$\langle T^{(\varepsilon_j)} \varphi, \psi\rangle \to \langle T\varphi, \psi\rangle. \tag{8.49}$$

Consider the assertions:

(i) *The following statement is valid:*

$$B_1 = \sup_B \sup_{\varepsilon>0} \left[\frac{\|T^{(\varepsilon)}\chi_B\|_2}{|B|^{1/2}} + \frac{\|(T^{(\varepsilon)})^t\chi_B\|_2}{|B|^{1/2}}\right] < \infty,$$

where the first supremum is taken over all balls B in $\mathbb{R}^n$.

(ii) *We have that*

$$\begin{aligned}B_2 = \sup_{\varepsilon,N,x_0} \Bigg[&\frac{1}{N^n}\int_{B(x_0,N)} \left|\int_{|x-y|<N} K^{(\varepsilon)}(x,y)dy\right|^2 dx \\ &+ \frac{1}{N^n}\int_{B(x_0,N)} \left|\int_{|x-y|<N} K^{(\varepsilon)}(y,x)dy\right|^2 dx\Bigg]^{1/2} < \infty,\end{aligned}$$

where the supremum is taken over all $0 < \varepsilon, N < \infty$ with $\varepsilon < N$, and all $x_0 \in \mathbb{R}^n$.

(iii) *The following statement is valid:*

$$B_3 = \sup_{\varphi} \sup_{x_0\in\mathbb{R}^n} \sup_{R>0} R^{n/2} \left[\|T(\tau^{x_0}\varphi_R)\|_2 + \|T^t(\tau^{x_0}\varphi_R)\|_2\right] < \infty,$$

where the first supremum is taken over all normalized bumps φ.

(iv) *The operator T satisfies the weak boundedness property and the distributions $T(1)$ and $T^t(1)$ coincide with BMO functions, that is,*

$$B_4 = \|T(1)\|_{\mathrm{BMO}} + \|T^t(1)\|_{\mathrm{BMO}} + \|T\|_{WB} < \infty.$$

(v) *For every $\xi \in \mathbb{R}^n$ the distributions $T(e^{i(\cdot)\cdot\xi})$ and $T^t(e^{i(\cdot)\cdot\xi})$ coincide* *with* BMO *functions such that*

$$B_5 = \sup_{\xi\in\mathbb{R}^n} \|T(e^{i(\cdot)\cdot\xi})\|_{\mathrm{BMO}} + \sup_{\xi\in\mathbb{R}^n} \|T^t(e^{i(\cdot)\cdot\xi})\|_{\mathrm{BMO}} < \infty.$$

(vi) *The following statement is valid:*

$$B_6 = \sup_{\varphi} \sup_{x_0\in\mathbb{R}^n} \sup_{R>0} R^n \left[\|T(\tau^{x_0}\varphi_R)\|_{\mathrm{BMO}} + \|T^t(\tau^{x_0}\varphi_R)\|_{\mathrm{BMO}}\right] < \infty,$$

where the first supremum is taken over all normalized bumps φ.

Then assertions (i)-(vi) *are all equivalent to each other and to the L^2 boundedness of T, and we have the following equivalence of the previous quantities:*

$$c_{n,\delta}(A + B_j) \leqslant \|T\|_{L^2\to L^2} \leqslant C_{n,\delta}(A + B_j),$$

for all $j \in \{1, 2, 3, 4, 5, 6\}$, for some constants $c_{n,\delta}$, $C_{n,\delta}$ that depend only on the dimension n and on the parameter $\delta > 0$.

Remark 8.27. Condition (8.49) says that the operator T is the weak limit of a sequence of its truncations. We already know from Proposition 8.12 that if T is bounded on L^2, then it must be equal to an operator that satisfies (8.49) plus a bounded function times the identity operator. Therefore, it is not a serious restriction to assume condition (8.49).

One should always keep in mind the following pathological situation: consider the distribution $W_0 \in \mathscr{S}'(\mathbb{R}^n \times \mathbb{R}^n)$ defined for F in $\mathscr{S}(\mathbb{R}^{2n})$ by

$$\langle W_0, F\rangle = \int_{\mathbb{R}^n} F(t,t)h(t)dt,$$

where $h(t) = |t|^2$. In this case, for all $\varepsilon > 0$, $F^{(\varepsilon)}(t,t) = F(t,t)\chi_{|t-t|>\varepsilon} = 0$ then $T^{(\varepsilon)} = 0$; hence, $T^{(\varepsilon)}$ is uniformly bounded on L^2, but $\langle Tf, \varphi\rangle = \int_{\mathbb{R}^n} \varphi(t)f(t)h(t)dt$; thus, Tf can be identified with fh for all $f \in \mathscr{S}$, which is certainly an unbounded operator on $L^2(\mathbb{R}^n)$. Note that (8.49) fails in this case; indeed,

$$\|T\|_{L^2\to L^2} = \sup_{f\in\mathscr{S}} \frac{\|Tf\|_2}{\|f\|_2} = \sup \frac{\|fh\|_2}{\|f\|_2} = \|h\|_\infty = \infty.$$

Before we begin the lengthy proof of this theorem, we state a lemma that we need.

Lemma 8.28. *Let $K \in \mathrm{SK}(\delta, A)$; then, there is a constant C_n such that for all normalized bumps φ, we have*

$$\sup_{x_0\in\mathbb{R}^n} \int_{|x-x_0|\geqslant 20R} \left|\int_{\mathbb{R}^n} K(x,y)\tau^{x_0}\varphi_R(y)dy\right|^2 dx \leqslant \frac{C_n A^2}{R^n}. \tag{8.50}$$

Proof. Note that the interior integral in (8.50) is absolutely convergent since $\tau^{x_0}\varphi_R$ is supported in the ball $B(x_0, 10R)$ and x lies in the complement of the double of this ball. To prove (8.50), simply observe that since $|K(x,y)| \leqslant A|x-y|^{-n}$, we have that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n} K(x,y)\tau^{x_0}\varphi_R(y)dy\right| = \left|\int_{\mathbb{R}^n} K(x,y)\varphi_R(y-x_0)dy\right| \\
=&\left|\int_{\mathbb{R}^n} K(x,y+x_0)\varphi_R(y)dy\right| = \left|\int_{\mathbb{R}^n} K(x,Ry+x_0)\varphi(y)dy\right| \\
\leqslant&\int_{B(0,10)} \frac{A}{|x-Ry-x_0|^n}|\varphi(y)|dy \leqslant \int_{B(0,10)} \frac{2^n A}{|x-x_0|^n}|\varphi(y)|dy \\
\leqslant&\frac{2^n A}{|x-x_0|^n}\int_{B(0,10)} dy = \frac{20^n V_n A}{|x-x_0|^n},
\end{aligned}$$

since $|x-Ry-x_0| \geqslant |x-x_0| - R|y| \geqslant |x-x_0| - 10R \geqslant |x-x_0|/2$ whenever $|x-x_0| \geqslant 20R$. It follows that

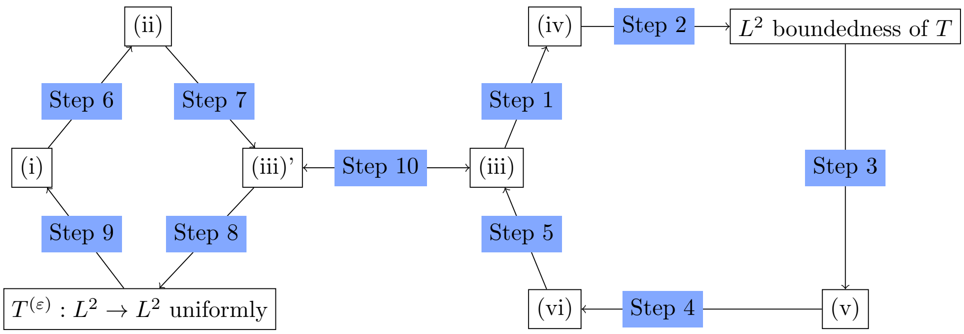
$$\begin{aligned}
&\sup_{x_0\in\mathbb{R}^n}\int_{|x-x_0|\geqslant 20R}\left|\int_{\mathbb{R}^n} K(x,y)\tau^{x_0}\varphi_R(y)dy\right|^2 dx \\
\leqslant&\sup_{x_0\in\mathbb{R}^n}\int_{|x-x_0|\geqslant 20R}\frac{20^{2n}V_n^2A^2}{|x-x_0|^{2n}}dx \\
=&20^{2n}V_n^2A^2\omega_{n-1}\int_{20R}^{\infty} r^{-2n+n-1}dr \\
=&\frac{20^{2n}V_n^2A^2\omega_{n-1}}{n}(20R)^{-n} \\
=&\frac{20^n V_n^3 A^2}{R^n}.
\end{aligned}$$

Therefore, we complete the proof. ❑

§ 8.3.2 Proof of the $T(1)$ theorem

This subsection is dedicated to the proof of Theorem 8.26.

Proof. The proof is based on a series of steps as described in the following picture.



Step 1. (iii) $\Longrightarrow$ (iv).

Fix a $\phi \in \mathscr{D}$ with $0 \leqslant \phi \leqslant 1$, supported in ball $B(0,4)$ and equal to 1 on ball $B(0,2)$. We consider the function $\phi(\cdot/R)$ that tends to 1 as $R \to \infty$, and we show that $T(1)$ is the weak limit of the functions $T(\phi(\cdot/R))$. This means that for all $g \in \mathscr{D}_0$, one has

$$\langle T(\phi(\cdot/R)), g\rangle \to \langle T(1), g\rangle \tag{8.51}$$

as $R \to \infty$. To prove (8.51), we fix an $\eta \in \mathscr{D}$ that is equal to 1 on a neighborhood of the support of g, which implies $\operatorname{dist}(\operatorname{supp} g, \operatorname{supp}(1-\eta)) = c > 0$ and $g = 0$ on $(\operatorname{supp}\eta)^c$. Then we write

$$\begin{aligned}\langle T(\phi(\cdot/R)), g\rangle =&\langle T(\eta\phi(\cdot/R)), g\rangle + \langle T((1-\eta)\phi(\cdot/R)), g\rangle\\ =&\langle T(\eta\phi(\cdot/R)), g\rangle &(8.52)\\ &+\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(K(x,y)-K(x_0,y))g(x)(1-\eta(y))\phi(\frac{y}{R})dydx, &(8.53)\end{aligned}$$

where $x_0 \in \operatorname{supp} g$. There exists $R_0 > 0$ such that $\operatorname{supp}\eta \subset \{y : |y| \leqslant 2R_0\}$, then for $R \geqslant R_0$, $\phi(\cdot/R)$ is equal to 1 on $\operatorname{supp}\eta$. By the similar argument (to the integration over I_0 and I_∞) behind Definition 8.16, we obtain that (8.53) converges to

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(K(x,y)-K(x_0,y))g(x)(1-\eta(y))dydx$$

as $R \to \infty$ by the dominated convergence theorem. By Definition 8.16, we obtain the validity of (8.51).

Next, we observe that the function $\phi(\cdot/R)$ is in L^2 since $\|\phi(\cdot/R)\|_2 = R^{n/2}\|\phi\|_2$. We show that

$$\|T(\phi(\cdot/R))\|_{\mathrm{BMO}} \leqslant C_{n,\delta}(A+B_3) \tag{8.54}$$

uniformly in $R > 0$. Once (8.54) is established, then the sequence

$$\{T(\phi(\cdot/j))\}_{j=1}^{\infty}$$

lies in a multiple of the unit ball of BMO $= (\mathcal{H}^1)^*$, and by the Banach-Alaoglu theorem, there is a subsequence of the positive integers R_j such that $T(\phi(\cdot/R_j))$ converges weakly to an element b in BMO. This means that

$$\langle T(\phi(\cdot/R_j)), g\rangle \to \langle b, g\rangle \tag{8.55}$$

as $j \to \infty$ for all $g \in \mathscr{D}_0$. Using (8.51), we conclude that $T(1)$ can be identified with the BMO function b, and as a consequence of (8.54), it satisfies

$$\|T(1)\|_{\mathrm{BMO}} \leqslant C_{n,\delta}(A+B_3).$$

In a similar fashion, we identify $T^t(1)$ with a BMO function with norm satisfying

$$\|T^t(1)\|_{\mathrm{BMO}} \leqslant C_{n,\delta}(A+B_3).$$

We return to the proof of (8.54). We fix a ball $B = B(x_0, r)$ with radius $r > 0$ centered at $x_0 \in \mathbb{R}^n$. If for all $R > 0$, we had a constant $c_{B,R}$ such that

$$\frac{1}{|B|}\int_B |T(\phi(\cdot/R))(x) - c_{B,R}|dx \leqslant c_{n,\delta}(A + B_3), \tag{8.56}$$

then property (iii) in Proposition 7.15 (adapted to balls) would yield (8.54). Obviously, (8.56) is a consequence of the two estimates

$$\frac{1}{|B|}\int_B \left| T\left[\phi\left(\frac{\cdot - x_0}{r}\right)\phi\left(\frac{\cdot}{R}\right)\right](x)\right| dx \leqslant c_n B_3, \tag{8.57}$$

$$\frac{1}{|B|}\int_B \left| T\left[\left(1-\phi\left(\frac{\cdot - x_0}{r}\right)\right)\phi\left(\frac{\cdot}{R}\right)\right](x) - T\left[\left(1-\phi\left(\frac{\cdot - x_0}{r}\right)\right)\phi\left(\frac{\cdot}{R}\right)\right](x_0)\right| dx \leqslant \frac{c_n}{\delta}A. \tag{8.58}$$

We bound the double integral in (8.58) by

$$\frac{1}{|B|}\int_B \int_{|y-x_0|\geqslant 2r} |K(x,y) - K(x_0,y)|\phi(y/R)dydx, \tag{8.59}$$

since $1 - \phi((y-x_0)/r) = 0$ when $|y - x_0| \leqslant 2r$. Since $|x - x_0| \leqslant r \leqslant \frac{1}{2}|y - x_0|$, condition (8.2) gives that (8.58) holds with $c_n = \omega_{n-1}$.

It remains to be proven (8.57). It is easy to verify that there is a constant $C_0 = C_0(n, \phi)$ such that for $0 < \varepsilon \leqslant 1$ and for all $a \in \mathbb{R}^n$, the functions

$$C_0^{-1}\phi(\varepsilon(x+a))\phi(x), \qquad C_0^{-1}\phi(x)\phi(-a+\varepsilon x) \tag{8.60}$$

are normalized bumps. The important observation is that with $a = x_0/r$ we have

$$\phi\left(\frac{x}{R}\right)\phi\left(\frac{x-x_0}{r}\right) = r^n \tau^{x_0}\left[\left(\phi\left(\frac{r}{R}(\cdot + a)\right)\phi(\cdot)\right)_r\right](x) \tag{8.61}$$

$$= R^n\left(\phi(\cdot)\phi\left(-a + \frac{R}{r}(\cdot)\right)\right)_R(x), \tag{8.62}$$

and thus in either case $r \leqslant R$ or $R \leqslant r$, one may express the product $\phi\left(\frac{x}{R}\right)\phi\left(\frac{x-x_0}{r}\right)$ as a multiple of a translation of an L^1 dilation of a normalized bump.

Case 1: $r \leqslant R$**.** In view of (8.61), we write

$$T\left[\phi\left(\frac{\cdot - x_0}{r}\right)\phi\left(\frac{\cdot}{R}\right)\right](x) = C_0 r^n T[\tau^{x_0}\varphi_r](x)$$

for some normalized bump φ. Using this fact and the Cauchy-Schwarz inequality, we estimate the expression on the left in (8.57) by

$$\frac{1}{|B|}\int_B |C_0 r^n T[\tau^{x_0}\varphi_r](x)|dx$$

$$\leqslant C_0 r^n \frac{1}{|B|}|B|^{1/2}\left(\int_B |T[\tau^{x_0}\varphi_r](x)|^2 dx\right)^{1/2}$$

$$\leqslant C_0 \frac{r^{n/2}}{|B|^{1/2}} B_3 \leqslant \frac{C_0}{V_n^{1/2}} B_3 = c_n B_3,$$

where the second inequality follows by applying hypothesis (iii).

Case 2: $R \leqslant r$**.** In view of (8.62), we write

$$T\left[\phi\left(\frac{\cdot - x_0}{r}\right)\phi\left(\frac{\cdot}{R}\right)\right](x) = C_0 R^n T(\varphi_R)(x)$$

for some other normalized bump φ. Using this fact and the Cauchy-Schwarz inequality, we estimate the expression on the left in (8.57) by

$$\begin{aligned}
&\frac{1}{|B|}\int_B |C_0 R^n T(\varphi_R)(x)|dx \\
\leqslant &\frac{C_0 R^n}{|B|^{1/2}}\left(\int_B |T(\varphi_R)(x)|^2 dx\right)^{1/2} \\
\leqslant &\frac{C_0 R^{n/2}}{|B|^{1/2}} B_3 = \frac{C_0 R^{n/2}}{V_n^{1/2} r^{n/2}} B_3 \leqslant \frac{C_0}{V_n^{1/2}} B_3 = c_n B_3
\end{aligned}$$

by applying hypothesis (iii) and recalling that $R \leqslant r$. This proves (8.57).

To finish the proof of (iv), we need to prove that T satisfies the weak boundedness property. However, this is elementary, since for all normalized bumps φ and ψ and all $x \in \mathbb{R}^n$ and $R > 0$, we have

$$\begin{aligned}
|\langle T(\tau^x \psi_R), \tau^x \varphi_R\rangle| &\leqslant \|T(\tau^x \psi_R)\|_2 \|\tau^x \varphi_R\|_2 \\
&\leqslant B_3 R^{-n/2} R^{-n/2} \|\varphi\|_2 \\
&\leqslant B_3 R^{-n}(V_n 10^n)^{1/2} = C_n B_3 R^{-n}.
\end{aligned}$$

This gives $\|T\|_{WB} \leqslant C_n B_3$, which implies the estimate $B_4 \leqslant C_{n,\delta}(A + B_3)$ and concludes the proof of the fact that condition (iii) implies (iv).

Step 2. (iv) $\Longrightarrow$ L^2 boundedness of T.

We now assume condition (iv), and we present the most important step of the proof, establishing the fact that T has an extension that maps $L^2(\mathbb{R}^n)$ to itself. The assumption that the distributions $T(1)$ and $T^t(1)$ coincide with BMO functions leads to the construction of Carleson measures that provide the key tool in the boundedness of T.

We pick a radial function $\Phi \in \mathscr{D}$ supported in the ball $B(0, 1/2)$ that satisfies $\int_{\mathbb{R}^n} \Phi(x)dx = 1$. For $t > 0$, we define $\Phi_t(x) = t^{-n}\Phi(x/t)$. Since Φ is a radial function, the operator

$$P_t(f) = f * \Phi_t \tag{8.63}$$

is self-transpose.

We now fix a Schwartz function f whose Fourier transform is supported away from a neighborhood of the origin. We discuss an integral representation for $T(f)$. We begin with the facts that can be found in Exercises 8.3 and 8.4:

$$T(f) = \lim_{s\to 0} P_s^2 T P_s^2(f) \quad \text{in } \mathscr{S}',$$
$$0 = \lim_{s\to\infty} P_s^2 T P_s^2(f) \quad \text{in } \dot{\mathscr{S}}',$$

where $\dot{\mathscr{S}}(\mathbb{R}^n) = \{f \in \mathscr{S}(\mathbb{R}^n) : \int_{\mathbb{R}^n} x^\alpha f(x)dx = 0, \forall \alpha \in \mathbb{N}_0^n\}$ is a subspace of $\mathscr{S}(\mathbb{R}^n)$ with the same topology, and its dual space $\dot{\mathscr{S}}'(\mathbb{R}^n) = \mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n)$. Thus, by the fundamental theorem of calculus and the product rule, we can write

$$\begin{aligned}
T(f) &= \lim_{s\to 0} P_s^2 T P_s^2(f) - \lim_{s\to\infty} P_s^2 T P_s^2(f)\\
&= -\lim_{\varepsilon\to 0}\int_\varepsilon^{1/\varepsilon} s\frac{d}{ds}(P_s^2 T P_s^2)(f)\frac{ds}{s}\\
&= -\lim_{\varepsilon\to 0}\int_\varepsilon^{1/\varepsilon}\left[\left(s\frac{d}{ds}P_s^2\right)TP_s^2(f) + P_s^2\left(Ts\frac{d}{ds}P_s^2\right)(f)\right]\frac{ds}{s},
\end{aligned} \tag{8.64}$$

where the limit is in the sense of $\dot{\mathscr{S}}'$. For $g \in \mathscr{S}$, we have by (8.63)

$$\begin{aligned}
\widehat{\left(s\tfrac{d}{ds}P_s^2(g)\right)}(\xi) &= (2\pi)^n \widehat{g}(\xi) s\frac{d}{ds}(\widehat{\Phi_s(\xi)}^2)\\
&= (2\pi)^n \widehat{g}(\xi) s\frac{d}{ds}((\widehat{\Phi}(s\xi))^2)\\
&= (2\pi)^n \widehat{g}(\xi)\widehat{\Phi}(s\xi)2s\xi\cdot\nabla\widehat{\Phi}(s\xi)\\
&= (2\pi)^n \widehat{g}(\xi)\sum_{k=1}^n \widehat{\Psi_k}(s\xi)\widehat{\Theta_k}(s\xi)\\
&= \sum_{k=1}^n \widehat{\widetilde{Q}_{k,s}Q_{k,s}(g)}(\xi) = \sum_{k=1}^n \widehat{Q_{k,s}\widetilde{Q}_{k,s}(g)}(\xi),
\end{aligned}$$

where for $1 \leqslant k \leqslant n$, $\widehat{\Psi_k}(\xi) = 2\xi_k\widehat{\Phi}(\xi)$, $\widehat{\Theta_k}(\xi) = \partial_k\widehat{\Phi}(\xi)$, and $Q_{k,s}$, $\widetilde{Q}_{k,s}$ are operators defined by

$$Q_{k,s}(g) = g * (\Psi_k)_s, \qquad \widetilde{Q}_{k,s}(g) = g * (\Theta_k)_s;$$

where $(\Theta_k)_s = s^{-n}\Theta_k(s^{-1}x)$ and $(\Psi_k)_s$ are defined similarly. Observe that Ψ_k and Θ_k are smooth odd bumps supported in $B(0,1/2)$ and have an integral of zero. Since it is easy to see that $\widehat{\Phi}$ is also radial, we obtain

$$\begin{aligned}
\Psi_k(-x) &= \int_{\mathbb{R}^n} e^{ix\cdot(-\xi)}2\xi_k\widehat{\Phi}(\xi)\,đ\xi\\
&= -\int_{\mathbb{R}^n} e^{ix\cdot\eta}2\eta_k\widehat{\Phi}(\eta)\,đ\eta = -\Psi_k(x),
\end{aligned}$$

and

$$\Psi_k(x) = -2i\partial_k\Phi(x), \quad \Theta_k(x) = -ix_k\Phi(x).$$

Since Ψ_k and Θ_k are odd, we have $(Q_{k,s})^t = -Q_{k,s}$ and $(\widetilde{Q}_{k,s})^t = -\widetilde{Q}_{k,s}$, that is, they are anti-self-transpose. We now write the expression in (8.64) as

$$-\lim_{\varepsilon\to 0}\sum_{k=1}^{n}\left[\int_{\varepsilon}^{1/\varepsilon}\widetilde{Q}_{k,s}Q_{k,s}TP_sP_s(f)\frac{ds}{s}+\int_{\varepsilon}^{1/\varepsilon}P_sP_sTQ_{k,s}\widetilde{Q}_{k,s}(f)\frac{ds}{s}\right], \tag{8.65}$$

where the limit is in the sense of $\dot{\mathscr{S}}'$. We set

$$T_{k,s}=Q_{k,s}TP_s,$$

and then taking the transpose, i.e.,

$$(T_{k,s})^t=(Q_{k,s}TP_s)^t=P_s^tT^t(Q_{k,s})^t=-P_sT^tQ_{k,s},$$

we obtain $P_sTQ_{k,s}=-((T^t)_{k,s})^t$.

Recall the notation $\tau^x h(z)=h(z-x)$. For a given $\varphi\in\mathscr{S}(\mathbb{R}^n)$, we have

$$\begin{aligned}
T_{k,s}(\varphi)(x)=&Q_{k,s}TP_s(\varphi)(x)=TP_s(\varphi)*(\Psi_k)_s(x)\\
=&\int TP_s(\varphi)(y)(\Psi_k)_s(x-y)dy\\
=&-\int TP_s(\varphi)(y)(\Psi_k)_s(y-x)dy\quad(\text{since }\Psi_k\text{ is odd})\\
=&-\int TP_s(\varphi)(y)\tau^x(\Psi_k)_s(y)dy\\
=&-\langle TP_s(\varphi),\tau^x(\Psi_k)_s\rangle\\
=&-\langle T(\Phi_s*\varphi),\tau^x(\Psi_k)_s\rangle\\
=&-\left\langle T\left(\int_{\mathbb{R}^n}\varphi(y)(\Phi_s(\cdot-y))dy\right),\tau^x(\Psi_k)_s\right\rangle\\
=&-\left\langle T\left(\int_{\mathbb{R}^n}\varphi(y)(\tau^y\Phi_s)(\cdot)dy\right),\tau^x(\Psi_k)_s\right\rangle\\
=&-\int_{\mathbb{R}^n}\langle T(\tau^y\Phi_s),\tau^x(\Psi_k)_s\rangle\varphi(y)dy.
\end{aligned}\tag{8.66}$$

The last equality is justified by the convergence of the Riemann sums R_N of the integral $I=\int_{\mathbb{R}^n}\varphi(y)(\tau^y\Phi_s)(\cdot)dy$ to itself in the topology of $\mathscr{S}$ (this is contained in the proof of Theorem 3.37); by the continuity of T, $T(R_N)$ converges to $T(I)$ in $\mathscr{S}'$ and thus $\langle T(R_N),\tau^x(\Psi_k)_s\rangle$ converges to $\langle T(I),\tau^x(\Psi_k)_s\rangle$. However, $\langle T(R_N),\tau^x(\Psi_k)_s\rangle$ is also a Riemann sum for the rapidly convergent integral in (8.66); hence it converges to it as well.

We have deduced that the operator $T_{k,s}=Q_{k,s}TP_s$ has kernel

$$K_{k,s}(x,y)=-\langle T(\tau^y\Phi_s),\tau^x(\Psi_k)_s\rangle=-\langle T^t(\tau^x(\Psi_k)_s),\tau^y\Phi_s\rangle. \tag{8.67}$$

Hence, the operator $P_sTQ_{k,s}=-((T^t)_{k,s})^t$ has kernel

$$-((K^t)_{k,s})^t(x,y)=(\langle T(\tau^x(\Psi_k)_s),\tau^y\Phi_s\rangle)^t=\langle T(\tau^y(\Psi_k)_s),\tau^x\Phi_s\rangle.$$

For $1\leqslant k\leqslant n$, we need the following facts regarding these kernels:

$$|\langle T(\tau^y(\Psi_k)_s),\tau^x\Phi_s\rangle|\leqslant C_{n,\delta}(\|T\|_{WB}+A)p_s(x-y), \tag{8.68}$$

$$|\langle T^t(\tau^x(\Psi_k)_s),\tau^y\Phi_s\rangle|\leqslant C_{n,\delta}(\|T\|_{WB}+A)p_s(x-y), \tag{8.69}$$

where

$$p_t(u) = \frac{1}{t^n}\frac{1}{(1+|u/t|)^{n+\delta}}$$

is the L^1 dilation of the function $p(u) = (1+|u|)^{-n-\delta}$.

To prove (8.69), we consider the following two cases.

Case 1: $|x-y| \leqslant 5s$. The weak boundedness property gives

$$|\langle T(\tau^y\Phi_s), \tau^x(\Psi_k)_s\rangle| = |\langle T(\tau^x((\tau^{(y-x)/s}\Phi)_s)), \tau^x(\Psi_k)_s\rangle| \leqslant \frac{C_{n,\Phi}\|T\|_{WB}}{s^n},$$

since

$$\begin{aligned}\tau^y\Phi_s(z) =&\Phi_s(z-y) = s^{-n}\Phi((z-y)/s) = s^{-n}\Phi((z-x)/s-(y-x)/s)\\ =&s^{-n}\tau^{(y-x)/s}\Phi((z-x)/s) = (\tau^{(y-x)/s}\Phi)_s(z-x)\\ =&\tau^x(\tau^{(y-x)/s}\Phi)_s(z),\end{aligned}$$

and both Ψ_k and $\tau^{(y-x)/s}\Phi$ are multiples of normalized bumps. Note here that $\operatorname{supp}\Psi_k \subset B(0,1/2) \subset B(0,10)$, and $\operatorname{supp}\tau^{(y-x)/s}\Phi \subset B(0,10)$ due to $|z| \leqslant |z-s^{-1}(y-x)| + |s^{-1}(y-x)| \leqslant 1/2+5 \leqslant 10$ from $\operatorname{supp}\Phi \subset B(0,1/2)$ and $s^{-1}|x-y| \leqslant 5$. This estimate proves (8.69) when $|x-y| \leqslant 5s$ due to $s^{-n} \leqslant Cp_s(x-y)$.

Case 2: $|x-y| \geqslant 5s$. For $z_1 \in \operatorname{supp}\tau^y\Phi_s$ and $z_2 \in \tau^x(\Psi_k)_s$, we have

$$\begin{aligned}|z_1-z_2| =&|z_1-y+y-x+x-z_2| \geqslant |y-x|-|z_1-y|-|x-z_2|\\ \geqslant&5s-s/2-s/2 = 4s > 0,\end{aligned}$$

i.e., the functions $\tau^y\Phi_s$ and $\tau^x(\Psi_k)_s$ have disjoint supports and so we have the integral representation by Proposition 8.10

$$\langle T^t(\tau^x(\Psi_k)_s), \tau^y\Phi_s\rangle = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\Phi_s(v-y)K(u,v)(\Psi_k)_s(u-x)dudv.$$

Because Ψ_k has a mean value of zero, we can write the previous expression as

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\Phi_s(v-y)(K(u,v)-K(x,v))(\Psi_k)_s(u-x)dudv.$$

We observe that $|u-x| \leqslant s$ and $|v-y| \leqslant s$ in the preceding double integral. Since $|x-y| \geqslant 5s$, $|u-v| \geqslant |x-y|-2s \geqslant 3s$, which implies that $|u-x| \leqslant |u-v|/2$. Using (8.2), we obtain

$$|K(u,v)-K(x,v)| \leqslant \frac{A|x-u|^\delta}{(|u-v|+|x-v|)^{n+\delta}} \leqslant \frac{C_{n,\delta}As^\delta}{(s+|x-y|)^{n+\delta}},$$

where we used the fact that $|x-v| \geqslant |x-y|-|y-v| \geqslant 4s$ and $|u-v| \approx |x-y|$ due to $|u-v| \leqslant |u-x|+|x-y|+|y-v| \leqslant |u-v|/2+|x-y|+s \leqslant 5|u-v|/6+|x-y|$ and $|u-v| \geqslant |x-y|-2s \geqslant |x-y|-2|x-y|/5 = 3|x-y|/5$. Inserting this estimate in the double integral, we obtain (8.69). Estimate (8.68) is proved similarly.

At this point, we drop the dependence of $Q_{k,s}$ and $\widetilde{Q}_{k,s}$ on the index

k since we can concentrate on one term of the sum in (8.65). We have managed to express $-T(f)$ as a finite sum of operators of the form

$$\int_0^\infty \widetilde{Q}_s T_s P_s(f) \frac{ds}{s} \tag{8.70}$$

and of the form

$$\int_0^\infty P_s T_s \widetilde{Q}_s(f) \frac{ds}{s}, \tag{8.71}$$

where the preceding integrals converge in $\dot{\mathscr{S}}'(\mathbb{R}^n)$ and T_s's have kernels $K_s(x,y)$, which are pointwise dominated by a constant multiple of $(A + B_4)p_s(x-y)$.

It suffices to prove L^2 bounds for an operator of the form (8.70) with constant at most a multiple of $A + B_4$, since by duality the same estimate also holds for the operators of the form (8.71). We make one more observation. Using (8.67) (recall that we have dropped the indices k), we obtain

$$\begin{aligned} T_s(1)(x) = \int_{\mathbb{R}^n} K_s(x,y)dy &= -\int_{\mathbb{R}^n} \langle T(\tau^y\Phi_s), \tau^x(\Psi_s)\rangle dy \\ &= -\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} T(\tau^y\Phi_s)(z)\Psi_s(z-x)dzdy \\ &= -\int_{\mathbb{R}^n} T\left(\int_{\mathbb{R}^n} \Phi_s(z-y)dy\right)\Psi_s(z-x)dz \\ &= -\int_{\mathbb{R}^n} T(1)\Psi_s(z-x)dz \\ &= \int_{\mathbb{R}^n} T(1)\Psi_s(x-z)dz \\ &= (\Psi_s * T(1))(x), \end{aligned} \tag{8.72}$$

where all integrals converge absolutely.

We can therefore concentrate on the L^2 boundedness of the operator in (8.70). We pair this operator with $g \in \dot{\mathscr{S}}$ and we use the convergence of the integral in $\dot{\mathscr{S}}'(\mathbb{R}^n)$ and the property $(\widetilde{Q}_s)^t = -\widetilde{Q}_s$ to obtain

$$\begin{aligned} \left\langle \int_0^\infty \widetilde{Q}_s T_s P_s(f)\frac{ds}{s}, g\right\rangle &= \int_0^\infty \left\langle \widetilde{Q}_s T_s P_s(f), g\right\rangle \frac{ds}{s} \\ &= -\int_0^\infty \left\langle T_s P_s(f), \widetilde{Q}_s(g)\right\rangle \frac{ds}{s}. \end{aligned}$$

The intuition here is as follows: T_s is an averaging operator at scale s, and $P_s(f)$ is essentially constant on that scale. Therefore, the expression $T_sP_s(f)$ must look like $T_s(1)P_s(f)$. To be precise, we introduce this term and try to estimate the error that occurs. We have

$$T_sP_s(f) = T_s(1)P_s(f) + [T_sP_s(f) - T_s(1)P_s(f)]. \tag{8.73}$$

We estimate the terms that arise from this splitting. Recalling (8.72), we write by using the Cauchy-Schwarz inequality

$$\begin{aligned}
&\left|\int_0^\infty \left\langle (\Psi_s * T(1))P_s(f), \widetilde{Q}_s(g)\right\rangle \frac{ds}{s}\right| \\
\leqslant &\left(\int_0^\infty \|P_s(f)(\Psi_s * T(1))\|_2^2 \frac{ds}{s}\right)^{1/2} \left(\int_0^\infty \|\widetilde{Q}_s(g)\|_2^2 \frac{ds}{s}\right)^{1/2} \\
= &\left(\int_0^\infty \int_{\mathbb{R}^n} |P_s(f)(\Psi_s * T(1))|^2 dx \frac{ds}{s}\right)^{1/2} \left\|\left(\int_0^\infty |\widetilde{Q}_s(g)|^2 \frac{ds}{s}\right)^{1/2}\right\|_2. \quad (8.74)
\end{aligned}$$

Since $T(1) \in \mathrm{BMO}$, $\Psi \in L^1$ with $\int_{\mathbb{R}^n} \Psi(x)dx = 0$ satisfying $|\Psi(x)| \leqslant C(1+|x|)^{-n-\delta}$, and due to $\Phi \in \mathscr{D} \subset \mathscr{S}$, we obtain for some $N \in \mathbb{N}$

$$\begin{aligned}
\sup_{\xi \in \mathbb{R}^n} \int_0^\infty |\widehat{\Psi}(s\xi)|^2 \frac{ds}{s} &= \sup_{\xi} \int_0^\infty |2s\xi \widehat{\Phi}(s\xi)|^2 \frac{ds}{s} \\
&= \sup_{\xi} \int_0^\infty |\widehat{\Phi}(s\xi)|^2 d(s^2|\xi|^2) \\
&\leqslant C \sup_{\xi} \int_0^\infty \frac{d(s^2|\xi|^2)}{(1+|s\xi|^2)^N} < \infty,
\end{aligned}$$

We obtain from Theorem 7.40 (ii) that $|(\Psi_s * T(1))(x)|^2 dx \frac{ds}{s}$ is a Carleson measure on $\mathbb{R}^{n+1}_+$ with norm at most $C_{n,\delta}\|T(1)\|^2_{\mathrm{BMO}}$. Then, from Theorem 7.39, it follows

$$\left(\int_0^\infty \int_{\mathbb{R}^n} |P_s(f)(x)|^2 |(\Psi_s * T(1))(x)|^2 dx \frac{ds}{s}\right)^{1/2} \leqslant C_n \|T(1)\|_{\mathrm{BMO}} \|f\|_2.$$

For the second factor in (8.74), by the continuous version of the Littlewood-Paley theorem (Exercise 5.10), we have

$$\left\|\left(\int_0^\infty |\widetilde{Q}_s(g)|^2 \frac{ds}{s}\right)^{1/2}\right\|_2 \leqslant C_n \|g\|_2.$$

Thus, we obtain

$$(8.74) \leqslant C_n \|T(1)\|_{\mathrm{BMO}} \|f\|_2 \|g\|_2 \leqslant C_n B_4 \|f\|_2 \|g\|_2.$$

This gives the sought estimate for the first term in (8.73). For the second term in (8.73), we have by the Cauchy-Schwarz inequality, Exercise 5.10, (8.67) and (8.69)

$$\begin{aligned}
&\left|\int_0^\infty \int_{\mathbb{R}^n} \widetilde{Q}_s(g)(x)[T_s P_s(f) - T_s(1)P_s(f)](x) dx \frac{ds}{s}\right| \\
\leqslant &\left(\int_0^\infty \int_{\mathbb{R}^n} |\widetilde{Q}_s(g)(x)|^2 dx \frac{ds}{s}\right)^{1/2} \\
&\cdot \left(\int_0^\infty \int_{\mathbb{R}^n} |[T_s P_s(f) - T_s(1)P_s(f)](x)|^2 dx \frac{ds}{s}\right)^{1/2} \\
\leqslant &C_n \|g\|_2 \left(\int_0^\infty \int_{\mathbb{R}^n} \left|\int_{\mathbb{R}^n} K_s(x,y)[P_s(f)(y) - P_s(f)(x)]dy\right|^2 dx \frac{ds}{s}\right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
&\leqslant C_n(\|T\|_{WB}+A)\|g\|_2\\
&\qquad\left(\int_0^\infty\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}p_s(x-y)|P_s(f)(y)-P_s(f)(x)|dy\right)^2dx\frac{ds}{s}\right)^{1/2}\\
&\leqslant C_n(A+B_4)\|g\|_2\Big(\int_0^\infty\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}p_s(x-y)|P_s(f)(y)-P_s(f)(x)|^2dy\\
&\quad\Big(\int_{\mathbb{R}^n}p_s(x-y)dy\Big)dx\frac{ds}{s}\Big)^{1/2}\\
&\leqslant C_n(A+B_4)\|g\|_2\\
&\quad\cdot\left(\int_0^\infty\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}p_s(x-y)|P_s(f)(y)-P_s(f)(x)|^2dydx\frac{ds}{s}\right)^{1/2}.
\end{aligned}$$

It suffices to estimate the last displayed square root. Changing variables $u=x-y$, applying Plancherel's theorem and the inequality $|1-e^{i\theta}|\leqslant|\theta|$, we express this square root as

$$\begin{aligned}
&\left(\int_0^\infty\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}p_s(u)|P_s(f)(y)-P_s(f)(y+u)|^2dudy\frac{ds}{s}\right)^{1/2}\\
=&(2\pi)^{n/2}\left(\int_0^\infty\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}p_s(u)|\widehat{\Phi}(s\xi)(1-e^{iu\cdot\xi})|^2|\hat{f}(\xi)|^2dud\xi\frac{ds}{s}\right)^{1/2}\\
\leqslant&(2\pi)^{n/2}\left(\int_0^\infty\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}p_s(u)|\widehat{\Phi}(s\xi)|^2 2^{2-\delta/2}|u|^{\delta/2}|\xi|^{\delta/2}|\hat{f}(\xi)|^2dud\xi\frac{ds}{s}\right)^{1/2}\\
=&(2\pi)^{n/2}2^{1-\delta/4}\\
&\cdot\left(\int_{\mathbb{R}^n}\int_0^\infty\left(\int_{\mathbb{R}^n}p_s(u)|u/s|^{\delta/2}du\right)|\widehat{\Phi}(s\xi)|^2|s\xi|^{\delta/2}\frac{ds}{s}|\hat{f}(\xi)|^2d\xi\right)^{1/2},
\end{aligned}$$

and we claim that this last expression is bounded by $C_{n,\delta}\|f\|_2$. Indeed, we first bound the quantity

$$\begin{aligned}
\int_{\mathbb{R}^n}p_s(u)|u/s|^{\delta/2}du=&\int_{\mathbb{R}^n}\frac{|u/s|^{\delta/2}}{(1+|u/s|)^{n+\delta}}d(u/s)\\
\leqslant&\int_{\mathbb{R}^n}\frac{1}{(1+|v|)^{n+\delta/2}}dv=C<\infty,
\end{aligned}$$

and then we use the estimate (due to $\widehat{\Phi}\in\mathscr{S}$)

$$\int_0^\infty|\widehat{\Phi}(s\xi)|^2|s\xi|^{\delta/2}\frac{ds}{s}=\int_0^\infty|\widehat{\Phi}(se_1)|^2s^{\delta/2}\frac{ds}{s}\leqslant C'_{n,\delta}<\infty$$

where $e_1=(1,0,\cdots,0)$, and Plancherel's theorem to obtain the claim. Since $\dot{\mathscr{S}}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ by Exercise 5.12, we deduce by duality that

$$\|Tf\|_2=\sup_{g\in\dot{\mathscr{S}},\,\|g\|_2\leqslant1}|\langle Tf,g\rangle|\leqslant C_{n,\delta}(A+B_4)\|f\|_2,$$

for all $f\in\dot{\mathscr{S}}$. Thus T admits an extension on L^2 that satisfies

$$\|T\|_{L^2\to L^2}\leqslant C_{n,\delta}(A+B_4).$$

Step 3. L^2 boundedness of $T \Longrightarrow$ (v).

If T has an extension that maps L^2 to itself, then by Theorem 8.23, we have

$$B_5 \leqslant C_{n,\delta}(A + \|T\|_{L^2 \to L^2}).$$

Thus, the boundedness of T on L^2 implies (v).

Step 4. (v) $\Longrightarrow$ (vi).

At a formal level the proof of this fact is clear, since we can write a normalized bump as the inverse Fourier transform of its Fourier transform and interchange the integration with the action of T to obtain

$$T(\tau^{x_0}\varphi_R) = \int_{\mathbb{R}^n} \widehat{\tau^{x_0}\varphi_R}(\xi) T(e^{i\xi\cdot(\cdot)})\,đ\xi. \tag{8.75}$$

The conclusion follows by taking BMO norms. To make identity (8.75) precise, we provide the following argument.

Let us fix a normalized bump φ and a function $g \in \mathscr{D}_0$. We select a function $\eta \in \mathscr{D}$ that is equal to 1 on the unit ball and vanishes outside the double of that ball. Define $\eta_k(\xi) = \eta(\xi/k)$ and note that η_k tends pointwise to 1 as $k \to \infty$. Observe that $\eta_k \tau^{x_0}\varphi_R$ converges to $\tau^{x_0}\varphi_R$ in $\mathscr{S}(\mathbb{R}^n)$ as $k \to \infty$, and by the continuity of T we obtain

$$\lim_{k\to\infty} \langle T(\eta_k \tau^{x_0}\varphi_R), g\rangle = \langle T(\tau^{x_0}\varphi_R), g\rangle.$$

We have

$$\begin{aligned} T(\eta_k \tau^{x_0}\varphi_R) &= T\left(\int_{\mathbb{R}^n} \widehat{\tau^{x_0}\varphi_R}(\xi)\eta_k(\cdot)e^{i\xi\cdot(\cdot)}\,đ\xi\right) \\ &= \int_{\mathbb{R}^n} \widehat{\tau^{x_0}\varphi_R}(\xi) T(\eta_k(\cdot)e^{i\xi\cdot(\cdot)})\,đ\xi, \end{aligned} \tag{8.76}$$

where the second equality is justified by the continuity and linearity of T along with the fact that the Riemann sums of the integral in (8.76) converge to that integral in $\mathscr{S}$ (a proof of this fact is essentially contained in the proof of Theorem 3.37). Consequently,

$$\langle T(\tau^{x_0}\varphi_R), g\rangle = \lim_{k\to\infty} \int_{\mathbb{R}^n} \widehat{\tau^{x_0}\varphi_R}(\xi)\langle T(\eta_k e^{i\xi\cdot(\cdot)}), g\rangle\,đ\xi. \tag{8.77}$$

We show that $\langle T(\eta_k e^{i\xi\cdot(\cdot)}), g\rangle$ is uniformly bounded in k for k large. Suppose that g is supported in the ball $\overline{B(0,M)}$. Let $k_0 = 2M$. Then, for $k \geqslant k_0$ write

$$\langle T(\eta_k e^{i\xi\cdot(\cdot)}), g\rangle = \langle T(e^{i\xi\cdot(\cdot)}), g\rangle - \langle T((1-\eta_k)e^{i\xi\cdot(\cdot)}), g\rangle. \tag{8.78}$$

The first expression on the r.h.s. of (8.78) is bounded by $B_5\|g\|_{\mathcal{H}^1}$, while the second expression can be written as

$$\int_{|y|\geqslant k} \left[\int_{\mathbb{R}^n} (K(x,y) - K(0,y))g(x)dx\right] (1-\eta_k(y))e^{i\xi\cdot y}dy,$$

in view of Definition 8.16. As $|x| \leqslant \max(|x-y|, |y|)/2$ when $|y| \geqslant k \geqslant$

$k_0 = 2M$ and $|x| \leqslant M$, we use (8.2) to bound the absolute value of the preceding expression by

$$\begin{aligned}
&2\|g\|_\infty \int_{|y|\geqslant 2M}\int_{|x|\leqslant M} \frac{A|x|^\delta}{|x-y|^{n+\delta}}dxdy\\
=&2\|g\|_\infty \int_{|x|\leqslant M} A|x|^\delta \int_{|x-y|\geqslant |x|} \frac{dy}{|x-y|^{n+\delta}}dx = C' < \infty.
\end{aligned}$$

The dominated convergence theorem allows us to pass the limit inside the integrals in (8.77) to obtain

$$\langle T(\tau^{x_0}\varphi_R), g\rangle = \int_{\mathbb{R}^n} \widehat{\tau^{x_0}\varphi_R}(\xi)\langle T(e^{i\xi\cdot(\cdot)}), g\rangle đ\xi.$$

We now use assumption (v). The distribution $T(e^{i\xi\cdot(\cdot)})$ coincides with a BMO function whose norm is at most B_5. It follows that

$$\begin{aligned}
|\langle T(\tau^{x_0}\varphi_R), g\rangle| \leqslant &(2\pi)^{-n/2}\|\widehat{\tau^{x_0}\varphi_R}\|_1 \sup_{\xi\in\mathbb{R}^n}\|T(e^{i\xi\cdot(\cdot)})\|_{\mathrm{BMO}}\|g\|_{\mathcal{H}^1}\\
\leqslant &C_n B_5\|\widehat{\varphi}(R\cdot)\|_1\|g\|_{\mathcal{H}^1}\\
\leqslant &C_n B_5 R^{-n}\|g\|_{\mathcal{H}^1}, \qquad (8.79)
\end{aligned}$$

where the constant C_n is independent of the normalized bump φ in view of (8.47). It follows from (8.79) that

$$g \mapsto \langle T(\tau^{x_0}\varphi_R), g\rangle$$

is a bounded linear functional on BMO with norm at most a multiple of $B_5 R^{-n}$. It follows from Theorem 7.27 that $T(\tau^{x_0}\varphi_R)$ coincides with a BMO function that satisfies

$$R^n\|T(\tau^{x_0}\varphi_R)\|_{\mathrm{BMO}} \leqslant C_n B_5.$$

The same argument is valid for T^t, which shows that

$$B_6 \leqslant C_{n,\delta}(A + B_5)$$

and concludes the proof that (v) implies (vi).

Step 5. (vi) $\Longrightarrow$ (iii).

We fix $x_0 \in \mathbb{R}^n$ and $R > 0$. Pick $z_0 \in \mathbb{R}^n$ such that $|x_0 - z_0| = 40R$. Then, if $|y - x_0| \leqslant 10R$ and $|x - z_0| \leqslant 20R$, we have

$$\begin{aligned}
10R \leqslant &|z_0 - x_0| - |x - z_0| - |y - x_0|\\
\leqslant &|x - y|\\
\leqslant &|x - z_0| + |z_0 - x_0| + |x_0 - y| \leqslant 70R.
\end{aligned}$$

From this it follows that when $|x - z_0| \leqslant 20R$, we have

$$\begin{aligned}
|T(\tau^{x_0}\varphi_R)| = &\left|\int_{|y-x_0|\leqslant 10R} K(x,y)\tau^{x_0}\varphi_R(y)dy\right|\\
\leqslant &\int_{10R\leqslant|x-y|\leqslant 70R} |K(x,y)|\frac{dy}{R^n}
\end{aligned}$$

$$\leqslant \frac{C_{n,\delta} A}{R^n},$$

and thus

$$\left| \operatorname*{Avg}_{B(z_0,20R)} T(\tau^{x_0}\varphi_R) \right| \leqslant \frac{C_{n,\delta} A}{R^n}, \tag{8.80}$$

where $\operatorname*{Avg}_B g$ denotes the average of g over B. Because of assumption (vi), the BMO norm of the function $T(\tau^{x_0}\varphi_R)$ is bounded by a multiple of $B_6 R^{-n}$, a fact used in the following sequence of implications. We have by Corollary 7.26, Exercise 7.5 with $|x_0 - z_0| = 40R$ and (8.80)

$$\begin{aligned}
&\|T(\tau^{x_0}\varphi_R)\|_{L^2(B(x_0,20R))}\\
\leqslant &\left\| T(\tau^{x_0}\varphi_R) - \operatorname*{Avg}_{B(x_0,20R)} T(\tau^{x_0}\varphi_R) \right\|_{L^2(B(x_0,20R))}\\
&+ V_n^{1/2}(20R)^{n/2} \left| \operatorname*{Avg}_{B(x_0,20R)} T(\tau^{x_0}\varphi_R) - \operatorname*{Avg}_{B(z_0,20R)} T(\tau^{x_0}\varphi_R) \right|\\
&+ V_n^{1/2}(20R)^{n/2} \left| \operatorname*{Avg}_{B(z_0,20R)} T(\tau^{x_0}\varphi_R) \right|\\
\leqslant &C_n \|T(\tau^{x_0}\varphi_R)\|_{\mathrm{BMO}} |B(x_0,20R)|^{1/2}\\
&+ V_n^{1/2}(20R)^{n/2} C_n \ln(|x_0 - z_0|/(20R) + 1) \|T(\tau^{x_0}\varphi_R)\|_{\mathrm{BMO}}\\
&+ V_n^{1/2}(20R)^{n/2} \frac{C_{n,\delta} A}{R^n}\\
\leqslant &C_{n,\delta} \left(R^{n/2} \|T(\tau^{x_0}\varphi_R)\|_{\mathrm{BMO}} + R^{-n/2} A \right)\\
\leqslant &C_{n,\delta} R^{-n/2}(B_6 + A).
\end{aligned}$$

Now, we have from Lemma 8.28 that

$$\|T(\tau^{x_0}\varphi_R)\|_{L^2(B(x_0,20R)^c)} \leqslant C_{n,\delta} A R^{-n/2}.$$

Since the same computations can apply to T^t, it follows that

$$R^{n/2} \left(\|T(\tau^{x_0}\varphi_R)\|_{L^2} + \|T^t(\tau^{x_0}\varphi_R)\|_{L^2} \right) \leqslant C_{n,\delta}(A + B_6), \tag{8.81}$$

which proves that $B_3 \leqslant C_{n,\delta}(A + B_6)$ and hence (iii). This concludes the proof of the fact that (vi) implies (iii).

We have now completed the proof of the following equivalence of statements:

$$L^2 \text{ boundedness of } T \iff \text{(iii)} \iff \text{(iv)} \iff \text{(v)} \iff \text{(vi)} \tag{8.82}$$

and we have established that

$$\|T\|_{L^2 \to L^2} \approx A + B_3 \approx A + B_4 \approx A + B_5 \approx A + B_6.$$

Step 6. (i) $\Longrightarrow$ (ii).

We show that the quantity B_2 is bounded by a multiple of $A+B_1$; if so, then so would do the quantity $A+B_2$. We set

$$I_{\varepsilon,N}(x)=\int_{\varepsilon<|x-y|<N}K(x,y)dy \quad \text{and} \quad I^t_{\varepsilon,N}(x)=\int_{\varepsilon<|x-y|<N}K^t(x,y)dy.$$

It suffices to show that there is a constant C_n such that for any $x_1\in\mathbb{R}^n$, we have

$$\sup_{\varepsilon,N}\left[\frac{1}{N^n}\int_{|x-x_1|<N/2}|I_{\varepsilon,N}(x)|^2dx\right]^{1/2}\leqslant C_n(A+B_1). \tag{8.83}$$

If (8.83) holds, then we can cover the ball $B(x_0,N)$ by finitely many balls $B(x_1,N/2)$ and thus deduce

$$\sup_{x_0\in\mathbb{R}^n}\sup_{\varepsilon,N}\left[\frac{1}{N^n}\int_{|x-x_0|<N}|I_{\varepsilon,N}(x)|^2dx\right]^{1/2}\leqslant C'_n(A+B_1) \tag{8.84}$$

with a larger constant C'_n in place of C_n.

We estimate the expression on the left in (8.83) by $I+II$, where

$$I=\sup_{\varepsilon,N}\left[\frac{1}{N^n}\int_{|x-x_1|<N/2}|I_{\varepsilon,N}(x)-T^{(\varepsilon)}(\chi_{B(x_1,N)})(x)|^2dx\right]^{1/2},$$

$$II=\sup_{\varepsilon,N}\left[\frac{1}{N^n}\int_{|x-x_1|<N}|T^{(\varepsilon)}(\chi_{B(x_1,N)})(x)|^2dx\right]^{1/2}.$$

By hypothesis, we have that II is bounded by B_1. Additionally, for $|x-x_1|<N/2$ we have by (8.1)

$$\begin{aligned}\left|I_{\varepsilon,N}(x)-T^{(\varepsilon)}(\chi_{B(x_1,N)})(x)\right|=&\left|\int_{\varepsilon<|x-y|<N}K(x,y)dy-\int_{\substack{\varepsilon<|x-y|\\|x_1-y|<N}}K(x,y)dy\right|\\ \leqslant&\int_{\left\{\substack{\varepsilon<|x-y|<N\\|x_1-y|\geqslant N}\right\}\cup\left\{\substack{|x-y|\geqslant N\\|x_1-y|<N}\right\}}|K(x,y)|dy\\ \leqslant&\int_{N/2\leqslant|x-y|\leqslant 3N/2}\frac{A}{|x-y|^n}dy\\ =&A\omega_{n-1}\ln 3.\end{aligned}$$

Thus, I is bounded by $\omega_{n-1}(\ln 3)A2^{-n/2}$. Combining the estimates for I and II yields the proof of (8.83) and hence of (8.84). Similarly, we can prove that

$$\sup_{x_0\in\mathbb{R}^n}\sup_{\varepsilon,N}\left[\frac{1}{N^n}\int_{|x-x_0|<N}|I^t_{\varepsilon,N}(x)|^2dx\right]^{1/2}\leqslant C'_n(A+B_1),$$

which together with (8.84) implies that $B_2\leqslant 2C'_n(A+B_1)$.

We now consider the following condition analogous to (iii):

(iii)′ $$B'_3=\sup_{\varphi}\sup_{x_0\in\mathbb{R}^n}\sup_{\substack{\varepsilon>0\\R>0}}R^{n/2}\left[\|T^{(\varepsilon)}(\tau^{x_0}\varphi_R)\|_2+\|(T^{(\varepsilon)})^t(\tau^{x_0}\varphi_R)\|_2\right]<\infty,$$

where the first supremum is taken over all normalized bumps φ. We continue the proof by showing that this condition is a consequence of (ii).

Step 7. (ii) $\Longrightarrow$ (iii)′.

More precisely, we prove that $B_3' \leqslant C_{n,\delta}(A+B_2)$. To prove (iii)′, fix a normalized bump φ, a point $x_0 \in \mathbb{R}^n$, and $R>0$. Additionally, fix $x \in \mathbb{R}^n$ with $|x-x_0| \leqslant 20R$. Then, we have

$$T^{(\varepsilon)}(\tau^{x_0}\varphi_R)(x) = \int_{\varepsilon<|x-y|\leqslant 30R} K(x,y)\tau^{x_0}\varphi_R(y)dy = U_1(x)+U_2(x),$$

where

$$\begin{aligned} U_1(x) &= \int_{\varepsilon<|x-y|\leqslant 30R} K(x,y)(\tau^{x_0}\varphi_R(y)-\tau^{x_0}\varphi_R(x))dy,\\ U_2(x) &= \tau^{x_0}\varphi_R(x)\int_{\varepsilon<|x-y|\leqslant 30R} K(x,y)dy. \end{aligned}$$

However, we have that

$$\begin{aligned} |\tau^{x_0}\varphi_R(y)-\tau^{x_0}\varphi_R(x)| &= R^{-n}\left[\varphi((y-x_0)/R)-\varphi((x-x_0)/R)\right]\\ &\leqslant R^{-n}\|\nabla\varphi\|_\infty R^{-1}|x-y|\\ &\leqslant C_n R^{-n-1}|x-y|. \end{aligned}$$

Thus, we obtain

$$|U_1(x)| \leqslant C_n R^{-n-1}A\int_{\varepsilon<|x-y|\leqslant 30R}\frac{1}{|x-y|^{n-1}}dy \leqslant C_n A R^{-n}$$

on $B(x_0,20R)$; hence,

$$\|U_1\|_{L^2(B(x_0,20R))} \leqslant C_n A R^{-n/2}.$$

Condition (ii) gives that

$$\|U_2\|_{L^2(B(x_0,20R))} \leqslant R^{-n}\|I_{\varepsilon,30R}\|_{L^2(B(x_0,30R))} \leqslant B_2(30R)^{n/2}R^{-n}.$$

Combining these two, we obtain

$$\|T^{(\varepsilon)}(\tau^{x_0}\varphi_R)\|_{L^2(B(x_0,20R))} \leqslant C_n(A+B_2)R^{-n/2} \tag{8.85}$$

and likewise for $(T^{(\varepsilon)})^t$. It follows from Lemma 8.28 that

$$\|T^{(\varepsilon)}(\tau^{x_0}\varphi_R)\|_{L^2(B(x_0,20R)^c)} \leqslant C_{n,\delta}AR^{-n/2},$$

which combined with (8.85) gives condition (iii)′ with constant $B_3' \leqslant C_{n,\delta}(A+B_2)$. This concludes the proof that condition (ii) implies (iii)′.

Step 8. (iii)′ $\Longrightarrow$ $T^{(\varepsilon)}: L^2 \to L^2$ uniformly in $\varepsilon>0$.

For $\varepsilon>0$, we introduce the smooth truncation $T_\zeta^{(\varepsilon)}$ of T by setting

$$T_\zeta^{(\varepsilon)}f(x) = \int_{\mathbb{R}^n} K(x,y)\zeta\left(\frac{x-y}{\varepsilon}\right)f(y)dy,$$

where $\zeta(x)\in[0,1]$ is a smooth function that is equal to 1 for $|x|\geqslant 1$ and vanishes for $|x| \leqslant 1/2$. We observe that

$$\begin{aligned}
|T_\zeta^{(\varepsilon)} f(x) - T^{(\varepsilon)} f(x)| &= \left| \int_{\mathbb{R}^n} K(x,y) \left(\zeta\left(\frac{x-y}{\varepsilon}\right) - \chi_{|x-y|\geqslant\varepsilon} \right) f(y) dy \right| \\
&\leqslant \int_{\varepsilon/2\leqslant|x-y|\leqslant\varepsilon} |K(x,y)||f(y)|dy \\
&\leqslant A \int_{\varepsilon/2\leqslant|x-y|\leqslant\varepsilon} \frac{1}{|x-y|^n}|f(y)|dy \\
&\leqslant A2^n \frac{1}{\varepsilon^n} \int_{|x-y|\leqslant\varepsilon} |f(y)|dy \\
&\leqslant C_n AMf(x),
\end{aligned} \tag{8.86}$$

thus, the uniform boundedness of $T^{(\varepsilon)}$ on L^2 is equivalent to the uniform boundedness of $T_\zeta^{(\varepsilon)}$. In view of Exercise 8.1, the kernels of the operators $T_\zeta^{(\varepsilon)}$ lie in $\mathrm{SK}(\delta, cA)$ uniformly in $\varepsilon > 0$ (for some constant c), since $\delta \leqslant 1$. Moreover, because of (8.86), we see that the operators $T_\zeta^{(\varepsilon)}$ satisfy (iii)′ with constant $C_n A + B_3'$.

A careful examination of the proof of the implications

$$\text{(iii)} \implies \text{(iv)} \implies L^2 \text{ boundedness of } T$$

reveals that all the estimates obtained depend only on the constants B_3, B_4 and A but not on the specific operator T. Therefore, these estimates are valid for the operators $T_\zeta^{(\varepsilon)}$ that satisfy condition (iii)′. This gives the uniform boundedness of $T_\zeta^{(\varepsilon)}$ on $L^2(\mathbb{R}^n)$ with bounds at most a constant multiple of $A + B_3'$. The same conclusion also holds for the operators $T^{(\varepsilon)}$.

Step 9. $T^{(\varepsilon)} : L^2 \to L^2$ uniformly in $\varepsilon > 0 \Longrightarrow$ (i).

This implication holds trivially.

We have now established the equivalence of the following statements:

$$\text{(i)} \iff \text{(ii)} \iff \text{(iii)}' \iff T^{(\varepsilon)} : L^2 \to L^2 \text{ uniformly in } \varepsilon > 0, \tag{8.87}$$

so that

$$A + B_1 \approx A + B_2 \approx A + B_3' \approx \sup_{\varepsilon>0} \|T^{(\varepsilon)}\|_{L^2\to L^2}.$$

Finally, it remains to link the sets of equivalent conditions (8.82) and (8.87). We do this by proving the equivalence of (iii) and (iii)′.

Step 10. (iii) $\iff$ (iii)′.

We will prove this by the following steps:

$$\begin{aligned}
\text{(iii)}' &\implies \text{(iii)} \implies T : L^2 \to L^2 \\
&\implies T^{(\varepsilon)} : L^2 \to L^2 \text{ uniformly in } \varepsilon > 0 \implies \text{(iii)}'.
\end{aligned}$$

We first observe that (iii)′ implies (iii). Indeed, using duality and (8.49), we obtain

$$\begin{aligned}\|T(\tau^{x_0}\varphi_R)\|_2 &= \sup_{\substack{h\in\mathscr{S}\\ \|h\|_2\leqslant 1}} \left|\int_{\mathbb{R}^n} T(\tau^{x_0}\varphi_R)(x)h(x)dx\right| \\ &= \sup_{\substack{h\in\mathscr{S}\\ \|h\|_2\leqslant 1}} \lim_{j\to\infty} \left|\int_{\mathbb{R}^n} T^{(\varepsilon_j)}(\tau^{x_0}\varphi_R)(x)h(x)dx\right| \\ &\leqslant \sup_{\substack{h\in\mathscr{S}\\ \|h\|_2\leqslant 1}} \limsup_{j\to\infty} \|T^{(\varepsilon_j)}(\tau^{x_0}\varphi_R)\|_2\|h\|_2 \\ &\leqslant B_3' R^{-n/2},\end{aligned}$$

which gives $B_3 \leqslant B_3'$.

We have shown that (iii) implies the L^2 boundedness of T. However, in view of Corollary 8.21, the boundedness of T on L^2 implies the boundedness of $T^{(\varepsilon)}$ on L^2 uniformly in $\varepsilon > 0$, which implies (iii)′. Moreover, B_3' is bounded by a constant multiple of $A + B_3$.

This completes the proof of the equivalence of the six statements (i)–(vi) in such a way that

$$\|T\|_{L^2\to L^2} \approx A + B_j$$

for all $j \in \{1,2,3,4,5,6\}$. The proof of the theorem is now complete. ❑

Remark 8.29. Suppose that condition (8.49) is removed from the hypothesis of Theorem 8.26. Then the given proof of Theorem 8.26 actually shows that (i) and (ii) are equivalent to each other and to the statement that the $T^{(\varepsilon)}$'s have bounded extensions on $L^2(\mathbb{R}^n)$ that satisfy

$$\sup_{\varepsilon>0} \|T^{(\varepsilon)}\|_{L^2\to L^2} < \infty.$$

Additionally, without hypothesis (8.49), the proof of Theorem 8.26 also shows that conditions (iii), (iv), (v) and (vi) are equivalent to each other and to the statement that T has an extension that maps $L^2(\mathbb{R}^n)$ to itself.

Exercises

Exercise 8.1. [Gra14b, Exercise 4.1.3] Let $\varphi(x)$ be a smooth radial function that is equal to 1 when $|x| \geqslant 1$ and vanishes when $|x| \leqslant 1/2$. Let $0 < \delta \leqslant 1$. Show that there is a constant $c > 0$ that depends only on n, φ, and δ such that if $K \in \mathrm{SK}(\delta, A)$, then all the smooth truncations $K_\varphi^{(\varepsilon)} = K(x,y)\varphi(\frac{x-y}{\varepsilon})$ lie in $\mathrm{SK}(\delta, cA)$ uniformly in $\varepsilon > 0$.

Exercise 8.2. [Gra14b, Exercise 4.2.1] Let $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ be a continuous linear operator whose Schwartz kernel coincides with a function

$K(x,y)$ on $(\mathbb{R}^n \times \mathbb{R}^n) \setminus \{(x,x) : x \in \mathbb{R}^n\}$. Suppose that the function $K(x,y)$ satisfies

$$\sup_{R>0} \int_{R \leqslant |x-y| \leqslant 2R} |K(x,y)| dy = A < \infty.$$

(i) Show that the previous condition is equivalent to

$$\sup_{R>0} \frac{1}{R} \int_{|x-y| \leqslant R} |x-y||K(x,y)| dy = A' < \infty$$

by proving that $A'/2 \leqslant A \leqslant 2A'$.

(ii) For $\varepsilon > 0$, let $T^{(\varepsilon)}$ be the truncated linear operators with kernels $K^{(\varepsilon)}(x,y) = K(x,y)\chi_{|x-y|>\varepsilon}$. Show that the integral defining $T^{(\varepsilon)} f$ converges absolutely for $f \in \mathscr{S}$.

Hint (ii) Consider the annuli $\varepsilon 2^j \leqslant |x| \leqslant \varepsilon 2^{j+1}$ for $j \geqslant 0$.

Exercise 8.3. [Gra14b, Exercise 4.3.1] Let T be a continuous linear operator from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ and $f \in \mathscr{S}(\mathbb{R}^n)$. Let P_t be as in (8.63).

(i) Show that $P_t(f)$ converges to f in $\mathscr{S}(\mathbb{R}^n)$ ad $t \to 0$.
(ii) Conclude that $TP_t(f) \to T(f)$ in $\mathscr{S}'(\mathbb{R}^n)$ as $t \to 0$.
(iii) Conclude that $P_t T P_t(f) \to T(f)$ in $\mathscr{S}'(\mathbb{R}^n)$ as $t \to 0$.
(iv) Observe that (i)-(iii) are also valid if P_t is replaced by P_t^2.

Hint (i) Use that $g_k \to g$ in $\mathscr{S}$ iff $\widehat{g_k} \to \widehat{g}$ in $\mathscr{S}$.

Exercise 8.4. [Gra14b, Exercise 4.3.2] Let T and P_t be as in Exercise 8.3 and let f be a Schwartz function whose Fourier transform vanishes in a neighborhood of the origin.

(i) Show that $P_t(f)$ converges to 0 in $\mathscr{S}(\mathbb{R}^n)$ ad $t \to \infty$.
(ii) Conclude that $TP_t(f) \to 0$ in $\mathscr{S}'(\mathbb{R}^n)$ as $t \to \infty$.
(iii) Conclude that $P_t T P_t(f) \to 0$ in $\dot{\mathscr{S}}'(\mathbb{R}^n)$ as $t \to \infty$.
(iv) Observe that (i)-(iii) are also valid if P_t is replaced by P_t^2.

Hint (i) Use the hint in Exercise 8.3 and the observation that $|\widehat{\Phi}(t\xi)\widehat{f}(\xi)| \leqslant C(1+tc_0)^{-1}|\widehat{f}(\xi)|$ if $\widehat{f}$ is support outside the ball $B(0,c_0)$. (iii) Pair with a function $g \in \dot{\mathscr{S}}(\mathbb{R}^n)$ and use (i) and the fact that all Schwartz seminorms of $P_t(g)$ are bounded uniformly in $t > 0$ (iff all Schwartz seminorms of $\widehat{P_t(g)}$ are bounded uniformly in $t > 0$).

9

Besov and Triebel-Lizorkin Spaces

§ 9.1 The smooth dyadic decomposition

In this section, we will introduce smooth 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 perform some operations such as differentiation and multiplier. 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 $0 < R_1 < R_2$.

Now, we give the fundamental Bernstein inequalities.

Proposition 9.1 (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, we have*

$$\forall f \in L^p(\mathbb{R}^n) \text{ with } \operatorname{supp} \widehat{f} \subset \lambda B \Longrightarrow \sup_{|\alpha|=k} \|\partial^\alpha f\|_q \leqslant C^{k+1} \lambda^{k+n(\frac{1}{p}-\frac{1}{q})} \|f\|_p,$$

$$\forall f \in L^p(\mathbb{R}^n) \text{ with } \operatorname{supp} \widehat{f} \subset \lambda \mathbb{A} \Longrightarrow$$
$$C^{-k-1} \lambda^k \|f\|_p \leqslant \sup_{|\alpha|=k} \|\partial^\alpha f\|_p \leqslant C^{k+1} \lambda^k \|f\|_p.$$

Proof. Since $\widehat{f} \in \mathscr{S}'$ has a compact support, we have $\widehat{f} \in \mathscr{E}'$ in view of the arguments below Definition 3.38. Then, it follows from Theorem 3.45 that $\widehat{\widehat{f}} \in \mathscr{C}^\infty$ which implies that f coincides with a $\mathscr{C}^\infty$ function by Fourier inversion in $\mathscr{S}'$.

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 $\widehat{f}(\xi) = \phi_\lambda(\xi)\widehat{f}(\xi)$ pointwise, we have

$$\partial^\alpha f = \partial^\alpha g_\lambda * f \quad \text{with} \quad g_\lambda = (\phi_\lambda)^\vee.$$

Thus, $g_\lambda(x) = \lambda^n \check{\phi}(\lambda x) = \lambda^n g(\lambda x)$, where we denote $g := g_1$.

Applying Young's inequality with $\frac{1}{r} := 1 - \frac{1}{p} + \frac{1}{q}$, we obtain

$$\begin{aligned}\|\partial^\alpha f\|_q =& \|\partial^\alpha g_\lambda * f\|_q \leqslant \|\partial^\alpha g_\lambda\|_r \|f\|_p \\ =& \lambda^{n+k} \|(\partial^\alpha g)(\lambda x)\|_r \|f\|_p = \lambda^{k+n/r'} \|\partial^\alpha g\|_r \|f\|_p\end{aligned}$$

$$=\lambda^{k+n(\frac{1}{p}-\frac{1}{q})}\|\partial^\alpha g\|_r\|f\|_p.$$

The first assertion follows from

$$\begin{aligned}
\|\partial^\alpha g\|_r \leqslant &\|\partial^\alpha g\|_\infty + \|\partial^\alpha g\|_1\\
\leqslant&\|\partial^\alpha g\|_\infty + \int_{\mathbb{R}^n} |\partial^\alpha g|(1+|x|^2)^n\frac{1}{(1+|x|^2)^n}dx\\
\leqslant&\|\partial^\alpha g\|_\infty + \|(1+|x|^2)^n\partial^\alpha g\|_\infty \int_{\mathbb{R}^n}\frac{1}{(1+|x|^2)^n}dx\\
\leqslant&C_n\|(1+|x|^2)^n\partial^\alpha g\|_\infty = C_n\|\mathscr{F}^{-1}\mathscr{F}((1+|x|^2)^n\partial^\alpha g)\|_\infty\\
\leqslant&C_n\|\mathscr{F}((1+|x|^2)^n\partial^\alpha g)\|_1 = C_n\|(1-\Delta)^n((i\xi)^\alpha\phi(\xi))\|_1\\
=&C_n\left\|\sum_{j=0}^n C_n^j(-1)^j\Delta^j(\xi^\alpha\phi(\xi))\right\|_1 \leqslant C_n\sum_{j=0}^n C_n^j\|\Delta^j(\xi^\alpha\phi(\xi))\|_1\\
\leqslant&C_n\sup_{0\leqslant|\beta|\leqslant|\alpha|,0\leqslant|\sigma|\leqslant 2n-|\beta|}\|\partial^\beta(\xi^\alpha)\partial^\sigma\phi\|_1\\
\leqslant&C_n\sup_{0\leqslant|\beta|\leqslant|\alpha|,0\leqslant|\sigma|\leqslant 2n-|\beta|}\|\xi^\beta\partial^\sigma\phi\|_1\\
\leqslant&C_nC^k\sup_{0\leqslant|\sigma|\leqslant 2n}\|\partial^\sigma\phi\|_1 \quad \text{(since } \phi \text{ is compactly supported)}\\
\leqslant&C_n^{k+1}.
\end{aligned}$$

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 $\widehat{f}=\tilde{\phi}\widehat{f}$ for $\lambda=1$, we deduce that there exists a family of integers $(a_\alpha)_{\alpha\in\mathbb{N}_0^n}$ such that

$$f=\sum_{|\alpha|=k}h_\alpha*\partial^\alpha f,\quad h_\alpha:=(2\pi)^{-n/2}a_\alpha\mathscr{F}^{-1}\left((-i\xi)^\alpha|\xi|^{-2k}\tilde{\phi}(\xi)\right)\in\mathscr{S}\subset L^1.$$

For $\lambda>0$, from $\operatorname{supp}\widehat{f}\subset\lambda\mathbb{A}$ we have

$$\begin{aligned}
\widehat{f}(\xi)=&\sum_{|\alpha|=k}a_\alpha\frac{(-i\xi)^\alpha}{|\xi|^{2k}}\tilde{\phi}(\xi/\lambda)(i\xi)^\alpha\widehat{f}(\xi)\\
=&\lambda^{-k}\sum_{|\alpha|=k}a_\alpha\frac{(-i\xi/\lambda)^\alpha}{|\xi/\lambda|^{2k}}\tilde{\phi}(\xi/\lambda)(i\xi)^\alpha\widehat{f}(\xi)\\
=&\lambda^{-k}\sum_{|\alpha|=k}(2\pi)^{n/2}a_\alpha\widehat{h_\alpha}(\xi/\lambda)(i\xi)^\alpha\widehat{f}(\xi)\\
=&\lambda^{-k}\sum_{|\alpha|=k}(2\pi)^{n/2}a_\alpha\lambda^n\widehat{h_\alpha(\lambda\cdot)}(\xi)\widehat{\partial^\alpha f}(\xi),
\end{aligned}$$

which implies that

$$f = \lambda^{-k} \sum_{|\alpha|=k} \lambda^n h_\alpha(\lambda\cdot) * \partial^\alpha f.$$

Then, by Young's inequality, we obtain

$$\|f\|_p \leqslant \lambda^{-k} \sum_{|\alpha|=k} \|h_\alpha\|_1 \|\partial^\alpha f\|_p \leqslant C^{k+1}\lambda^{-k} \sum_{|\alpha|=k} \|\partial^\alpha f\|_p,$$

by a similar argument for $\|h_\alpha\|_1$ as in $\|\partial^\alpha g\|_r$, and the result follows from the first inequality. ❑

Remark 9.2. 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 9.3. *Let $\mathbb{A}$ be an annulus, $m \in \mathbb{R}$, and $k > n/2$ be an integer. 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 $\operatorname{supp} \widehat{f} \subset \lambda\mathbb{A}$,

$$\|\sigma(D)f\|_p \leqslant C\lambda^m \|f\|_p, \quad \text{with } \sigma(D)f := \left(\sigma\widehat{f}\right)^\vee.$$

Proof. It is clear that

$$\|\sigma(\xi)\chi_{\lambda\mathbb{A}}(\xi)\|_2 = C_n \lambda^{m+n/2},$$
$$\|\partial^\alpha(\sigma(\xi)\chi_{\lambda\mathbb{A}}(\xi))\|_2 = C_n \lambda^{m-k+n/2}, \quad \text{and } |\alpha| = k.$$

Thus, we have by the Bernstein multiplier theorem for $p \in [1,\infty]$

$$\|\sigma\chi_{\lambda\mathbb{A}}\|_{\mathcal{M}_p} \leqslant C_n \left(\lambda^{m+n/2}\right)^{1-n/2k} \left(\lambda^{m-k+n/2}\right)^{n/2k} = C_n\lambda^m,$$

which implies the desired result. ❑

Let $\alpha \in (1, \sqrt{2})$ and $\psi : \mathbb{R}^n \to [0,1]$ be a real radial smooth bump function, e.g.,

$$\psi(\xi) = \begin{cases} 1, & |\xi| \leqslant \alpha^{-1}, \\ \text{smooth}, & \alpha^{-1} < |\xi| < \alpha, \\ 0, & |\xi| \geqslant \alpha. \end{cases} \tag{9.1}$$

Let $\varphi(\xi)$ be the function

$$\varphi(\xi) := \psi(\xi/2) - \psi(\xi). \tag{9.2}$$

Thus, φ is a bump function supported on the annulus

$$\mathbb{A} = \left\{\xi : \alpha^{-1} \leqslant |\xi| \leqslant 2\alpha\right\}. \tag{9.3}$$

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:

$$\begin{cases} \psi_k(\xi) = \psi(2^{-k}\xi), & k \in \mathbb{Z}, \\ \varphi_k(\xi) = \varphi(2^{-k}\xi) = \psi_{k+1}(\xi) - \psi_k(\xi), & k \in \mathbb{Z}. \end{cases} \tag{9.4}$$

Since $\operatorname{supp}\varphi \subset \mathbb{A}$, we have

$$\begin{aligned} &\operatorname{supp}\varphi_k \subset 2^k\mathbb{A} := \left\{\xi : 2^k\alpha^{-1} \leqslant |\xi| \leqslant 2^{k+1}\alpha\right\}, \quad k \in \mathbb{Z}, \\ &\operatorname{supp}\psi_k \subset \left\{\xi : |\xi| \leqslant 2^k\alpha\right\}, \quad k \in \mathbb{Z}. \end{aligned} \tag{9.5}$$

We now define the ***k-th homogeneous dyadic blocks*** $\dot{\Delta}_k$ and the ***homogeneous low-frequency cut-off operators*** $\dot{S}_k$ by

$$\dot{\Delta}_k f = \mathscr{F}^{-1}\varphi_k\mathscr{F}f, \quad \dot{S}_k f = \mathscr{F}^{-1}\psi_k\mathscr{F}f = \sum_{j\leqslant k-1}\dot{\Delta}_j f, \quad k \in \mathbb{Z}. \tag{9.6}$$

Informally, $\dot{\Delta}_k$ is a frequency projection[1] to the annulus

$$\left\{\xi : 2^k\alpha^{-1} \leqslant |\xi| \leqslant 2^{k+1}\alpha\right\},$$

while $\dot{S}_k$ is a frequency projection to the ball $\{\xi : |\xi| \leqslant 2^k\alpha\}$. The nonhomogeneous dyadic blocks Δ_k are defined by

$$\Delta_k f = 0 \text{ if } k \leqslant -2, \quad \Delta_{-1} f = \dot{S}_0 f, \quad \text{and } \Delta_k f = \dot{\Delta}_k f \text{ if } k \geqslant 0.$$

The nonhomogeneous low-frequency cut-off operator S_k is defined by

$$S_k f = \sum_{j\leqslant k-1} \Delta_j f.$$

Obviously, $S_k f = 0$ if $k \leqslant -1$, and $S_k f = \dot{S}_k f$ if $k \geqslant 0$.

Observe that $\dot{S}_{k+1} = \dot{S}_k + \dot{\Delta}_k$ from (9.4). Additionally, 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

[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 cut-off. However, the operator $\dot{\Delta}_k\dot{\Delta}_k$ is of the same form as $\dot{\Delta}_k$, and similarly for $\dot{S}_k$, and so it is still quite reasonable to think of these operators as (smoothed out) projection operators.

the series, we can thus write the following (formal) Littlewood-Paley (or dyadic) decomposition[2]

$$\mathrm{Id} = \sum_{k\in\mathbb{Z}} \dot{\Delta}_k \quad \text{and} \quad \mathrm{Id} = \sum_{k\in\mathbb{Z}} \Delta_k. \tag{9.7}$$

The homogeneous decomposition takes a single function and writes it as a superposition of a countably infinite family of functions $\dot{\Delta}_k f$, each one of which has a frequency of magnitude of approximately 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 9.4. *Let $f \in \mathscr{S}'(\mathbb{R}^n)$, then $f = \lim\limits_{k\to+\infty} S_k f$ 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 $g = \lim\limits_{k\to+\infty} S_k g$ in $\mathscr{S}(\mathbb{R}^n)$. Because the Fourier transform is an automorphism of $\mathscr{S}(\mathbb{R}^n)$, we can alternatively prove that $\psi(2^{-k}\cdot)\widehat{g}$ tends to $\widehat{g}$ in $\mathscr{S}(\mathbb{R}^n)$. This can easily be verified, so we left it to the interested reader. ❑

We now state another result of convergence.

Proposition 9.5. *Let $\{u_j\}_{j\in\mathbb{N}}$ be a sequence of bounded functions such that $\operatorname{supp}\widehat{u_j} \subset 2^j\widetilde{\mathbb{A}}$, where $\widetilde{\mathbb{A}}$ is a given annulus. Assume that for some $N \in \mathbb{N}$*

$$\|u_j\|_\infty \leqslant C2^{jN}, \quad \forall j \in \mathbb{N}, \tag{9.8}$$

then the series $\sum\limits_j u_j$ converges in $\mathscr{S}'$.

Proof. Taking $\phi(\xi) \in \mathscr{D}(\mathbb{R}^n \setminus \{0\})$ with value 1 near $\widetilde{\mathbb{A}}$, we have near $\widetilde{\mathbb{A}}$

[2]Actually, this decomposition works for just about any locally integrable function that 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, then this formula fails. For instance, if $f \equiv 1$, then all the projections $\dot{\Delta}_k f$ vanish because $\dot{\Delta}_k 1 = \int e^{ix\xi}\varphi_k(\xi)(2\pi)^{n/2}\delta_0(\xi)đ\xi = \varphi_k(0) = \varphi(0) = 0$.

and any $k \in \mathbb{N}$,

$$\widehat{u_0} = \phi(\xi)\widehat{u_0}(\xi) = \sum_{|\alpha|=k} a_\alpha \frac{(-i\xi)^\alpha}{|\xi|^{2k}} \phi(\xi)(i\xi)^\alpha \widehat{u_0}(\xi),$$

namely,

$$u_0 = \sum_{|\alpha|=k} g_\alpha * \partial^\alpha u_0, \quad g_\alpha = (2\pi)^{-n/2} a_\alpha \left(\frac{(-i\xi)^\alpha}{|\xi|^{2k}} \phi(\xi) \right)^\vee.$$

Similarly, on each $2^j \widetilde{\mathbb{A}}$, it holds

$$\widehat{u_j} = \sum_{|\alpha|=k} a_\alpha 2^{-jk} \frac{(-i\xi/2^j)^\alpha}{|\xi/2^j|^{2k}} \phi(\xi/2^j)(i\xi)^\alpha \widehat{u_j}(\xi),$$

that is,

$$u_j = 2^{-jk} \sum_{|\alpha|=k} 2^{jn} g_\alpha(2^j \cdot) * \partial^\alpha u_j. \tag{9.9}$$

For any $f \in \mathscr{S}$, we obtain from Definition 3.32 and Definition 3.27

$$\begin{aligned} |\langle u_j, f\rangle| =& 2^{-jk} \left| \sum_{|\alpha|=k} \langle u_j, 2^{jn} g_\alpha(-2^j \cdot) * (-\partial)^\alpha f \rangle \right| \\ \leqslant& 2^{-jk} \sum_{|\alpha|=k} \|u_j\|_\infty \|2^{jn} g_\alpha(-2^j \cdot) * \partial^\alpha f\|_1 \\ \leqslant& C 2^{-jk} \sum_{|\alpha|=k} 2^{jN} \|\partial^\alpha f\|_1. \end{aligned}$$

It is clear that

$$\begin{aligned} \|\partial^\alpha f\|_1 \leqslant& \int_{\mathbb{R}^n} \frac{dx}{(1+|x|)^{n+1}} \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1} |\partial^\alpha f(x)| \\ \leqslant& C \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1} |\partial^\alpha f(x)|. \end{aligned}$$

Taking $k = N+1$, we have

$$\left| \sum_{j\in\mathbb{N}} \langle u_j, f\rangle \right| \leqslant C \sum_{|\alpha|=N+1} \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1} |\partial^\alpha f(x)|,$$

which implies that the series converges in $\mathscr{S}'$ by the equivalent conditions of $\mathscr{S}'$ in Theorem 3.23. Thus, the convergent series

$$\langle u, f\rangle := \lim_{j\to\infty} \sum_{j'\leqslant j} \langle u_{j'}, f\rangle$$

defines a tempered distribution. ❑

We have some identities as follows:

Proposition 9.6. *Let $\alpha \in (1, \sqrt{2})$, k, $l \in \mathbb{Z}$, and $\dot{\Delta}_k, \dot{S}_k$ be defined as in*

(9.6). *For any $f, g \in \mathscr{S}'(\mathbb{R}^n)$, we have the following properties:*

$$\dot{S}_k\dot{\Delta}_{k+l}f \equiv 0, \text{if } l \geqslant 1, \tag{9.10}$$

$$\dot{\Delta}_k\dot{\Delta}_l f \equiv 0, \text{if } |k-l| \geqslant 2, \tag{9.11}$$

$$\dot{\Delta}_k(\dot{S}_{l-1}f\dot{\Delta}_l g) \equiv 0, \text{if } k-l \leqslant -2-\log_2\frac{\alpha^2}{2-\alpha^2}, \text{ or } k-l \geqslant -1+\log_2 5\alpha^2. \tag{9.12}$$

In particular, taking $\alpha = \frac{9}{8}$, (9.12) becomes

$$\dot{\Delta}_k(\dot{S}_{l-1}f\dot{\Delta}_l g) \equiv 0, \quad \text{if } k-l \leqslant -3, \text{ or } k-l \geqslant 2. \tag{9.13}$$

Remark 9.7. In these properties, we need the condition $\alpha^2 < 2$, which is the reason that we require $\alpha < \sqrt{2}$ at the beginning of the section. **From now on, we always take $\alpha = \frac{9}{8}$ and use (9.13) instead of (9.12) for simplicity** since there are at most four nonzero terms for this choice.

When dealing with the Littlewood-Paley decomposition, it is convenient to introduce the functions

$$\tilde{\psi}(\xi) = \psi(\xi/2), \quad \tilde{\varphi}(\xi) = \varphi_{-1}(\xi) + \varphi_0(\xi) + \varphi_1(\xi) = \psi(\xi/4) - \psi(2\xi).$$

and 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 $\dot{S}_k = \tilde{S}_k\dot{S}_k$, and $\dot{\Delta}_k = \tilde{\triangle}_k\dot{\Delta}_k$ from Proposition 9.6.

By Young's inequality, we can easily prove the following crucial properties of the operators $\dot{\Delta}_k$ and $\dot{S}_k$:

Proposition 9.8 (Boundedness)**.** *For any $1 \leqslant p \leqslant \infty$ and $k \in \mathbb{Z}$, it holds*

$$\|\dot{\Delta}_k f\|_p \leqslant C\|f\|_p, \quad \|\dot{S}_k f\|_p \leqslant C\|f\|_p,$$

for some constant C independent of p.

We now study how the Littlewood-Paley pieces $\dot{\Delta}_k f$ (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 $\dot{\Delta}_k f$ relates to the L^p behavior of f. One can already see this when $p = 2$, in which case we have

$$\|f\|_2 \sim \left(\sum_{k\in\mathbb{Z}} \|\dot{\Delta}_k f\|_2^2\right)^{1/2}. \tag{9.14}$$

In fact, we square both sides and take Plancherel to obtain

$$\int_{\mathbb{R}^n} |\widehat{f}(\xi)|^2 d\xi \sim \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^n} |\varphi_k(\xi)|^2 |\widehat{f}(\xi)|^2 d\xi.$$

Observe that for each $\xi \neq 0$ there are only three values of $\varphi_k(\xi)$ that do not vanish. That is, for $|\xi| \in [\frac{8}{9}2^\ell, \frac{9}{4}2^\ell]$,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} |\varphi_k(\xi)|^2 =&\varphi_{\ell-1}^2(\xi) + \varphi_\ell^2(\xi) + \varphi_{\ell+1}^2(\xi)\\ =&(\varphi_{\ell-1}(\xi) + \varphi_\ell(\xi) + \varphi_{\ell+1}(\xi))^2\\ &- 2(\varphi_{\ell-1}(\xi)\varphi_\ell(\xi) + \varphi_{\ell-1}(\xi)\varphi_{\ell+1}(\xi) + \varphi_\ell(\xi)\varphi_{\ell+1}(\xi))\\ =&1 - 2(\varphi_{\ell-1}(\xi) + \varphi_{\ell+1}(\xi))\varphi_\ell(\xi)\\ =&1 - 2(1 - \varphi_\ell(\xi))\varphi_\ell(\xi)\\ =&1 - 2\varphi_\ell(\xi) + 2\varphi_\ell^2(\xi)\\ =&\frac{1}{2} + 2\left(\frac{1}{2} - \varphi_\ell(\xi)\right)^2,\end{aligned}$$

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 (9.14) is

$$\|f\|_2 \sim \left\| \left(\sum_{k\in\mathbb{Z}} |\dot{\Delta}_k f|^2 \right)^{1/2} \right\|_2, \tag{9.15}$$

which is different from (5.38). More generally, another version of the Littlewood-Paley square function theorem (Theorem 5.25) is valid:

Theorem 9.9 (Littlewood-Paley square function theorem, another version). *For any $1 < p < \infty$, we have*

$$\left\| \left(\sum_{k\in\mathbb{Z}} |\dot{\Delta}_k f|^2 \right)^{1/2} \right\|_p \sim \|f\|_p$$

with the implicit constant depending on p.

We omit the proof. One can read the proof in [Ste93, Page 267], or [Gra14a, Pages 339-343].

§ 9.2 Definitions and embeddings

The Littlewood-Paley decomposition is very useful. For example, we can define (independently of the choice of the initial function ψ) the following notations.

Definition 9.10. Let $s \in \mathbb{R}$, $1 \leqslant p, r \leqslant \infty$. For $f \in \mathscr{S}'(\mathbb{R}^n)$, we write

$$\|f\|_{\dot{B}^s_{p,r}} = \left(\sum_{k=-\infty}^{\infty} \left(2^{sk}\|\dot{\Delta}_k f\|_p\right)^r\right)^{\frac{1}{r}}, \tag{9.16}$$

$$\|f\|_{B^s_{p,r}} = \|S_0 f\|_p + \left(\sum_{k=0}^{\infty} \left(2^{sk}\|\Delta_k f\|_p\right)^r\right)^{\frac{1}{r}}. \tag{9.17}$$

For $r = \infty$, it corresponds to the usual ℓ^∞ norm.

Observe that (9.16) does not satisfy the condition of the norms, since we have $\dot{\Delta}_k P(x) = 0$ in $\mathscr{S}'$ for any $P \in \mathscr{P}$. In fact,

$$\dot{\Delta}_k P(x) = 0 \text{ in } \mathscr{S}' \iff \langle \dot{\Delta}_k P, g\rangle = 0, \ \forall g \in \mathscr{S}.$$

It follows from $0 \notin \operatorname{supp} \varphi_k$ for any $k \in \mathbb{Z}$ that for any $\alpha \in \mathbb{N}_0^n$

$$\begin{aligned}\int_{\mathbb{R}^n} x^\alpha \dot{\Delta}_k g(x)dx = \int_{\mathbb{R}^n} x^\alpha \left(\widehat{\dot{\Delta}_k g}\right)^\vee (x)dx = \int_{\mathbb{R}^n} e^{-ix\cdot 0} i^{|\alpha|} \left(\partial_\xi^\alpha \widehat{\dot{\Delta}_k g}\right)^\vee (x)dx \\ =(2\pi)^{n/2} i^{|\alpha|} \left[\partial_\xi^\alpha \widehat{\dot{\Delta}_k g}\right](0) \\ =(2\pi)^{n/2}(-1)^{|\alpha|}\left[\partial_\xi^\alpha(\varphi_k \widehat{g})\right](0) = 0.\end{aligned}$$

Thus, by the property of φ_k, we obtain

$$\int_{\mathbb{R}^n} (\dot{\Delta}_k x^\alpha) g(x)dx = 0.$$

Now, we can use $\dot{\mathscr{S}}(\mathbb{R}^n)$ to give the following definition.

Definition 9.11. Let $s \in \mathbb{R}$, $1 \leqslant p, r \leqslant \infty$. 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 9.12. Let $s \in \mathbb{R}$, $1 \leqslant p < \infty$, $1 \leqslant r \leqslant \infty$. We write

$$\|f\|_{\dot{F}^s_{p,r}} = \left\|\left(\sum_{k=-\infty}^{\infty} \left(2^{sk}|\dot{\Delta}_k f|\right)^r\right)^{\frac{1}{r}}\right\|_p, \quad \forall f \in \dot{\mathscr{S}}'(\mathbb{R}^n),$$

$$\|f\|_{F^s_{p,r}} = \|S_0 f\|_p + \left\| \left(\sum_{k=0}^{\infty} \left(2^{sk} |\Delta_k f| \right)^r \right)^{\frac{1}{r}} \right\|_p, \quad \forall f \in \mathscr{S}'(\mathbb{R}^n).$$

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 9.13. It is easy to see that the above quantities define a quasinorm 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 $p = \infty$ in the definition of Triebel-Lizorkin space because the L^∞ norm has to be replaced here by a more complicated Carleson measure.

Besov spaces and Triebel-Lizorkin spaces were constructed between the 1960s and 1980s. Recently, they have been 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 9.9, we immediately have the following relations involving Sobolev spaces and Triebel-Lizorkin spaces:

Theorem 9.14. *Let $s \in \mathbb{R}$ and $1 < p < \infty$. Then*

$$H^s_p = F^s_{p,2}, \quad \dot{H}^s_p = \dot{F}^s_{p,2}, \tag{9.18}$$

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.). However, 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 9.15. *Let X denote B or F. Then, we have the following embedding:*

$$\begin{aligned} X^s_{p,r_1} \hookrightarrow X^s_{p,r_2}, \quad \dot{X}^s_{p,r_1} &\hookrightarrow \dot{X}^s_{p,r_2}, \quad \text{if } r_1 \leqslant r_2, \\ X^{s+\varepsilon}_{p,r_1} &\hookrightarrow X^s_{p,r_2}, \quad \text{if } \varepsilon > 0, \\ B^s_{p,\min(p,r)} \hookrightarrow F^s_{p,r} &\hookrightarrow B^s_{p,\max(p,r)}, \quad \text{if } 1 \leqslant p < \infty, \end{aligned}$$

$$\dot{B}^s_{p,\min(p,r)} \hookrightarrow \dot{F}^s_{p,r} \hookrightarrow \dot{B}^s_{p,\max(p,r)}, \quad \text{if } 1 \leqslant p < \infty.$$

Proof. It is clear that the first one is valid because of $\ell^r \hookrightarrow \ell^{r+a}$ 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 $a_k = \|\Delta_k f\|_p$ or $a_k = |\Delta_k f|$, we can obtain

$$X^{s+\varepsilon}_{p,\infty} \hookrightarrow X^s_{p,r_2},$$

which yields the second result in view of the first one.

For the third and last one, we separate into two cases and denote $b_k = 2^{sk}|\dot{\Delta}_k f|$ and $j = 0$ for the third or $j = -\infty$ for the last one.

Case I: $r \leqslant p$. In this case, we have $\ell^r \hookrightarrow \ell^p$ 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,[a], we obtain

$$\left\|\left(\sum_{k=j}^{\infty} b_k^r\right)^{\frac{1}{r}}\right\|_p^r = \left\|\sum_{k=j}^{\infty} b_k^r\right\|_{\frac{p}{r}} \leqslant \sum_{k=j}^{\infty} \|b_k^r\|_{\frac{p}{r}} = \sum_{k=j}^{\infty} \|b_k\|_p^r,$$

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 $\ell^p \hookrightarrow \ell^r$ 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. We complete the proof. ❑

---

[a] 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 $f_j \geqslant 0$.

From Theorems 9.14 and 9.15, we can obtain the following corollary.

Corollary 9.16. *Let $s \in \mathbb{R}$. Then we have*

(i) *For $1 < p < \infty$, $B^s_{p,\min(p,2)} \hookrightarrow H^s_p \hookrightarrow B^s_{p,\max(p,2)}$ and $\dot{B}^s_{p,\min(p,2)} \hookrightarrow \dot{H}^s_p \hookrightarrow \dot{B}^s_{p,\max(p,2)}$. 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 $1 \leqslant p \leqslant \infty$, $B^s_{p,1} \hookrightarrow H^s_p \hookrightarrow B^s_{p,\infty}$ and $\dot{B}^s_{p,1} \hookrightarrow \dot{H}^s_p \hookrightarrow \dot{B}^s_{p,\infty}$.*

Proof. It obviously follows from Theorems 9.14 and 9.15 except the endpoint cases $p = 1$ or ∞ in (ii). For the proof of the endpoint cases, one can see [BL76a, Chapter 6]. ❑

Theorem 9.17. *Let X denote B or F. Then,*

(i) *$X^s_{p,r}$ and $\dot{X}^s_{p,r}$ are complete;*

(ii) *$\mathscr{S}(\mathbb{R}^n) \hookrightarrow X^s_{p,r} \hookrightarrow \mathscr{S}'(\mathbb{R}^n)$, $\dot{\mathscr{S}}(\mathbb{R}^n) \hookrightarrow \dot{X}^s_{p,r} \hookrightarrow \dot{\mathscr{S}}'(\mathbb{R}^n)$;*

(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 (cf. [Jaw77; Saw18]). 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 proven in the future. Let us first prove the second result. We divide the proofs into four steps.

Step 1: To prove $\mathscr{S} \hookrightarrow B^s_{p,\infty}$. In fact, for some integer $\sigma \geqslant \max(s, 0)$ and sufficiently large[a] $L \in \mathbb{N}_0$, we have for any $f \in \mathscr{S}$, from Propositions 9.1 and 9.8, that

$$\begin{aligned}
\|f\|_{B^s_{p,\infty}} =&\|S_0 f\|_p + \sup_{k\geqslant 0} 2^{sk}\|\Delta_k f\|_p \\
\leqslant& C\|f\|_p + \sup_{k\geqslant 0} 2^{sk}2^{-\sigma k}2^{\sigma k}\|\Delta_k f\|_p \\
\lesssim& \sum_{\alpha,\beta} |f|_{\alpha,\beta} + \sup_{k\geqslant 0} 2^{sk}2^{-\sigma k} \sup_{|\gamma|=\sigma} \|\partial^\gamma f\|_p \\
\lesssim& \sum_{\alpha,\beta} |f|_{\alpha,\beta} + \sup_{|\gamma|=\sigma} \|(1+|x|^2)^L \partial^\gamma f\|_\infty \lesssim \sum_{\alpha,\beta} |f|_{\alpha,\beta},
\end{aligned}$$

where $|f|_{\alpha,\beta}$ is one of the seminorm sequences of $\mathscr{S}$. Thus, we obtain the result.

Step 2: To prove $\mathscr{S} \hookrightarrow X^s_{p,r}$. From Step 1, we know $\mathscr{S} \hookrightarrow B^{s+\varepsilon}_{p,\infty}$ for any $\varepsilon > 0$. From Theorem 9.15, we obtain $B^{s+\varepsilon}_{p,\infty} \hookrightarrow B^s_{p,\min(p,r)} \hookrightarrow B^s_{p,r} \cap F^s_{p,r}$. Therefore, $\mathscr{S} \hookrightarrow X^s_{p,r}$.

Step 3: To prove $B^s_{p,\infty} \hookrightarrow \mathscr{S}'$. For simplicity, we temporarily denote $\Delta_{-1} \equiv 0$. For any $f \in B^s_{p,\infty}$ and $g \in \mathscr{S}$, we have, from Schwarz's inequality, Proposition 9.8 and the result in Step 1, that

$$\begin{aligned}
|\langle f,g\rangle| &= \left|\left\langle (S_0+\sum_{k=0}^{\infty}\Delta_k)f,(S_0+\sum_{l=0}^{\infty}\Delta_l)g\right\rangle\right| \\
&\leqslant |\langle S_0f,S_0g\rangle| + |\langle S_0f,\Delta_0g\rangle| + |\langle \Delta_0f,S_0g\rangle| \\
&\quad + \sum_{k=0}^{\infty}\sum_{l=-1}^{1} |\langle \Delta_kf,\Delta_{k+l}g\rangle| \\
&\lesssim \|f\|_p\|g\|_{p'} + \sum_{k=0}^{\infty}\sum_{l=-1}^{1} \|\Delta_kf\|_p\|\Delta_{k+l}g\|_{p'} \\
&\lesssim \|f\|_p\|g\|_{p'} + \sum_{k=0}^{\infty}\sum_{l=-1}^{1} 2^{sk}\|\Delta_kf\|_p 2^{-sk}\|\Delta_{k+l}g\|_{p'} \\
&\lesssim \|f\|_p\|g\|_{p'} + \sup_{k\geqslant 0} 2^{sk}\|\Delta_kf\|_p \sum_{k=0}^{\infty} 2^{-\varepsilon k}2^{(-s+\varepsilon)k}\|\Delta_kg\|_{p'} \\
&\lesssim \|f\|_{B^s_{p,\infty}}\|g\|_{B^{-s+\varepsilon}_{p',\infty}} \\
&\lesssim \|f\|_{B^s_{p,\infty}} \sum_{\alpha,\beta} |g|_{\alpha,\beta}.
\end{aligned}$$

Thus, we have proven the result.

Step 4: To prove $X^s_{p,r} \hookrightarrow \mathscr{S}'$. From Theorem 9.15, we have $X^s_{p,r} \hookrightarrow B^s_{p,\max(p,r)} \hookrightarrow B^s_{p,\infty} \hookrightarrow \mathscr{S}'$.

Finally, let us prove the completeness of $B^s_{p,r}$. The completeness of $F^s_{p,r}$ can be proved at a similar way. Let $\{f_l\}_1^\infty$ be a Cauchy sequence in $B^s_{p,r}$. So does it 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 $\{\Delta_k f_l\}_{l=1}^\infty$ is a Cauchy sequence in L^p. From the completeness of L^p, there is a $g_k \in L^p$ such that

$$\|\Delta_k f_l - g_k\|_p \to 0, \quad l \to \infty. \tag{9.19}$$

Since $L^p \hookrightarrow \mathscr{S}'$ and $\Delta_k f_l \to \Delta_k f$ as $l \to \infty$ in $\mathscr{S}'$, we obtain $g_k = \Delta_k f$. Hence, (9.19) implies

$$\|\Delta_k(f_l - f)\|_p \to 0, \quad l \to \infty.$$

which yields $\sup_{k\geqslant 0} 2^{(s+\varepsilon)k}\|\Delta_k(f_l-f)\|_p \to 0$ as $l \to \infty$ for any $\varepsilon > 0$.

Similarly, we have

$$\|S_0(f_l - f)\|_p \to 0, \quad l \to \infty.$$

Therefore,

$$\|f_l - f\|_{B^s_{p,r}} \lesssim \|f_l - f\|_{B^{s+\varepsilon}_{p,\infty}} \to 0, \quad l \to \infty.$$

Similarly, we can obtain the density statement in (iii). We omit the details. ❑

[a] It is enough to assume that $L > \frac{n}{2p}$. 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 $2pL > n$.

Theorem 9.18 (The embedding theorem). *Let $1 \leqslant p, p_1, r, r_1 \leqslant \infty$ and $s, s_1 \in \mathbb{R}$. Assume that $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$. The following conclusions hold*

$$B^s_{p,r} \hookrightarrow B^{s_1}_{p_1,r_1}, \quad \dot{B}^s_{p,r} \hookrightarrow \dot{B}^{s_1}_{p_1,r_1}, \quad \forall p \leqslant p_1 \text{ and } r \leqslant r_1;$$
$$F^s_{p,r} \hookrightarrow F^{s_1}_{p_1,r_1}, \quad \dot{F}^s_{p,r} \hookrightarrow \dot{F}^{s_1}_{p_1,r_1}, \quad \forall p < p_1 < \infty.$$

Proof. We only give the proof of the nonhomogeneous cases, and the homogeneous cases can be treated in a similar way.

Let us prove the first conclusion. From the Bernstein inequality in Proposition 9.1, we immediately have

$$\|\Delta_k f\|_{p_1} \lesssim 2^{kn(\frac{1}{p}-\frac{1}{p_1})}\|\Delta_k f\|_p, \quad \|S_0 f\|_{p_1} \lesssim \|S_0 f\|_p, \tag{9.20}$$

since $1 \leqslant p \leqslant p_1 \leqslant \infty$. Thus, with the help of the embedding $B^s_{p,r} \hookrightarrow B^s_{p,r_1}$ for $r \leqslant r_1$ in Theorem 9.15, we obtain

$$\begin{aligned}\|f\|_{B^{s_1}_{p_1,r_1}} =& \|S_0 f\|_{p_1} + \left(\sum_{k=0}^\infty \left(2^{s_1 k}\|\Delta_k f\|_{p_1}\right)^{r_1}\right)^{\frac{1}{r_1}}\\ \lesssim& \|S_0 f\|_p + \left(\sum_{k=0}^\infty \left(2^{sk}\|\Delta_k f\|_p\right)^{r_1}\right)^{\frac{1}{r_1}} = \|f\|_{B^s_{p,r_1}} \lesssim \|f\|_{B^s_{p,r}}.\end{aligned}$$

This gives the first conclusion.

Next, we prove the second conclusion. In view of Theorem 9.15, we need only prove $F^s_{p,\infty} \hookrightarrow F^{s_1}_{p_1,1}$. Without loss of generality, we assume $\|f\|_{F^s_{p,\infty}} = 1$ and consider the norm

$$\|f\|_{F^{s_1}_{p_1,1}} = \|S_0 f\|_{p_1} + \left\|\sum_{k=0}^\infty 2^{s_1 k}|\Delta_k f|\right\|_{p_1}.$$

We use the following equivalent norm (i.e., Theorem 1.17) on L^p for $1 \leqslant p < \infty$:

$$\|f\|_p^p = p\int_0^\infty t^{p-1}|\{x : |f(x)| > t\}|dt.$$

Thus, we have

$$\left\|\sum_{k=0}^\infty 2^{s_1 k}|\Delta_k f|\right\|_{p_1}^{p_1} = p_1\int_0^A t^{p_1-1}\left|\left\{x : \sum_{k=0}^\infty 2^{s_1 k}|\Delta_k f(x)| > t\right\}\right|dt$$

$$+ p_1 \int_A^\infty t^{p_1-1} \left| \left\{ x : \sum_{k=0}^\infty 2^{s_1 k} |\Delta_k f(x)| > t \right\} \right| dt$$
$$=: I + II,$$

where $A \gg 1$ is a constant that can be chosen as below. Note that $p < p_1$ and $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$ imply $s > s_1$, we have

$$\sum_{k=K}^\infty 2^{s_1 k} |\Delta_k f| \lesssim 2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk} |\Delta_k f|, \quad \forall K \in \mathbb{N}_0. \tag{9.21}$$

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 obtain

$$\begin{aligned} I \lesssim & \int_0^A t^{p_1-1} \left| \left\{ x : \sup_{k\geqslant 0} 2^{sk} |\Delta_k f(x)| > ct \right\} \right| dt \\ \lesssim & \int_0^{cA} \tau^{p-1} \left| \left\{ x : \sup_{k\geqslant 0} 2^{sk} |\Delta_k f(x)| > \tau \right\} \right| d\tau \lesssim \left\| \sup_{k\geqslant 0} 2^{sk} |\Delta_k f| \right\|_p^p \lesssim 1, \end{aligned}$$

where the implicit constant depends on A, but it is a fixed constant.

Now we estimate II. By the Bernstein inequality in Proposition 9.1, we have

$$\|\Delta_k f\|_\infty \lesssim 2^{kn/p} \|\Delta_k f\|_p \lesssim 2^{k(n/p-s)} \left\| \sup_{k\geqslant 0} 2^{sk} |\Delta_k f| \right\|_p .$$

Hence, for $K \in \mathbb{N}$, we obtain

$$\begin{aligned} \sum_{k=0}^{K-1} 2^{s_1 k} |\Delta_k f| \lesssim & \sum_{k=0}^{K-1} 2^{k(s_1-s+n/p)} \left\| \sup_{k\geqslant 0} 2^{sk} |\Delta_k f| \right\|_p \\ \lesssim & 2^{Kn/p_1} \left\| \sup_{k\geqslant 0} 2^{sk} |\Delta_k f| \right\|_p \lesssim 2^{Kn/p_1}. \end{aligned} \tag{9.22}$$

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 $\sum_{k=0}^\infty 2^{s_1 k} |(\Delta_k f)(x)| > t$, we have, from (9.21) and (9.22), that

$$C 2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk} |\Delta_k f| \geqslant \sum_{k=K}^\infty 2^{s_1 k} |\Delta_k f| > t/2. \tag{9.23}$$

Hence, from (9.22) and (9.23), we obtain

$$\begin{aligned} II =& p_1 \int_A^\infty t^{p_1-1} \left| \left\{ x : \sum_{k=0}^\infty 2^{s_1 k} |\Delta_k f(x)| > t \right\} \right| dt \\ \lesssim & \int_A^\infty t^{p_1-1} \left| \left\{ x : \sum_{k=0}^{K-1} 2^{s_1 k} |\Delta_k f(x)| > t/2 \right\} \right| dt \end{aligned}$$

$$
\begin{aligned}
&\quad + \int_A^\infty t^{p_1-1} \left|\left\{x : \sum_{k=K}^\infty 2^{s_1 k}|\Delta_k f(x)| > t/2\right\}\right| dt \\
&\lesssim \int_A^\infty t^{p_1-1} \left|\left\{x : C2^{Kn/p_1} > t/2\right\}\right| dt \\
&\quad + \int_A^\infty t^{p_1-1} \left|\left\{x : C2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > t/2\right\}\right| dt \\
&\lesssim \int_A^\infty t^{p_1-1} \left|\left\{x : \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > ct^{p_1/p}\right\}\right| dt \\
&\lesssim \int_{A'}^\infty \tau^{p-1} \left|\left\{x : \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > \tau\right\}\right| d\tau \\
&\lesssim \|\sup_{k\geqslant 0} 2^{sk}|\Delta_k f|\|_p^p \lesssim 1.
\end{aligned}
$$

That is,

$$
\left\|\sum_{k=0}^\infty 2^{s_1 k}|\Delta_k f|\right\|_{p_1} \lesssim 1.
$$

However, from (9.20), we have $\|S_0 f\|_{p_1} \lesssim 1$. Therefore, we have obtained $\|f\|_{F^{s_1}_{p_1,1}} \lesssim 1$ under the assumption $\|f\|_{F^s_{p,\infty}} = 1$. This completes the proof. ❑

Theorem 9.19. *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*

$$
X^s_{p,r} \hookrightarrow B^0_{\infty,1} \hookrightarrow L^\infty.
$$

Proof. By Bernstein's inequality and Theorem 9.15, we have

$$
\begin{aligned}
\|f\|_\infty &\leqslant \sum_{k=-1}^\infty \|\Delta_k f\|_\infty \lesssim \sum_{k=-1}^\infty 2^{kn/p}\|\Delta_k f\|_p \\
&\lesssim \left(\sum_{k=-1}^\infty 2^{k(n/p-s)}\right) \|f\|_{B^s_{p,\infty}} \lesssim \|f\|_{X^s_{p,r}}.
\end{aligned}
$$

❑

Now, we give some fractional Gagliardo-Nirenberg inequalities in homogeneous Besov spaces.

Theorem 9.20. *Let $1 \leqslant p, p_0, p_1, r, r_0, r_1 \leqslant \infty$, $s, s_0, s_1 \in \mathbb{R}$, $0 \leqslant \theta \leqslant 1$. Suppose that the following conditions hold:*

$$
s - \frac{n}{p} = (1-\theta)\left(s_0 - \frac{n}{p_0}\right) + \theta\left(s_1 - \frac{n}{p_1}\right), \tag{9.24}
$$

$$
s \leqslant (1-\theta)s_0 + \theta s_1, \tag{9.25}
$$

$$\frac{1}{r} \leqslant \frac{1-\theta}{r_0} + \frac{\theta}{r_1}. \tag{9.26}$$

Then the fractional GN inequality of the following type

$$\|u\|_{\dot{B}^s_{p,r}} \lesssim \|u\|^{1-\theta}_{\dot{B}^{s_0}_{p_0,r_0}} \|u\|^{\theta}_{\dot{B}^{s_1}_{p_1,r_1}} \tag{9.27}$$

holds for all $u \in \dot{B}^{s_0}_{p_0,r_0} \cap \dot{B}^{s_1}_{p_1,r_1}$.

Proof. Let $s^* = (1-\theta)s_0 + \theta s_1$, $1/p^* = (1-\theta)/p_0 + \theta/p_1$ and $1/r^* = (1-\theta)/r_0 + \theta/r_1$. By (9.25), we have $s \leqslant s^*$ and $r^* \leqslant r$. Applying the convexity Hölder inequality, we have

$$\|f\|_{\dot{B}^{s^*}_{p^*,r^*}} \leqslant \|f\|^{1-\theta}_{\dot{B}^{s_0}_{p_0,r_0}} \|f\|^{\theta}_{\dot{B}^{s_1}_{p_1,r_1}}. \tag{9.28}$$

Using the embedding $\dot{B}^{s^*}_{p^*,r^*} \hookrightarrow \dot{B}^s_{p,r}$, we obtain the conclusion. ❑

Now, we give the duality theorem:

Theorem 9.21 (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 [BL76a; Tri83] for details. ❑

§9.3 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

$$\omega^m_p(t,f) = \sup_{|y|\leqslant t} \| \triangle^m_y f\|_p,$$

where $\triangle^m_y$ is the m-th order difference operator:

$$\triangle^m_y f(x) = \sum_{k=0}^{m} C^k_m (-1)^k f(x+ky).$$

Theorem 9.22. *Assume that* $s > 0$, *and let* m *and* N *be integers, such that* $m + N > s$ *and* $0 \leqslant N < s$. *Then, for* $1 \leqslant p, r \leqslant \infty$,

$$\|f\|_{B^s_{p,r}} \sim \|f\|_p + \sum_{j=1}^{n} \left(\int_0^\infty \left(t^{N-s} \omega^m_p \left(t, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \frac{dt}{t} \right)^{1/r}.$$

Proof. 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

$$\begin{aligned}
\omega_p^m(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N}) &= \sup_{|y|\leqslant 2^{-\ell}} \left\| \triangle_y^m \frac{\partial^N f}{\partial x_j^N} \right\|_p = \sup_{|y|\leqslant 2^{-\ell}} \left\| \sum_{k=0}^m C_m^k (-1)^k \frac{\partial^N f}{\partial x_j^N}(x+ky) \right\|_p \\
&= \sup_{|y|\leqslant 2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \left(\sum_{k=0}^m C_m^k (-1)^k f(x+ky) \right) \right\|_p \\
&= \sup_{|y|\leqslant 2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \left(\sum_{k=0}^m C_m^k (-1)^k \left(e^{iky\cdot\xi} \widehat{f} \right)^\vee \right) \right\|_p \\
&= \sup_{|y|\leqslant 2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \left(\sum_{k=0}^m C_m^k (-1)^k e^{iky\cdot\xi} \widehat{f} \right)^\vee \right\|_p \\
&= \sup_{|y|\leqslant 2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \left((1-e^{iy\cdot\xi})^m \widehat{f} \right)^\vee \right\|_p.
\end{aligned}$$

Denote $\rho_y(\xi) = (1-e^{iy\cdot\xi})^m$. By the Littlewood-Paley decomposition and the Bernstein inequalities, we have

$$\begin{aligned}
&\omega_p^m(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N}) \\
= &\sup_{|y|\leqslant 2^{-\ell}} \left\| \left(S_0 + \sum_{k=0}^\infty \Delta_k \right) \frac{\partial^N}{\partial x_j^N} \left(\rho_y(\xi) \widehat{f} \right)^\vee \right\|_p \\
\lesssim &\sup_{|y|\leqslant 2^{-\ell}} \left\| (\rho_y)^\vee * S_0 f \right\|_p + \sup_{|y|\leqslant 2^{-\ell}} \sum_{k=0}^\infty 2^{kN} \left\| (\rho_y)^\vee * \Delta_k f \right\|_p.
\end{aligned}$$

If we can prove that for all integers k

$$\| (\rho_y)^\vee * S_0 f\|_p \lesssim \min(1, |y|^m) \|S_0 f\|_p, \tag{9.29}$$

and

$$\| (\rho_y)^\vee * \Delta_k f\|_p \lesssim \min(1, |y|^m 2^{mk}) \|\Delta_k f\|_p. \tag{9.30}$$

Then, we can obtain

$$\begin{aligned}
&\sum_{j=1}^n \left(\sum_{\ell=-\infty}^{\infty} \left(2^{\ell(s-N)} \omega_p^m \left(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \right)^{1/r} \\
\lesssim &\Big(\sum_{\ell=-\infty}^{\infty} \Big(2^{\ell(s-N)} \sup_{|y|\leqslant 2^{-\ell}} \min(1, |y|^m) \|S_0 f\|_p
\end{aligned}$$

$$+ \sup_{|y|\leqslant 2^{-\ell}} \sum_{k=0}^{\infty} 2^{(\ell-k)(s-N)} 2^{ks} \min(1, |y|^m 2^{mk}) \|\Delta_k f\|_p \Big)^r \Big)^{1/r}$$

$$\lesssim \Big(\sum_{\ell=-\infty}^{\infty} \Big(2^{\ell(s-N)} \min(1, 2^{-\ell m}) \|S_0 f\|_p$$

$$+ \sum_{k=0}^{\infty} 2^{(\ell-k)(s-N)} \min(1, 2^{-(\ell-k)m}) 2^{ks} \|\Delta_k f\|_p \Big)^r \Big)^{1/r}$$

$$\lesssim \|(2^{k(s-N)} \min(1, 2^{-km})) * (\alpha_k)\|_{\ell^r}$$

$$\lesssim \|(2^{k(s-N)} \min(1, 2^{-km}))\|_{\ell^1} \|(\alpha_k)\|_{\ell^r} \lesssim \|f\|_{B^s_{p,r}},$$

where the sequence $(\alpha_k)_{k=-\infty}^{\infty}$ with $\alpha_k = 2^{sk}\|\Delta_k f\|_p$ 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

$$\|f\|_p \lesssim \|S_0 f\|_p + \sum_{k=0}^{\infty} \|\Delta_k f\|_p$$

$$\lesssim \|S_0 f\|_p + \Big(\sum_{k=0}^{\infty} 2^{-skr'} \Big)^{1/r'} \Big(\sum_{k=0}^{\infty} (2^{sk} \|\Delta_k f\|_p)^r \Big)^{1/r} \lesssim \|f\|_{B^s_{p,r}},$$

which implies the desired conclusion.

Now, we turn to prove (9.29) and (9.30). We only need to show $\rho_y \in \mathcal{M}_p$, $\rho_y(\cdot)\langle y, \cdot\rangle^{-m} \in \mathcal{M}_p$ for $p \in [1, \infty]$ and

$$\|\rho_y\|_{\mathcal{M}_p} \leqslant C, \quad \|\rho_y(\cdot)\langle y, \cdot\rangle^{-m}\|_{\mathcal{M}_p} \leqslant C, \quad \forall y \neq 0. \tag{9.31}$$

From the definition of ρ_y, we obtain

$$\|\rho_y\|_{\mathcal{M}_p} = (2\pi)^{-n/2} \sup_{f \in \mathscr{S}} \frac{\| (\rho_y)^{\vee} * f\|_p}{\|f\|_p} = \sup_{f\in\mathscr{S}} \frac{\|\sum_{k=0}^m C_m^k (-1)^k f(x+ky)\|_p}{\|f\|_p}$$

$$\leqslant \sum_{k=0}^{m} C_m^k = 2^m.$$

By Theorem 3.53, we have

$$\begin{aligned} \|\rho_y(\xi)\langle y, \xi\rangle^{-m}\|_{\mathcal{M}_p(\mathbb{R}^n)} &= \|(1 - e^{i\langle y, \xi\rangle})^m \langle y, \xi\rangle^{-m}\|_{\mathcal{M}_p(\mathbb{R}^n)} \\ &= \|((1-e^{i\eta})/\eta)^m\|_{\mathcal{M}_p(\mathbb{R})} \\ &\leqslant \|(1-e^{i\eta})/\eta\|^m_{\mathcal{M}_p(\mathbb{R})}, \end{aligned}$$

since $\mathcal{M}_p$ is a Banach algebra and the integer $m \geqslant 1$ in view of the conditions $m + N > s$ and $0 \leqslant N < s$.

In view of the Bernstein multiplier theorem (i.e., Theorem 3.55), we only need to show $(1 - e^{i\eta})/\eta \in L^2(\mathbb{R})$ and $\partial_\eta((1 - e^{i\eta})/\eta) \in L^2(\mathbb{R})$. We split the L^2 integral into two parts $|\eta| < 1$ and $|\eta| \geqslant 1$. For $|\eta| < 1$, we can use $|1 - e^{i\eta}| \leqslant |\eta|$ to obtain $|(1 - e^{i\eta})/\eta| \leqslant 1$; while for its first order derivative, we can use Taylor's expansion $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ whenever $|z| < \infty$

($z \in \mathbb{C}$) to obtain

$$\begin{aligned}
\partial_\eta((1-e^{i\eta})/\eta) =& -\eta^{-2}(i\eta e^{i\eta}+1-e^{i\eta}) \\
=& -\eta^{-2}\left(i\eta\sum_{k=0}^{\infty}\frac{(i\eta)^k}{k!}-\sum_{k=1}^{\infty}\frac{(i\eta)^k}{k!}\right) \\
=& -\eta^{-2}\left(\sum_{k=0}^{\infty}\frac{(i\eta)^{k+1}}{k!}-\sum_{k=0}^{\infty}\frac{(i\eta)^{k+1}}{(k+1)!}\right) \\
=& \sum_{k=1}^{\infty}\frac{k(i\eta)^{k-1}}{(k+1)!} \\
=& \sum_{k=1}^{\infty}\frac{(i\eta)^{k-1}}{(k-1)!}\frac{1}{k+1},
\end{aligned}$$

which implies $|\partial_\eta((1-e^{i\eta})/\eta)| \leqslant \frac{1}{2}e^{|\eta|}$. Then it is easy to obtain the bound of the L^2 integral. Thus, $\|((1-e^{i\eta})/\eta)\|_{\mathcal{M}_p(\mathbb{R})} \leqslant C$ by Theorem 3.55, which completes the proof of (9.31).

Similarly, we can prove

$$\|\langle y/|y|,\cdot\rangle^m\widetilde{\psi}(\cdot)\|_{\mathcal{M}_p} \leqslant C, \text{ and } \|\langle y/|y|,\cdot\rangle^m\widetilde{\varphi}(\cdot)\|_{\mathcal{M}_p} \leqslant C,$$

which implies by Theorem 3.53

$$\|\langle y,\cdot\rangle^m\widetilde{\psi}(\cdot)\|_{\mathcal{M}_p} \leqslant C|y|^m, \quad \|\langle y,\cdot\rangle^m\widetilde{\varphi}(2^{-k}\cdot)\|_{\mathcal{M}_p} \leqslant C|y|^m 2^{mk}.$$

Thus, we obtain

$$\begin{aligned}
&\|\left(\rho_y\right)^\vee * S_0 f\|_p \lesssim \|S_0 f\|_p, \\
&\|\left(\rho_y\right)^\vee * S_0 f\|_p \\
=&(2\pi)^{-n/2}\|\left(\rho_y(\xi)\langle y,\xi\rangle^{-m}\right)^\vee * \left(\langle y,\xi\rangle^m\widetilde{\psi}(\xi)\right)^\vee * S_0 f\|_p \\
\lesssim&|y|^m\|S_0 f\|_p,
\end{aligned}$$

which yields (9.29). Similarly, we have

$$\begin{aligned}
&\|\left(\rho_y\right)^\vee * \Delta_k f\|_p \lesssim \|\Delta_k f\|_p, \\
&\|\left(\rho_y\right)^\vee * \Delta_k f\|_p \\
=&(2\pi)^{-n/2}\|\left(\rho_y(\xi)\langle y,\xi\rangle^{-m}\right)^\vee * \left(\langle y,\xi\rangle^m\widetilde{\varphi}(2^{-k}\xi)\right)^\vee * \Delta_k f\|_p \\
\lesssim&|y|^m 2^{mk}\|\Delta_k f\|_p,
\end{aligned}$$

which yields (9.30).

The converse inequality will follow if we can prove the estimate

$$\|\Delta_k f\|_p \lesssim 2^{-Nk}\sum_{j=1}^{n}\left\|\left(\rho_{jk}\right)^\vee * \frac{\partial^N f}{\partial x_j^N}\right\|_p, \tag{9.32}$$

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 (9.32) is valid, we have

$$\|f\|_{B^s_{p,r}} \lesssim \|f\|_p + \left(\sum_{k=0}^{\infty}\left(2^{k(s-N)}\sum_{j=1}^{n}\left\|(\rho_{jk})^{\vee} * \frac{\partial^N f}{\partial x_j^N}\right\|_p\right)^r\right)^{1/r}$$
$$\lesssim \|f\|_p + \sum_{j=1}^{n}\left(\sum_{k=0}^{\infty}\left(2^{k(s-N)}\omega_p^m\left(2^{-k}, \frac{\partial^N f}{\partial x_j^N}\right)\right)^r\right)^{1/r},$$

which implies the desired inequality.

To prove (9.32), we need the following lemma.

Lemma 9.23. *Assume that $n \geqslant 2$ and take φ as in (9.2). Then there exist some functions $\chi_j \in \mathscr{S}(\mathbb{R}^n)$ $(1 \leqslant j \leqslant n)$, such that*

$$\sum_{j=1}^{n}\chi_j = 1 \quad \text{on } \operatorname{supp}\varphi,$$
$$\operatorname{supp}\chi_j \subset \left\{\xi \in \mathbb{R}^n : |\xi_j| \geqslant (3\sqrt{n})^{-1}\right\}, \quad 1 \leqslant j \leqslant n.$$

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 $\sigma \in \mathscr{S}(\mathbb{R}^{n-1})$ with $\operatorname{supp}\sigma = \{\xi \in \mathbb{R}^{n-1} : |\xi| \leqslant 3\}$ and positive in the interior. Writing

$$\bar{\xi}^j = (\xi_1, \cdots, \xi_{j-1}, \xi_{j+1}, \cdots, \xi_n)$$

and

$$\chi_j(\xi) = \kappa(\xi_j)\sigma(\bar{\xi}^j) \Big/ \sum_{j=1}^{n}\kappa(\xi_j)\sigma(\bar{\xi}^j), \quad 1 \leqslant j \leqslant n,$$

where $\sum_{j=1}^{n}\kappa(\xi_j)\sigma(\bar{\xi}^j) > 0$ on $\operatorname{supp}\varphi$, only routine verification remains to complete the proof of the lemma. ❑

We now complete the proof of the theorem, i.e., we prove (9.32). By the previous lemma, we obtain the formula

$$\|\Delta_k f\|_p \lesssim \sum_{j=1}^{n}\left\|\left(\rho_{jk}^{-1}\chi_j(2^{-k}\cdot)\xi_j^{-N}\varphi(2^{-k}\cdot)\right)^{\vee} * \left(\rho_{jk}\widehat{\frac{\partial^N f}{\partial x_j^N}}\right)^{\vee}\right\|_p$$
$$\lesssim 2^{-kN}\sum_{j=1}^{n}\left\|\left(\rho_{jk}^{-1}\chi_j(2^{-k}\cdot)(2^{-k}\xi_j)^{-N}\varphi(2^{-k}\cdot)\right)^{\vee} * \left(\rho_{jk}\widehat{\frac{\partial^N f}{\partial x_j^N}}\right)^{\vee}\right\|_p$$
$$\lesssim 2^{-kN}\sum_{j=1}^{n}\left\|\rho_{j0}^{-1}\chi_j\xi_j^{-N}\varphi\right\|_{\mathcal{M}_p(\mathbb{R}^n)}\left\|(\rho_{jk})^{\vee} * \frac{\partial^N f}{\partial x_j^N}\right\|_p$$
$$\lesssim 2^{-kN}\sum_{j=1}^{n}\left\|(\rho_{jk})^{\vee} * \frac{\partial^N f}{\partial x_j^N}\right\|_p,$$

since, by Theorem 3.53 and 3.55, we have

$$(1-e^{i\xi_j})^{-m}\chi_j(\xi)\xi_j^{-N}\varphi(\xi)\in\mathcal{M}_p,$$

for $1\leqslant j\leqslant n$ and $1\leqslant p\leqslant\infty$. ❑

Now, we give a corollary that is very convenient for nonlinear estimates in PDEs.

Corollary 9.24. *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 obtain an equivalent norm for the homogeneous Besov space.

Theorem 9.25. *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

$$\|f\|_{\dot{B}^s_{p,r}}\sim\sum_{j=1}^n\left(\int_0^\infty\left(t^{[s]-s}\sup_{|h|\leqslant t}\left\|\triangle_h\,\partial^{[s]}_{x_j}f\right\|_p\right)^r\frac{dt}{t}\right)^{1/r},$$

Exercises

Exercise 9.1. [Gra14b, Exercise 2.2.3] Let $\alpha\in\mathbb{R}$, $\beta>0$ and $p\in[1,\infty)$. Let $1'=\infty$ and $p'=p/(p-1)$ for $p\neq1$.

(a) Suppose that the Fourier transform of function g is $\mathcal{C}^\infty$ and is equal to $|\xi|^{-\alpha}$ for $|\xi|\geqslant10$. Show that $g\in B^s_{p,r}(\mathbb{R}^n)$ iff $1\leqslant r<\infty$ and $s<\alpha-n/p'$ or $r=\infty$ and $s\leqslant\alpha-n/p'$.

(b) If the Fourier transform of function g is $\mathcal{C}^\infty$ and is equal to $|\xi|^{-\alpha}(\ln|\xi|)^{-\beta}$ for $|\xi|\geqslant10$, then show that $g\in B^{\alpha-n/p'}_{p,r}(\mathbb{R}^n)$ iff $r>1/\beta$.

Exercise 9.2. [Gra14b, Exercise 2.2.5] Let $s\in\mathbb{R}$, $p,r\in[1,\infty)$, and $N=\left[\frac{n}{2}+\frac{n}{\min(p,r)}\right]+1$. Assume that m is a $\mathcal{C}^N$ function on $\mathbb{R}^n\setminus\{0\}$ that satisfies

$$|\partial^\alpha m(\xi)|\leqslant C_\alpha|\xi|^{-|\alpha|}$$

for all $|\alpha|\leqslant N$. Show that there exists a constant C such that for all $f\in\dot{\mathscr{S}}'$

we have

$$\left\| \left(m\widehat{f} \right)^{\vee} \right\|_{\dot{B}^s_{p,r}} \leqslant C\|f\|_{\dot{B}^s_{p,r}}.$$

10

Paraproducts - An Introduction

§ 10.1 The realization of homogeneous Besov spaces for PDEs

When we consider partial differential equations, it is not conformable to work on the quotient space. One of the reasons is that the quotient space does not give us any information on the value of functions. Therefore, at least we want to return to the subspace of $\mathscr{S}'$. Although the evaluation does not make sense in $\mathscr{S}'$, we feel that the situation becomes better in $\mathscr{S}'$ than in $\dot{\mathscr{S}}' = \mathscr{S}'/\mathscr{P}$. Such a situation is available when s is small enough.

Theorem 10.1. *Let $1 \leqslant p, r \leqslant \infty$. Assume*

$$s < \frac{n}{p}, \quad \text{or} \quad s = \frac{n}{p} \text{ and } r = 1. \tag{10.1}$$

Then, for all $f \in \dot{B}^s_{p,r}$, $\sum_{k=-\infty}^{0} \dot{\Delta}_k f$ is convergent in L^∞ and $\sum_{k=1}^{\infty} \dot{\Delta}_k f$ is convergent in $\mathscr{S}'$.

Proof. From Bernstein's inequality, we have $\|\dot{\Delta}_k f\|_\infty \lesssim 2^{kn/p}\|\dot{\Delta}_k f\|_p$. It follows that

$$\left\|\sum_{k=-\infty}^{0} \dot{\Delta}_k f\right\|_\infty \lesssim \sum_{k=-\infty}^{0} \|\dot{\Delta}_k f\|_\infty \lesssim \sum_{k=-\infty}^{0} 2^{k(n/p-s)} 2^{ks}\|\dot{\Delta}_k f\|_p$$
$$\lesssim \begin{cases} \|f\|_{\dot{B}^s_{p,\infty}} \lesssim \|f\|_{\dot{B}^s_{p,r}}, & \text{if } s < n/p, \\ \|f\|_{\dot{B}^{n/p}_{p,1}}, & \text{if } s = n/p \text{ and } r = 1. \end{cases}$$

The fact that $\sum_{k=1}^{\infty} \dot{\Delta}_k f$ is convergent in $\mathscr{S}'$ is a general fact. ❑

There is a way to modify the definition of homogeneous Besov spaces regarding the regularity index. For convenience, we first define a subspace of $\mathscr{S}'(\mathbb{R}^n)$ that will play an important role in studying PDEs.

Definition 10.2. We denote by $\mathscr{S}'_h(\mathbb{R}^n)$ the space of tempered distributions f such that

$$\lim_{\lambda\to\infty} \|\theta(\lambda D)f\|_\infty = 0, \quad \forall \theta \in \mathscr{D}(\mathbb{R}^n), \tag{10.2}$$

where the operator $\theta(D)$ is defined by $\theta(D)f := \left(\theta \widehat{f}\right)^\vee$ for a measurable function f on $\mathbb{R}^n$ with at most polynomial growth at infinity.

Remark 10.3. We have the following facts about $\mathscr{S}'_h(\mathbb{R}^n)$.
1) It holds

$$\mathscr{S}'_h(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) : \lim_{k\to-\infty} \dot{S}_k f = 0 \text{ in } L^\infty(\mathbb{R}^n) \right\}, \tag{10.3}$$

and

$$\mathscr{S}'_h(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) : f = \sum_{k\in\mathbb{Z}} \dot{\Delta}_k f \text{ in } \mathscr{S}'(\mathbb{R}^n) \right\}. \tag{10.4}$$

In fact, since $\psi \in \mathscr{D}$ given in (9.1), we have $\dot{S}_k f = \psi_k(D) f = \psi(2^{-k}D) f \to 0$ in L^∞ as $k \to -\infty$ if f satisfies (10.2).

Conversely, for a given $\theta \in \mathscr{D}$, we may assume $\operatorname{supp}\theta \subset \{\xi : |\xi| \leqslant C\}$. It follows that $\varphi_k(\xi) = 0$ if $2^k\alpha^{-1} > C/\lambda$, i.e., $k > \log_2 \frac{C}{\lambda\alpha}$. It holds for any $g \in \mathscr{S}$,

$$\begin{aligned}
|\langle \theta(\lambda D)f, g\rangle| =& |\langle \widehat{f}(\xi), \theta(\lambda\xi)\check{g}\rangle| \\
=& \left| \left\langle \sum_{k\in\mathbb{Z}} \varphi_k(\xi)\widehat{f}, \theta(\lambda\xi)\check{g} \right\rangle \right| \\
=& \left| \left\langle \sum_{k\leqslant[\log_2 \frac{C}{\lambda\alpha}]} \varphi_k(\xi)\widehat{f}, \theta(\lambda\xi)\check{g} \right\rangle \right| \\
=& (2\pi)^{-n/2} \left| \left\langle \dot{S}_{[\log_2 \frac{C}{\lambda\alpha}]+1} f, \widehat{\theta(\lambda\cdot)} * g \right\rangle \right| \to 0 \text{ as } \lambda \to \infty,
\end{aligned}$$

by (10.3) due to $\widehat{\theta(\lambda\cdot)} * g \in \mathscr{S}$ and the fact that $\|\widehat{\theta(\lambda\cdot)} * g\|_1 \leqslant \|\widehat{\theta(\lambda\cdot)}\|_1 \|g\|_1 = \|\widehat{\theta}\|_1 \|g\|_1$ by Young's inequality, i.e., $\|\widehat{\theta(\lambda\cdot)} * g\|_1$ is uniformly bounded w.r.t. λ. Taking supremum over all $g \in \mathscr{S}$ with $\|g\|_1 \leqslant 1$, we obtain $\|\theta(\lambda D)f\|_\infty \to 0$ as $\lambda \to \infty$ since the Lebesgue measure on $\mathbb{R}^n$ is obviously semifinite.

For (10.4), noticing $\dot{\Delta}_k = \dot{S}_{k+1} - \dot{S}_k$ and by Proposition 9.4 and (10.3), we have for any $g \in \mathscr{S}$

$$\begin{aligned}
\left\langle \sum_{k\in\mathbb{Z}} \dot{\Delta}_k f, g \right\rangle =& \left\langle \sum_{k\in\mathbb{Z}} (\dot{S}_{k+1} f - \dot{S}_k f), g \right\rangle \\
=& \left\langle \lim_{k\to+\infty} \dot{S}_{k+1} f - \lim_{k\to-\infty} \dot{S}_k f, g \right\rangle \\
=& \langle f, g\rangle.
\end{aligned}$$

On the other hand, from Proposition 9.4 and (10.4), from the above equality, we obtain $\left\langle \lim_{k\to-\infty} \dot{S}_k f, g \right\rangle = 0$ for any $g \in \mathscr{S}$, and it follows (10.3).

2) It is clear that whether a tempered distribution f belongs to $\mathscr{S}'_h$ depends only on low frequencies. If a tempered distribution f is such that its Fourier transform $\widehat{f}$ is locally integrable near 0, then $f \in \mathscr{S}'_h$. In particular, the space $\mathscr{E}'$ of compactly supported distributions is included in $\mathscr{S}'_h$. In fact, for any $g \in \mathscr{S}$, we obtain

$$\begin{aligned}|\langle \dot{S}_k f, g\rangle| =&|\langle \psi(2^{-k}\xi)\widehat{f}(\xi), \check{g}(\xi)\rangle| \\ \leqslant& \int_{|\xi|\leqslant 2^k\alpha} |\widehat{f}(\xi)||\check{g}(\xi)|d\xi \\ \leqslant& C\int_{|\xi|\leqslant 2^k\alpha} |\widehat{f}(\xi)|d\xi \to 0, \text{ as } k \to -\infty,\end{aligned}$$

since $\widehat{f}$ is locally integrable near 0. Thus, $f \in \mathscr{S}'_h$.

3) $f \in \mathscr{S}'_h(\mathbb{R}^n) \Leftrightarrow \exists \theta \in \mathscr{D}(\mathbb{R}^n)$, s.t. $\lim_{\lambda\to\infty} \|\theta(\lambda D)f\|_\infty = 0$ and $\theta(0) \neq 0$. Indeed, the necessity is clear from the definition. For the sufficiency, by assumption, there is an $\ell \in \mathbb{Z}$ small enough such that $\operatorname{supp}\psi_\ell \subset \operatorname{supp}\theta$, then

$$\begin{aligned}|\langle \dot{S}_k f, g\rangle| =& \left|\left\langle \theta(2^{\ell-k}\xi)\widehat{f}(\xi), \frac{\psi(2^{-k}\xi)}{\theta(2^{\ell-k}\xi)}\check{g}(\xi)\right\rangle\right| \\ \leqslant& (2\pi)^{-n/2}\|\theta(2^{\ell-k}D)f\|_\infty \left\|\mathscr{F}\left(\frac{\psi(2^{-k}\xi)}{\theta(2^{\ell-k}\xi)}\right)\right\|_1 \|g\|_1 \\ =& (2\pi)^{-n/2}\|\theta(2^{\ell-k}D)f\|_\infty \left\|\mathscr{F}\left(\frac{\psi_\ell}{\theta}\right)\right\|_1 \|g\|_1 \qquad (10.5) \\ \leqslant& C\|\theta(2^{\ell-k}D)f\|_\infty \to 0, \text{ as } k \to -\infty,\end{aligned}$$

since $\frac{\psi_\ell}{\theta} \in \mathscr{D} \subset \mathscr{S}$.

4) Obviously, $f \in \mathscr{S}'_h(\mathbb{R}^n) \Leftrightarrow \forall \theta \in \mathscr{D}(\mathbb{R}^n)$ with value 1 near the origin, we have

$$\lim_{\lambda\to\infty} \|\theta(\lambda D)f\|_\infty = 0.$$

5) If $f \in \mathscr{S}'$ satisfies $\theta(D)f \in L^p$ for some $p \in [1,\infty)$ and some function $\theta \in \mathscr{D}(\mathbb{R}^n)$ with $\theta(0) \neq 0$, then $f \in \mathscr{S}'_h$. In fact, similar to (10.5), we can also obtain for any $k < \ell$

$$\begin{aligned}|\langle \dot{S}_k f, g\rangle| =& \left|\left\langle \theta(\xi)\widehat{f}(\xi), \frac{\psi(2^{-k}\xi)}{\theta(\xi)}\check{g}(\xi)\right\rangle\right| \\ \leqslant& (2\pi)^{-n/2}\|\theta(D)f\|_p \left\|\mathscr{F}\left(\frac{\psi(2^{-k}\xi)}{\theta(\xi)}\right)\right\|_{p'} \|g\|_1 \\ =& (2\pi)^{-n/2}\|\theta(D)f\|_p \left\|2^{kn}\mathscr{F}\left(\frac{\psi(\cdot)}{\theta(2^k\cdot)}\right)(2^k\cdot)\right\|_{p'} \|g\|_1\end{aligned}$$

$$=(2\pi)^{-n/2}2^{kn/p}\|\theta(D)f\|_p \left\|\mathscr{F}\left(\frac{\psi(\cdot)}{\theta(2^k\cdot)}\right)\right\|_{p'} \|g\|_1 \to 0,$$

as $k \to -\infty$, with the help of $\theta(2^k\cdot) \to \theta(0) \neq 0$ as $k \to -\infty$ and the uniform continuity of the Fourier transform for L^1 functions.

6) A nonzero polynomial P does not belong to $\mathscr{S}'_h$ because for any $\theta \in \mathscr{D}(\mathbb{R}^n)$ with value 1 near 0 and any $\lambda > 0$, we may write $\theta(\lambda D)P = P$. In fact, $\forall \alpha \in \mathbb{N}_0^n$, $\forall g \in \mathscr{S}$,

$$\begin{aligned}
&\langle \theta(\lambda D)x^\alpha, g(x)\rangle = \langle \theta(\lambda\xi)\widehat{x^\alpha}(\xi), \check{g}(\xi)\rangle = \langle x^\alpha, \widehat{\theta(\lambda\xi)\check{g}(\xi)}\rangle \\
=&\langle 1, x^\alpha \widehat{\theta(\lambda\xi)\check{g}(\xi)}\rangle \\
=&\langle 1, \widehat{(-i\partial_\xi)^\alpha(\theta(\lambda\xi)\check{g}(\xi))}\rangle \\
=&\left\langle (2\pi)^{n/2}\delta_0(\xi), \sum_{\alpha=\beta+\gamma} C_\alpha^\beta(-i\lambda)^\beta(\partial_\xi^\beta\theta)(\lambda\xi)(-i\partial_\xi)^\gamma\check{g}(\xi)\right\rangle \\
=&(2\pi)^{n/2}\sum_{\alpha=\beta+\gamma} C_\alpha^\beta(-i\lambda)^\beta(\partial_\xi^\beta\theta)(0)\,(x^\gamma g)^\vee(0) \\
=&(2\pi)^{n/2}(x^\alpha g)^\vee(0) = \langle (2\pi)^{n/2}\delta_0, (x^\alpha g)^\vee\rangle \\
=&\langle 1, x^\alpha g\rangle = \langle x^\alpha, g(x)\rangle,
\end{aligned}$$

since $(\partial^\beta\theta)(0) = 0$ for any $\beta \neq 0$.

7) A nonzero constant function f does not belong to $\mathscr{S}'_h$ because $\dot{S}_k f = f$, $\forall k \in \mathbb{Z}$, i.e., $\lim_{k\to-\infty} \dot{S}_k f \neq 0$. Indeed, we have for any $g \in \mathscr{S}$

$$\begin{aligned}
\langle \dot{S}_k f, g\rangle =& \left\langle \varphi_k \widehat{f}, \check{g}\right\rangle = (2\pi)^{n/2}\langle \varphi_k f\delta_0, \check{g}\rangle = (2\pi)^{n/2}\varphi_k(0)f\check{g}(0) \\
=&(2\pi)^{n/2} f\check{g}(0) = (2\pi)^{n/2}\langle f\delta_0, \check{g}\rangle = \langle f, g\rangle.
\end{aligned}$$

We note that this example implies that $\mathscr{S}'_h$ is not a closed subspace of $\mathscr{S}'$ for the topology of weak-* convergence, a fact that must be kept in mind in the applications. For example, taking $f \in \mathscr{S}(\mathbb{R}^n)$ with $f(0) = 1$ and constructing the sequence

$$f_k(x) = f\left(\frac{x}{k}\right) \in \mathscr{S}(\mathbb{R}^n) \subset \mathscr{S}'_h(\mathbb{R}^n),$$

we can prove

$$f_k(x) \xrightarrow{\mathscr{S}'(\mathbb{R}^n)} 1 \notin \mathscr{S}'_h(\mathbb{R}^n), \text{ as } k \to \infty.$$

Now, we redefine homogeneous Besov spaces that can be used in the context of PDEs.

Definition 10.4 (Realization of homogeneous Besov spaces). Let $s \in \mathbb{R}$, $1 \leqslant p, r \leqslant \infty$. The homogeneous Besov space $\dot{\mathcal{B}}^s_{p,r}$ is defined by

$$\dot{\mathcal{B}}^s_{p,r} = \left\{ f \in \mathscr{S}'_h(\mathbb{R}^n) : \|f\|_{\dot{\mathcal{B}}^s_{p,r}} := \|f\|_{\dot{B}^s_{p,r}} < \infty \right\}.$$

Proposition 10.5. *The space $\dot{\mathcal{B}}^s_{p,r}$ endowed with $\|\cdot\|_{\dot{\mathcal{B}}^s_{p,r}}$ is a normed space.*

Proof. It is clear that $\|\cdot\|_{\dot{\mathcal{B}}^s_{p,r}}$ is a seminorm. Assume that for some $f \in \mathscr{S}'_h$, we have $\|f\|_{\dot{\mathcal{B}}^s_{p,r}} = 0$. This implies that $\operatorname{supp} \widehat{f} \subset \{0\}$, and thus, for any $k \in \mathbb{Z}$, we have $\dot{S}_k f = f$. As $f \in \mathscr{S}'_h$, we conclude that $f = 0$. ❑

Remark 10.6. The definition of the realized Besov space $\dot{\mathcal{B}}^s_{p,r}$ is independent of the function φ used for defining the blocks $\dot{\Delta}_k$ and changing φ yields an equivalent norm. Indeed, if $\tilde{\varphi}$ is another dyadic partition of unity, then an integer N_0 exists such that $|k-k'| \geqslant N_0$ implies that $\operatorname{supp} \tilde{\varphi}(2^{-k}\cdot) \cap \operatorname{supp} \varphi(2^{-k'}\cdot) = \emptyset$. Thus,

$$\begin{aligned}2^{ks}\|\tilde{\varphi}(2^{-k}D)f\|_p =& 2^{ks}\left\|\sum_{|k-k'|\leqslant N_0} \tilde{\varphi}(2^{-k}D)\dot{\Delta}_{k'}f\right\|_p \\ \lesssim& 2^{N_0|s|}\sum_{k'}\chi_{[-N_0,N_0]}(k-k')2^{k's}\|\dot{\Delta}_{k'}f\|_p,\end{aligned}$$

which implies the result by Young's inequality. We also note that the previous embedding relations for $\dot{B}^s_{p,r}$ are valid for $\dot{\mathcal{B}}^s_{p,r}$.

The (realized) homogeneous Besov spaces have nice scaling properties. Indeed, if f is a tempered distribution, then consider the tempered distribution f_N defined by $f_N := f(2^N\cdot)$. We have the following proposition.

Proposition 10.7 (Scaling properties)**.** *Let $N \in \mathbb{N}_0$ and $f \in \mathscr{S}'_h(\mathbb{R}^n)$. Then, $\|f\|_{\dot{\mathcal{B}}^s_{p,r}}$ is finite iff $\|f_N\|_{\dot{\mathcal{B}}^s_{p,r}}$ is finite. Moreover, we have*

$$\|f_N\|_{\dot{\mathcal{B}}^s_{p,r}} = 2^{N(s-n/p)}\|f\|_{\dot{\mathcal{B}}^s_{p,r}}.$$

Proof. By the definition of $\dot{\Delta}_k$, we obtain

$$\begin{aligned}\dot{\Delta}_k f_N(x) =& \left(\varphi(2^{-k}\xi)\widehat{f(2^N x)}(\xi)\right)^\vee(x)\\ =& \left(\varphi(2^{-k}\xi)2^{-nN}\widehat{f}(2^{-N}\xi)\right)^\vee(x)\\ =& \left(\varphi(2^{-(k-N)}\xi)\widehat{f}(\xi)\right)^\vee(2^N x) = \dot{\Delta}_{k-N}f(2^N x).\end{aligned}$$

It turns out that $\|\dot{\Delta}_k f_N\|_p = 2^{-nN/p}\|\dot{\Delta}_{k-N}f\|_p$. We deduce from this that

$$2^{ks}\|\dot{\Delta}_k f_N\|_p = 2^{N(s-n/p)}2^{(k-N)s}\|\dot{\Delta}_{k-N}f\|_p,$$

and the proposition follows immediately by summation. ❑

In contrast with the standard function spaces (e.g., general Sobolev space H^s or L^p spaces with $p<\infty$), (realized) homogeneous Besov spaces contain nontrivial homogeneous distributions. This is illustrated by the following proposition.

Proposition 10.8. *Let $\sigma \in (0,n)$. Then for any $p \in [1,\infty]$, it holds*

$$\frac{1}{|x|^\sigma} \in \dot{\mathcal{B}}^{\frac{n}{p}-\sigma}_{p,\infty}(\mathbb{R}^n). \tag{10.6}$$

Proof. By Theorem 9.18, it is sufficient to prove that $\rho_\sigma := |\cdot|^{-\sigma} \in \dot{\mathcal{B}}^{n-\sigma}_{1,\infty} \hookrightarrow \dot{\mathcal{B}}^{\frac{n}{p}-\sigma}_{p,\infty}$. To do so, we introduce $\chi \in \mathscr{D}$ with value 1 near the unit ball and write

$$\rho_\sigma = \rho_0 + \rho_1, \text{ with } \rho_0(x) := \chi(x)|x|^{-\sigma} \quad \text{ and } \rho_1(x) := (1-\chi(x))|x|^{-\sigma}.$$

It is obvious that $\rho_0 \in L^1$ and that $\rho_1 \in L^q$ whenever $q > n/\sigma$. This implies that $\rho_\sigma \in \mathscr{S}'_h$. The homogeneity of ρ_σ gives

$$\begin{aligned}\dot{\Delta}_k\rho_\sigma =&(2\pi)^{-n/2}\left(\varphi(2^{-k}\cdot)\right)^\vee * \rho_\sigma = (2\pi)^{-n/2}2^{kn}\check{\varphi}(2^k\cdot) * \rho_\sigma\\ =&(2\pi)^{-n/2}2^{k(n+\sigma)}\check{\varphi}(2^k\cdot) * \rho_\sigma(2^k\cdot) = 2^{k\sigma}(\dot{\Delta}_0\rho_\sigma)(2^k\cdot).\end{aligned}$$

Therefore, $\|\dot{\Delta}_k\rho_\sigma\|_1 = 2^{k(\sigma-n)}\|\dot{\Delta}_0\rho_\sigma\|_1$, which reduces the problem to proving that $\dot{\Delta}_0\rho_\sigma \in L^1$. Due to $\rho_0 \in L^1$, we have $\dot{\Delta}_0\rho_0 \in L^1$ by the continuity of $\dot{\Delta}_0$ on Lebesgue spaces. By Bernstein's inequality, we obtain

$$\|\dot{\Delta}_0\rho_1\|_1 \leqslant C_k \sup_{|\alpha|=k} \|\partial^\alpha\dot{\Delta}_0\rho_1\|_1 \leqslant C_k \sup_{|\alpha|=k} \|\partial^\alpha\rho_1\|_1.$$

From Leibniz's formula, $\partial^\alpha\rho_1 - (1-\chi)\partial^\alpha\rho_\sigma \in \mathscr{D}$. Then, we complete the proof by choosing $k > n-\sigma$ for which $|\partial^\alpha|x|^{-\sigma}| \lesssim |x|^{-\sigma-k}$ is integrable outside the unit ball. ❑

The following lemma provides a useful criterion for determining whether the sum of a series belongs to a homogeneous Besov space.

Lemma 10.9. *Let $s \in \mathbb{R}$, $1 \leqslant p,r \leqslant \infty$ and $\mathbb{A}$ be an annulus in $\mathbb{R}^n$. Assume that $\{f_k\}_{k\in\mathbb{Z}}$ is a sequence of functions satisfying*

$$\operatorname{supp}\widehat{f}_k \subset 2^k\mathbb{A}, \quad \text{and} \quad \left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})} < \infty.$$

If the series $\sum\limits_{k\in\mathbb{Z}} f_k$ converges in $\mathscr{S}'$ to some $f \in \mathscr{S}'_h$, then $f \in \dot{\mathcal{B}}^s_{p,r}$ and

$$\|f\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C_s \left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})}.$$

Proof. It is clear that there exists some positive integer N_0 such that $\dot{\Delta}_j f_k = 0$ for $|j-k| \geqslant N_0$. Hence,

$$\|\dot{\Delta}_j f\|_p = \left\|\sum_{|j-k|<N_0} \dot{\Delta}_j f_k\right\|_p \lesssim \sum_{|j-k|<N_0} \|f_k\|_p.$$

Therefore, we obtain that

$$2^{js}\|\dot{\Delta}_j f\|_p \lesssim \sum_{|j-k|<N_0} 2^{(j-k)s}2^{ks}\|f_k\|_p = C\sum_{k\in\mathbb{Z}} 2^{(j-k)s}\chi_{|j-k|<N_0}(k)2^{ks}\|f_k\|_p.$$

Thus, by Young's inequality, we obtain

$$\|f\|_{\dot{\mathcal{B}}^s_{p,r}} \lesssim \left(\sum_{j=-N_0+1}^{N_0-1} 2^{js}\right)\left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})} \lesssim_s \left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})}.$$

As $f \in \mathscr{S}'_h$ by assumption, this proves the lemma. ❑

Remark 10.10. The above convergence assumption concerns $\{f_k\}_{k<0}$. We note that if (s,p,r) satisfies the condition (10.1), i.e.,

$$s < \frac{n}{p}, \quad \text{or} \quad s = \frac{n}{p} \text{ and } r = 1, \tag{10.1}$$

then, from the proof of Theorem 10.1, we have

$$\lim_{j\to-\infty}\sum_{k<j} f_k = 0 \text{ in } L^\infty.$$

Hence, $\sum_{k\in\mathbb{Z}} f_k$ converges to some $f \in \mathscr{S}'$, and $\dot{S}_k f$ tends to 0 when k goes to $-\infty$. In particular, we have $f \in \mathscr{S}'_h$.

Lemma 10.9 will enable us to establish the following important topological properties of homogeneous Besov spaces.

Theorem 10.11. *Let $s_1, s_2 \in \mathbb{R}$ and $1 \leqslant p_1, p_2, r_1, r_2 \leqslant \infty$. Assume that (s_1,p_1,r_1) satisfies the condition (10.1). Then the space $\dot{B}^{s_1}_{p_1,r_1} \cap \dot{B}^{s_2}_{p_2,r_2}$ endowed with the norm $\|\cdot\|_{\dot{B}^{s_1}_{p_1,r_1}} + \|\cdot\|_{\dot{B}^{s_2}_{p_2,r_2}}$ is complete and satisfies the Fatou property: If $\{f_k\}_{k\in\mathbb{N}_0}$ is a bounded sequence of $\dot{B}^{s_1}_{p_1,r_1} \cap \dot{B}^{s_2}_{p_2,r_2}$, then there exists a subsequence $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ and an element $f \in \dot{B}^{s_1}_{p_1,r_1} \cap \dot{B}^{s_2}_{p_2,r_2}$ such that*

$$\lim_{j\to\infty} f_{k_j} = f \text{ in } \mathscr{S}', \text{ and } \|f\|_{\dot{B}^{s_l}_{p_l,r_l}} \leqslant C \liminf_{j\to\infty}\|f_{k_j}\|_{\dot{B}^{s_l}_{p_l,r_l}} \text{ for } l = 1,2.$$

Proof. We first prove the Fatou property. According to Bernstein's inequality, for any $m \in \mathbb{Z}$, the C^∞ sequence $\{\dot{\Delta}_m f_k\}_{k\in\mathbb{N}_0}$ is bounded (uniformly in k) in L^p for any $\min(p_1,p_2) \leqslant p \leqslant \infty$, especially in $L^{\min(p_1,p_2)} \cap L^\infty$, since

$$\|\dot{\Delta}_m f_k\|_p \leqslant 2^{mn(1/p_l-1/p)}2^{-ms_l}2^{ms_l}\|\dot{\Delta}_m f_k\|_{p_l}, \text{ for } p \geqslant p_l,\ l = 1 \text{ or } 2.$$

Cantor's diagonal process supplies a subsequence $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ and a sequence $\{\tilde{f}_m\}_{m\in\mathbb{Z}} \subset C^\infty$ with Fourier transform supported in $2^m\mathbb{A}$ (where $\mathbb{A}$ has been defined in (9.3)) such that, for any $m \in \mathbb{Z}$, $\phi \in \mathscr{S}$, and $l = 1,2$,

$$\lim_{j\to\infty}\langle\dot{\Delta}_m f_{k_j},\phi\rangle = \langle\tilde{f}_m,\phi\rangle, \quad \text{and } \|\tilde{f}_m\|_{p_l} \leqslant \liminf_{j\to\infty}\|\dot{\Delta}_m f_{k_j}\|_{p_l}.$$

Now, the sequence $\left(\{2^{ms_l}\|\dot{\Delta}_m f_{k_j}\|_{p_l}\}_m\right)_{j\in\mathbb{N}_0}$ is bounded in $\ell^{r_l}(\mathbb{Z})$. Hence, there exists an element $\{\tilde{c}^l_m\}_{m\in\mathbb{Z}}$ of ℓ^{r_l} such that (up to an omitted extraction) we have, for any test sequence $\{d_m\}_{m\in\mathbb{Z}} \in c_{00}$[a] of nonnegative real numbers different from 0 for only a finite number of indices m,

$$\lim_{j\to\infty} \sum_{m\in\mathbb{Z}} 2^{ms_l}\|\dot{\Delta}_m f_{k_j}\|_{p_l} d_m = \sum_{m\in\mathbb{Z}} \tilde{c}^l_m d_m, \text{ and } \|\{\tilde{c}^l_m\}_m\|_{\ell^{r_l}} \leqslant \liminf_{j\to\infty} \|f_{k_j}\|_{\dot{B}^{s_l}_{p_l,r_l}}.$$

Passing to the limit in the sum and using the Hölder inequality gives that

$$\{2^{ms_l}\|\tilde{f}_m\|_{p_l}\}_m \in \ell^{r_l}(\mathbb{Z}).$$

From the definition of $\tilde{f}_m$, we easily deduce that $\mathscr{F}\tilde{f}_m$ is supported in the annulus $2^m\mathbb{A}$. As (s_1, p_1, r_1) satisfies (10.1), Lemma 10.9 thus guarantees that the series $\sum_{m\in\mathbb{Z}} \tilde{f}_m$ converges to some $f \in \mathscr{S}'_h$. By Proposition 9.6, we have, for all $M < N$ and $\phi \in \mathscr{S}$,

$$\langle \sum_{m=M}^{N} \dot{\Delta}_m f, \phi\rangle = \langle \sum_{m=M}^{N} \sum_{|m'-m|\leqslant 1} \dot{\Delta}_m \tilde{f}_{m'}, \phi\rangle.$$

Hence, by the definition of $\tilde{f}_m$ and Proposition 9.6 again, we have

$$\sum_{m=M}^{N} \dot{\Delta}_m f = \lim_{j\to\infty} \sum_{m=M}^{N} \sum_{|m'-m|\leqslant 1} \dot{\Delta}_m \dot{\Delta}_{m'} f_{k_j} = \lim_{j\to\infty} \sum_{m=M}^{N} \dot{\Delta}_m f_{k_j}, \quad \text{in } \mathscr{S}'.$$

Since the condition (10.1) is satisfied by (s_1, p_1, r_1), and $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ is bounded in $\dot{B}^{s_1}_{p_1,r_1}$, Lemma 10.9 ensures that $\dot{S}_M f_{k_j}$ tends uniformly to 0 when M goes to $-\infty$. Similarly, $(1 - \dot{S}_N) f_{k_j}$ tends uniformly to 0 in $\dot{B}^{s_2}_{p_2,r_2}$ as $N \to \infty$. Hence, f is indeed the limit of $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ in $\mathscr{S}'$, which completes the proof of the Fatou property.

We will now check that $\dot{B}^{s_1}_{p_1,r_1} \cap \dot{B}^{s_2}_{p_2,r_2}$ is complete. Consider a Cauchy sequence $\{f_k\}_{k\in\mathbb{N}_0}$. This sequence is of course bounded, so there exists some f in $\dot{B}^{s_1}_{p_1,r_1} \cap \dot{B}^{s_2}_{p_2,r_2}$ and a subsequence $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ such that $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ converges to f in $\mathscr{S}'$. For any positive ε, there exists an integer j_ε such that

$$j \geqslant j' \geqslant j_\varepsilon \Longrightarrow \|f_{k_{j'}} - f_{k_j}\|_{\dot{B}^{s_1}_{p_1,r_1}} + \|f_{k_{j'}} - f_{k_j}\|_{\dot{B}^{s_2}_{p_2,r_2}} < \varepsilon,$$

the Fatou property for $\{f_{k_{j'}} - f_{k_j}\}_{j\in\mathbb{N}_0}$ ensures that

$$\forall j' \geqslant j_\varepsilon, \quad \|f_{k_{j'}} - f\|_{\dot{B}^{s_1}_{p_1,r_1}} + \|f_{k_{j'}} - f\|_{\dot{B}^{s_2}_{p_2,r_2}} < C\varepsilon.$$

Hence, $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ tends to f in $\dot{B}^{s_1}_{p_1,r_1} \cap \dot{B}^{s_2}_{p_2,r_2}$. This completes the proof. ❑

[a] c_{00} is the space of all infinite sequences with only a finite number of nonzero terms (sequence with finite support), which is dense in ℓ^r for any $r \in [1, \infty)$.

In the case of negative indices of regularity, homogeneous Besov spaces may be characterized in terms of operator $\dot{S}_k$, as follows.

Theorem 10.12. *Let $s < 0$ and $1 \leqslant p, r \leqslant \infty$. Assume that $f \in \mathscr{S}'_h(\mathbb{R}^n)$; then, $f \in \dot{\mathrm{B}}^s_{p,r}(\mathbb{R}^n)$ iff*

$$\left\{2^{ks}\|\dot{S}_k f\|_p\right\}_{k\in\mathbb{Z}} \in \ell^r.$$

Moreover, we have

$$2^{-|s|-1}\|f\|_{\dot{\mathrm{B}}^s_{p,r}} \leqslant \left\|\{2^{ks}\|\dot{S}_k f\|_p\}_k\right\|_{\ell^r} \leqslant \frac{1}{|s|\ln 2}\|f\|_{\dot{\mathrm{B}}^s_{p,r}}.$$

Proof. We write

$$2^{ks}\|\dot{\Delta}_k f\|_p \leqslant 2^{ks}(\|\dot{S}_{k+1}f\|_p + \|\dot{S}_k f\|_p) \leqslant 2^{-s}2^{(k+1)s}\|\dot{S}_{k+1}f\|_p + 2^{ks}\|\dot{S}_k f\|_p.$$

Then, the left inequality is proved. To obtain the right inequality, we write

$$\begin{aligned}2^{ks}\|\dot{S}_k f\|_p \leqslant & 2^{ks}\sum_{j\leqslant k-1}\|\dot{\Delta}_j f\|_p = \sum_{j\leqslant k-1} 2^{(k-j)s}2^{js}\|\dot{\Delta}_j f\|_p \\ &= \sum_{j\in\mathbb{Z}} 2^{(k-j)s}\chi_{\{k-j\geqslant 1\}}2^{js}\|\dot{\Delta}_j f\|_p,\end{aligned}$$

where $\chi_{\{k-j\geqslant 1\}} = 1$ if $k-j \geqslant 1$ and is zero otherwise. As s is negative, the result follows by Young's inequality for ℓ^r spaces. Precisely, the coefficient is

$$\sum_{j\geqslant 1} 2^{js} = \frac{2^s}{1-2^s} = \frac{1}{2^{|s|}-1} \leqslant \frac{1}{|s|\ln 2},$$

since $2^{|s|} = \sum_{k=0}^{\infty}\frac{(|s|\ln 2)^k}{k!}$ implies $2^{|s|} - 1 \geqslant |s|\ln 2$. ❑

Theorem 10.13. *Let $s \in \mathbb{R}$, $p, r \in [1,\infty]$. Then, $\dot{\mathrm{B}}^s_{p,r}(\mathbb{R}^n)$ is a Banach space when $s < \frac{n}{p}$. In addition, $\dot{\mathrm{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n)$ is also a Banach space.*

Proof. By Proposition 10.5, both $\dot{\mathrm{B}}^s_{p,r}(\mathbb{R}^n)$ and $\dot{\mathrm{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n)$ are normed spaces.

Step 1. To prove the embedding: $\dot{\mathrm{B}}^s_{p,r}(\mathbb{R}^n) \hookrightarrow \mathscr{S}'$ for $s < \frac{n}{p}$, and $\dot{\mathrm{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n) \hookrightarrow \mathscr{S}'$.

We know that $\dot{\mathrm{B}}^s_{p,r}(\mathbb{R}^n) \subset \mathscr{S}'$ for $s < \frac{n}{p}$ and $\dot{\mathrm{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n) \subset \mathscr{S}'$ by the definition of Besov spaces due to $\mathscr{S}'_h \subset \mathscr{S}'$, but the embedding relation in the topological sense needs to be proven. From Bernstein's inequality, it follows that

$$\|\dot{\Delta}_k u\|_\infty \lesssim 2^{k\frac{n}{p}}\|\dot{\Delta}_k u\|_p. \tag{10.7}$$

For $u \in \dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}$, we have

$$\|u\|_\infty \leqslant \sum_{k\in\mathbb{Z}} \|\dot{\Delta}_k u\|_\infty \lesssim \sum_{k\in\mathbb{Z}} 2^{k\frac{n}{p}} \|\dot{\Delta}_k u\|_p = C\|u\|_{\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}},$$

which yields $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1} \hookrightarrow L^\infty \hookrightarrow \mathscr{S}'$.

For $s < \frac{n}{p}$, we first consider the part of low frequencies $k < 0$. For any $f \in \mathscr{S}$, we obtain

$$\begin{aligned} |\langle \dot{\Delta}_k u, f\rangle| \leqslant &\|\dot{\Delta}_k u\|_\infty \|f\|_1 \lesssim 2^{k\frac{n}{p}} \|\dot{\Delta}_k u\|_p \|f\|_1 \\ \lesssim &2^{k\left(\frac{n}{p}-s\right)} \|u\|_{\dot{\mathcal{B}}^s_{p,\infty}} \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1}|f(x)|. \end{aligned} \tag{10.8}$$

Thus,

$$\left|\left\langle \sum_{k<0} \dot{\Delta}_k u, f \right\rangle\right| \lesssim \|u\|_{\dot{\mathcal{B}}^s_{p,r}} \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1}|f(x)|.$$

For high frequencies $k \geqslant 0$, we can use, as in (9.9),

$$\dot{\Delta}_k u = 2^{-kl} \sum_{|\alpha|=l} \partial^\alpha (2^{kn} g_\alpha(2^k\cdot) * \dot{\Delta}_k u), \tag{10.9}$$

$$g_\alpha := (2\pi)^{-n/2} a_\alpha \left(\frac{(-i\xi)^\alpha}{|\xi|^{2l}} \varphi(\xi)\right)^\vee. \tag{10.10}$$

Then, it holds for $l \in \mathbb{N}_0$ and any $f \in \mathscr{S}$,

$$\begin{aligned} \langle \dot{\Delta}_k u, f\rangle =&2^{-kl} \sum_{|\alpha|=l} \langle \partial^\alpha (2^{kn} g_\alpha(2^k\cdot) * \dot{\Delta}_k u), f\rangle \\ =&2^{-kl} \sum_{|\alpha|=l} \langle \dot{\Delta}_k u, 2^{kn} g_\alpha(-2^k\cdot) * (-\partial)^\alpha f\rangle \\ \lesssim&\|\dot{\Delta}_k u\|_\infty 2^{-kl} \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|=l}} (1+|x|)^{n+1}|\partial^\alpha f(x)| \\ \lesssim&2^{k\left(\frac{n}{p}-s-l\right)} 2^{ks} \|\dot{\Delta}_k u\|_p \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|=l}} (1+|x|)^{n+1}|\partial^\alpha f(x)|. \end{aligned}$$

Thus, for large $l > \frac{n}{p} - s$, it follows that

$$\left|\left\langle \sum_{k\geqslant 0} \dot{\Delta}_k u, f \right\rangle\right| \lesssim \|u\|_{\dot{\mathcal{B}}^s_{p,r}} \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|=l}} (1+|x|)^{n+1}|\partial^\alpha f(x)|.$$

Therefore, we obtain for any $f \in \mathscr{S}$

$$|\langle u, f\rangle| \leqslant \sum_{k\in\mathbb{Z}} |\langle \dot{\Delta}_k u, f\rangle| \lesssim \|u\|_{\dot{\mathcal{B}}^s_{p,r}} \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|\leqslant l}} (1+|x|)^{n+1}|\partial^\alpha f(x)|, \tag{10.11}$$

which implies $\dot{\mathcal{B}}^s_{p,r} \hookrightarrow \mathscr{S}'$.

Step 2. To prove the completeness. Let $\{u_\ell\}_{\ell\in\mathbb{N}}$ be a Cauchy sequence in $\dot{\mathcal{B}}^s_{p,r}$, where $s < \frac{n}{p}$ or $s = \frac{n}{p}$ and $r = 1$. Replacing u by $u_\ell - u_j$ in

(10.11), there exists a $u \in \mathscr{S}'$ such that

$$u_\ell \xrightarrow{\mathscr{S}'} u \in \mathscr{S}', \text{ as } \ell \to \infty.$$

Step 2.1. To show $u \in \mathscr{S}'_h$. For $s < \frac{n}{p}$, by the assumption, it is clear that $u_\ell \in \mathscr{S}'_h$ for any $\ell \in \mathbb{N}$. Similar to (10.8), we have for any $\ell \in \mathbb{N}$ and $j \in \mathbb{Z}$

$$\begin{aligned}|\langle \dot{S}_j u_\ell, f\rangle| \leqslant& \sum_{k\leqslant j-1} |\langle \dot{\Delta}_k u_\ell, f\rangle| \leqslant \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_\ell\|_\infty \|f\|_1 \\ \lesssim_s& 2^{j\left(\frac{n}{p}-s\right)} \sup_\ell \|u_\ell\|_{\dot{\mathcal{B}}^s_{p,r}} \|f\|_1.\end{aligned}$$

From $u_\ell \xrightarrow{\mathscr{S}'} u \in \mathscr{S}'$, it follows that

$$|\langle \dot{S}_j u, f\rangle| \lesssim_s 2^{j\left(\frac{n}{p}-s\right)} \sup_\ell \|u_\ell\|_{\dot{\mathcal{B}}^s_{p,r}} \|f\|_1, \quad \forall f \in \mathscr{S}.$$

Hence, we obtain

$$\lim_{j\to-\infty} \dot{S}_j u = 0, \quad \text{i.e., } u \in \mathscr{S}'_h.$$

For the case $s = \frac{n}{p}$ and $r = 1$, since $\{u_\ell\}$ is Cauchy in $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1} \hookrightarrow \dot{\mathcal{B}}^0_{\infty,1}$, we have $\forall \varepsilon > 0$, $\exists \ell_0 \in \mathbb{N}$, s.t. $\forall j \in \mathbb{Z}$ and $\ell \geqslant \ell_0$

$$\begin{aligned}\sum_{k\leqslant j-1} \|\dot{\Delta}_k u_\ell\|_\infty \leqslant& \sum_{k\leqslant j-1} \|\dot{\Delta}_k (u_\ell - u_{\ell_0})\|_\infty + \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty \\ \leqslant& \|u_\ell - u_{\ell_0}\|_{\dot{\mathcal{B}}^0_{\infty,1}} + \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty \\ \leqslant& \frac{\varepsilon}{2} + \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty.\end{aligned}$$

We can choose j_0 so small that

$$\sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty < \frac{\varepsilon}{2}, \quad \forall j \leqslant j_0.$$

Thus, it follows that for $u_\ell \in \mathscr{S}'_h$, we have, $\forall j \leqslant j_0$, $\forall \ell \geqslant \ell_0$

$$\|\dot{S}_j u_\ell\|_\infty \leqslant \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_\ell\|_\infty < \varepsilon. \tag{10.12}$$

Since $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1} \hookrightarrow \dot{\mathcal{B}}^0_{\infty,1} \hookrightarrow L^\infty$, $\{u_\ell\}_{\ell\in\mathbb{N}}$ is also a Cauchy sequence in L^∞, i.e., $u_\ell \to u \in L^\infty$ as $\ell \to \infty$. Taking $\ell \to \infty$ in (10.12) yields

$$\|\dot{S}_j u\|_\infty \leqslant \varepsilon, \quad \forall j \leqslant j_0,$$

which indicates $u \in \mathscr{S}'_h$.

Step 2.2. To show $u \in \dot{\mathcal{B}}^s_{p,r}$. From the definition of Besov spaces, it follows that for any fixed k, $\{\dot{\Delta}_k u_\ell\}_{\ell\in\mathbb{N}}$ is a Cauchy sequence in L^p. By the completeness of L^p, there exists $\bar{u}_k \in L^p$ such that

$$\lim_{\ell\to\infty} \|\dot{\Delta}_k u_\ell - \bar{u}_k\|_p = 0.$$

Since $u_\ell \xrightarrow{\mathscr{S}'} u$ as $\ell \to \infty$, we have $\dot\Delta_k u_\ell \xrightarrow{\text{a.e.}} \dot\Delta_k u$ as $\ell \to \infty$. Then, $\bar u_k = \dot\Delta_k u$. Thus,

$$\lim_{\ell\to\infty} 2^{ks}\|\dot\Delta_k u_\ell\|_p = 2^{ks}\|\dot\Delta_k u\|_p, \quad \forall k \in \mathbb{Z}.$$

For $\ell \in \mathbb{N}$, $\{2^{ks}\|\dot\Delta_k u_\ell\|_p\}$ is bounded in $\ell^r(\mathbb{Z})$, then so is $\{2^{ks}\|\dot\Delta_k u\|_p\}$. It follows that $u \in \dot{\mathcal{B}}^s_{p,r}$ from Lemma 10.9.

Step 2.3. To show the convergence in $\dot{\mathcal{B}}^s_{p,r}$. For any given $K > 0$, due to $\dot\Delta_k u_m \to \dot\Delta_k u$ in L^p as $m \to \infty$, we obtain

$$\left(\sum_{|k|\leqslant K}\left(2^{ks}\|\dot\Delta_k(u_\ell - u)\|_p\right)^r\right)^{\frac1r} = \lim_{m\to\infty}\left(\sum_{|k|\leqslant K}\left(2^{ks}\|\dot\Delta_k(u_\ell - u_m)\|_p\right)^r\right)^{\frac1r}.$$

Noticing that $\{u_\ell\}_{\ell\in\mathbb{N}}$ is Cauchy in $\dot{\mathcal{B}}^s_{p,r}$, thus, for any $\varepsilon > 0$, there exists a $\ell_0 \in \mathbb{N}$ independent of K such that for all $\ell > \ell_0$, we have

$$\left(\sum_{|k|\leqslant K}\left(2^{ks}\|\dot\Delta_k(u_\ell - u)\|_p\right)^r\right)^{\frac1r} < \varepsilon.$$

Taking $K \to \infty$, it yields that $u_\ell \to u$ in $\dot{\mathcal{B}}^s_{p,r}$ as $\ell \to \infty$. Thus, we complete the proof. ❏

Remark 10.14. The realization $\dot{\mathcal{B}}^s_{p,r}$ coincides with the general definition $\dot B^s_{p,r}$ when $s < n/p$, or $s = n/p$ and $r = 1$. However, if $s > n/p$ (or $s = n/p$ and $r > 1$), then $\dot{\mathcal{B}}^s_{p,r}$ is no longer a Banach space. This is due to a breakdown of convergence for low frequencies, the so-called infrared divergence.

Example 10.15. Let $\chi(\xi) \in \mathscr{D}(\mathbb{R})$ with value 1 when $|\xi| < 8/9$ and $\operatorname{supp}\chi = \{\xi : |\xi| \leqslant 9/10\}$. Define

$$\widehat{f_k}(\xi) = \begin{cases} \dfrac{\chi(\xi)}{\xi\ln|\xi|}, & |\xi| \geqslant 2^{-k}, \\ 0, & \text{otherwise.}\end{cases}$$

It is clear that for $k > \ell > 0$

$$\widehat{f_k}(\xi) - \widehat{f_\ell}(\xi) = \begin{cases} 0, & |\xi| \geqslant 2^{-\ell}, \\ \dfrac{1}{\xi\ln|\xi|}, & 2^{-k} < |\xi| < 2^{-\ell}, \\ 0, & |\xi| \leqslant 2^{-k}.\end{cases}$$

Thus, we have

$$\begin{aligned}\|f_k - f_\ell\|_{\dot{\mathcal{B}}^{1/2}_{2,\infty}} &= \sup_{j\in\mathbb{Z}} 2^{j/2}\|\dot\Delta_j(f_k - f_\ell)\|_2 \\ &= \sup_{j\in\mathbb{Z}} 2^{j/2}\left(\int_{2^{-k}<|\xi|<2^{-\ell}}\left|\frac{\varphi_j(\xi)}{\xi\ln|\xi|}\right|^2 d\xi\right)^{1/2}\end{aligned}$$

$$\leqslant \sup_{j\in\mathbb{Z}} 2^{j/2}\frac{1}{\ell \ln 2}\left(\int_{\mathbb{R}}\left|\frac{\varphi(2^{-j}\xi)}{\xi}\right|^2 d\xi\right)^{1/2}$$
$$=\frac{1}{\ell \ln 2}\left(\int_{\mathbb{R}}\left|\frac{\varphi(\xi)}{\xi}\right|^2 d\xi\right)^{1/2} \to 0, \text{ as } k,\ell\to\infty,$$

namely, $\{f_k\}$ is Cauchy in $\dot{\mathcal{B}}^{1/2}_{2,\infty}$. However, it holds

$$\lim_{k\to\infty}\widehat{f_k}(\xi) = \frac{\chi(\xi)}{\xi \ln|\xi|}, \quad \forall \xi > 0,$$

which is not integrable near $\{0\}$, therefore $\lim_{k\to\infty} f_k \notin \mathscr{S}'_h$ and then $\lim_{k\to\infty} f_k \notin \dot{\mathcal{B}}^{1/2}_{2,\infty}$.

Finally, we give the dual of realized homogeneous Besov spaces. Observe that in Littlewood-Paley theory, the duality on $\mathscr{S}'_h$ reads for $\phi \in \mathscr{S}$,

$$\langle u,\phi\rangle = \sum_{|k-j|\leqslant 1}\langle \dot{\Delta}_k u, \dot{\Delta}_j \phi\rangle = \sum_{|k-j|\leqslant 1}\int_{\mathbb{R}^n}\dot{\Delta}_k u(x)\dot{\Delta}_j\phi(x)dx.$$

For the L^p space, we can estimate the norm in $\dot{\mathcal{B}}^s_{p,r}$ by duality.

Proposition 10.16. *For all $s \in \mathbb{R}$ and $p,r \in [1,\infty]$,*

$$\begin{cases} \dot{\mathcal{B}}^s_{p,r}\times \dot{\mathcal{B}}^{-s}_{p',r'} \longrightarrow \mathbb{R} \\ \qquad (u,\phi) \mapsto \displaystyle\sum_{|k-j|\leqslant 1}\langle \dot{\Delta}_k u, \dot{\Delta}_j\phi\rangle \end{cases}$$

defines a continuous bilinear functional on $\dot{\mathcal{B}}^s_{p,r}\times \dot{\mathcal{B}}^{-s}_{p',r'}$. Let

$$Q^{-s}_{p',r'} := \left\{\phi\in \mathscr{S}\cap \dot{\mathcal{B}}^{-s}_{p',r'} : \|\phi\|_{\dot{\mathcal{B}}^{-s}_{p',r'}} \leqslant 1\right\}.$$

If $u \in \mathscr{S}'_h$, then we have for $p,r\in(1,\infty]$,

$$\|u\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C \sup_{\phi\in Q^{-s}_{p',r'}} \langle u,\phi\rangle.$$

Proof. For $|k-j|\leqslant 1$, by Hölder's inequality, we have

$$|\langle \dot{\Delta}_k u, \dot{\Delta}_j\phi\rangle| \leqslant 2^{|s|}2^{ks}\|\dot{\Delta}_k u\|_p 2^{-js}\|\dot{\Delta}_j\phi\|_{p'}.$$

Again using Hölder's inequality, we deduce that

$$|\langle u,\phi\rangle| \lesssim_s \|u\|_{\dot{\mathcal{B}}^s_{p,r}}\|\phi\|_{\dot{\mathcal{B}}^{-s}_{p',r'}}.$$

To prove the second part, for $N\in\mathbb{N}$, let

$$Q^{r'}_N := \left\{(\alpha_k)\in \ell^{r'}(\mathbb{Z}) : \|(\alpha_k)\|_{\ell^{r'}} \leqslant 1, \text{ with } \alpha_k = 0 \text{ for } |k|>N\right\}.$$

By the definition of the Besov norm and the dual properties of ℓ^r, we obtain

$$\begin{aligned}
\|u\|_{\dot{\mathcal{B}}^s_{p,r}} &= \sup_{N\in\mathbb{N}} \left\| \left(\chi_{|k|\leqslant N}(k) 2^{ks} \|\dot{\Delta}_k u\|_p\right)_k \right\|_{\ell^r} \\
&= \sup_{N\in\mathbb{N}} \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \|\dot{\Delta}_k u\|_p 2^{ks}\alpha_k \quad (\text{by duality of } \ell^r) \\
&= \sup_{N\in\mathbb{N}} \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} 2^{ks}\alpha_k \sup_{\substack{\tilde{\phi}\in\mathscr{S}\\ \|\tilde{\phi}\|_{p'}\leqslant 1}} \langle \dot{\Delta}_k u, \tilde{\phi}\rangle, \quad (\text{by duality of } L^p).
\end{aligned}$$

By definition of supremum, for $|k| \leqslant N$ and any $\varepsilon > 0$, there is a $\phi_k \in \mathscr{S}$ with $\|\phi_k\|_{p'} \leqslant 1$ such that

$$\sup_{\substack{\tilde{\phi}\in\mathscr{S}\\ \|\tilde{\phi}\|_{p'}\leqslant 1}} \langle \dot{\Delta}_k u, \tilde{\phi}\rangle < \langle \dot{\Delta}_k u, \phi_k\rangle + \frac{\varepsilon 2^{-ks}}{(1+|\alpha_k|)(1+|k|^2)}.$$

Let

$$\Phi_N := \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \alpha_k 2^{ks} \dot{\Delta}_k \phi_k.$$

Note that

$$\left(\sum_{k=1}^K a_k\right)^\alpha \leqslant K^{\max(0,\alpha-1)} \sum_{k=1}^K a_k^\alpha$$

for $\alpha \geqslant 0$ and $a_k \geqslant 0$ (cf. [Dan10, p.391]). Then, for $r' \in [1,\infty)$, we obtain

$$\begin{aligned}
\|\Phi_N\|_{\dot{\mathcal{B}}^{-s}_{p',r'}} &= \left(\sum_{j\in\mathbb{Z}} 2^{-jsr'} \left\| \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \alpha_k 2^{ks} \dot{\Delta}_j \dot{\Delta}_k \phi_k \right\|_{p'}^{r'} \right)^{1/r'} \\
&= \left(\sum_{j\in\mathbb{Z}} 2^{-jsr'} \left\| \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \chi_{[j-1,j+1]}(k) \alpha_k 2^{ks} \dot{\Delta}_j \dot{\Delta}_k \phi_k \right\|_{p'}^{r'} \right)^{1/r'} \\
&\lesssim \left(3^{r'-1} \sum_{j\in\mathbb{Z}} \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} |\alpha_k|^{r'} \chi_{[j-1,j+1]}(k) 2^{(k-j)sr'} \|\phi_k\|_{p'}^{r'} \right)^{1/r'} \\
&\lesssim \left(3^{r'-1} \sum_{j\in\mathbb{Z}} \sup_{(\alpha_k)\in Q^{r'}_N} \left(\sum_{|k|\leqslant N} |\alpha_k|^{r'} \right) \sup_{|k|\leqslant N} \chi_{[j-1,j+1]}(k) 2^{(k-j)sr'} \right)^{1/r'} \\
&\lesssim 2^{|s|} \left(3\cdot 3^{r'-1}\right)^{1/r'} \\
&\lesssim 32^{|s|},
\end{aligned}$$

which is independent of N.

Thus, for any N,

$$\left\| \left(\chi_{|k|\leqslant N}(k) 2^{ks} \|\dot{\Delta}_k u\|_p\right)_k \right\|_{\ell^r}$$

$$<\langle u,\Phi_N\rangle + \sup_{(\alpha_k)\in Q_N^{r'}} \sum_{|k|\leqslant N} 2^{ks}|\alpha_k| \frac{\varepsilon 2^{-ks}}{(1+|\alpha_k|)(1+|k|^2)}$$
$$\leqslant \langle u,\Phi_N\rangle + \varepsilon.$$

Therefore, we complete the proof. ❑

§ 10.2 More results on nonhomogeneous Besov spaces

For the nonhomogeneous Besov space, we have the following lemma, the proof of which is analogous to that of Lemma 10.9.

Lemma 10.17. *Let $s \in \mathbb{R}$, $1 \leqslant p,r \leqslant \infty$ and $\mathbb{A}$ be an annulus in $\mathbb{R}^n$. Assume that $\{f_k\}_{k\in\mathbb{N}_0}$ is a sequence of smooth functions satisfying*

$$\operatorname{supp}\widehat{f}_k \subset 2^k\mathbb{A}, \quad \textit{and} \quad \left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{N}_0)} < \infty.$$

Then, we have

$$f := \sum_{k\in\mathbb{N}_0} f_k \in B^s_{p,r}, \quad \textit{and } \|f\|_{B^s_{p,r}} \lesssim_s \left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{N}_0)}.$$

Remark 10.18. If $1 \leqslant r < \infty$, then for any $f \in B^s_{p,r}$, we have

$$\lim_{k\to\infty} \|S_k f - f\|_{B^s_{p,r}} = 0,$$

since

$$\lim_{k\to\infty} \sum_{j\geqslant k} 2^{jsr}\|\Delta_j f\|_p^r = 0, \quad \forall r \in [1,\infty),\ f \in B^s_{p,r}.$$

Theorem 10.19. *Let $s < 0$ and $1 \leqslant p,r \leqslant \infty$. Assume that $f \in \mathscr{S}'(\mathbb{R}^n)$; then, $f \in B^s_{p,r}(\mathbb{R}^n)$ iff*

$$\left\{2^{ks}\|S_k f\|_p\right\}_{k\in\mathbb{N}_0} \in \ell^r.$$

Moreover, for some constant C depending only on n, we have

$$2^{-|s|-1}\|f\|_{B^s_{p,r}} \leqslant \left\|\{2^{ks}\|S_k f\|_p\}_k\right\|_{\ell^r} \leqslant \frac{1}{|s|\ln 2}\|f\|_{B^s_{p,r}}.$$

Proof. The proof is very close to that of Theorem 10.12 and is thus omitted. ❑

Similar to Theorem 10.11, we have the following.

Theorem 10.20. *Let $s \in \mathbb{R}$ and $1 \leqslant p,r \leqslant \infty$. Then $B^s_{p,r}$ is a Banach space and satisfies the Fatou property, namely, if $\{f_k\}_{k\in\mathbb{N}_0}$ is a bounded sequence* *of $B^s_{p,r}$, then there exists an $f \in B^s_{p,r}$ and a subsequence $\{f_{k_j}\}_{j\in\mathbb{N}_0}$ such that*

$$\lim_{j\to\infty} f_{k_j} = f \quad \text{in } \mathscr{S}', \quad \text{and} \quad \|f\|_{B^s_{p,r}} \lesssim \liminf_{j\to\infty} \|f_{k_j}\|_{B^s_{p,r}}.$$

We will now examine the way Fourier multipliers act on nonhomogeneous Besov spaces. Before stating our result, we need to define the multipliers we are going to consider.

Definition 10.21. Let $m \in \mathbb{R}$. A smooth function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be an ***S^m-multiplier*** if, for each multi-index α, there exists a constant C_α such that

$$|\partial^\alpha f(\xi)| \lesssim_\alpha (1+|\xi|)^{m-|\alpha|}, \quad \forall \xi \in \mathbb{R}^n.$$

Theorem 10.22. *Let $m \in \mathbb{R}$ and f be an S^m-multiplier. Then for all $s \in \mathbb{R}$ and $1 \leqslant p, r \leqslant \infty$, operator $f(D)$ is continuous from $B^s_{p,r}$ to $B^{s-m}_{p,r}$.*

Proof. According to Lemma 10.17, it suffices to prove that

$$2^{k(s-m)}\|\Delta_k f(D)u\|_p \leqslant C2^{ks}\|\Delta_k u\|_p, \quad \forall k \geqslant -1. \tag{10.13}$$

Obviously, we can find the smooth function $\sigma_k := \tilde{\varphi}_k f$ satisfying the assumptions of Lemma 9.3, i.e., on supp $\tilde{\varphi}_k$

$$\begin{aligned}
|\partial^\alpha \sigma_k(\xi)| =& |\partial^\alpha(\tilde{\varphi}(2^{-k}\xi)f(\xi))| \\
\lesssim& \sum_{\alpha=\beta+\gamma} |2^{-k|\beta|}(\partial^\beta\tilde{\varphi})(2^{-k}\xi)\partial^\gamma f(\xi)| \\
\lesssim& \sum_{\alpha=\beta+\gamma} |2^{-k|\beta|}(\partial^\beta\tilde{\varphi})(2^{-k}\xi)(1+|\xi|)^{m-|\gamma|}| \\
\lesssim& \sum_{\beta\leqslant\alpha} |(\partial^\beta\tilde{\varphi})(2^{-k}\xi)|2^{k(m-|\alpha|)} \\
\lesssim& 2^{k(m-|\alpha|)},
\end{aligned}$$

and such that

$$\Delta_k f(D)u = \tilde{\Delta}_k f(D)\Delta_k u = \sigma_k(D)\Delta_k u, \quad \forall k \geqslant 0.$$

Hence, Lemma 9.3 guarantees that (10.13) is satisfied for $k \geqslant 0$.

Next, introducing $\theta \in \mathscr{D}(\mathbb{R}^n)$ such that $\theta \equiv 1$ on supp ψ, we see that

$$\Delta_{-1} f(D)u = (\theta f)(D)\Delta_{-1}u.$$

As $(\theta f)^\vee \in L^1$, Young's inequality yields (10.13) for $k = -1$. ❏

§ 10.3 Paraproduct and Bony decomposition

In this section, we study the way that the product acts on Besov spaces. Let f and g be tempered distributions in $\mathscr{S}'_h(\mathbb{R}^n)$. We have

$$f = \sum_k \dot{\Delta}_k f \text{ and } g = \sum_j \dot{\Delta}_j g,$$

hence, at least formally,

$$fg = \sum_{k,j} \dot{\Delta}_k f \dot{\Delta}_j g.$$

Paradifferential calculus is a mathematical tool for splitting the above sum into three parts:

1. The first part concerns the indices (k, j) for which the size of $\operatorname{supp} \widehat{\dot{\Delta}_k f}$ is small compared to the size of $\operatorname{supp} \widehat{\dot{\Delta}_j g}$ (i.e., $k < j - N_0$ for some suitable positive integer N_0).
2. The second part contains the indices corresponding to those frequencies of f that are large compared with the frequencies of g (i.e., $k > j + N_0$).
3. In the last part, we keep the indices (k, j) for which $\operatorname{supp} \widehat{\dot{\Delta}_k f}$ and $\operatorname{supp} \widehat{\dot{\Delta}_j g}$ have comparable sizes (i.e., $|k - j| \leqslant N_0$).

The suitable choice for N_0 depends on the assumptions made on the support of the function φ used in the definition of the Littlewood-Paley decomposition, i.e., (9.2).

In what follows, we shall always assume that φ has been chosen according to (9.2) so that taking $N_0 = 1$ is appropriate in view of Proposition 9.6. This leads to the following definition.

Definition 10.23. The homogeneous paraproduct of g by f is defined as follows:

$$\dot{T}_f g := \sum_j \dot{S}_{j-1} f \dot{\Delta}_j g.$$

The homogeneous remainder of f and g is defined by

$$\dot{R}(f, g) = \sum_{|k-j| \leqslant 1} \dot{\Delta}_k f \dot{\Delta}_j g.$$

Remark 10.24. It can be checked that $\dot{T}_f g$ makes sense in $\mathscr{S}'$ whenever f and g are in $\mathscr{S}'_h$ and that $\dot{T} : (f, g) \mapsto \dot{T}_f g$ is a bilinear operator. Of course, the remainder operator $\dot{R} : (f, g) \mapsto \dot{R}(f, g)$, when restricted to sufficiently smooth distributions, is also bilinear.

The main motivation for using operators $\dot{T}$ and $\dot{R}$ is that, at least formally, the following so-called ***Bony's paraproduct decomposition*** holds true:

$$fg = \dot{T}_f g + \dot{T}_g f + \dot{R}(f,g). \tag{10.14}$$

Therefore, to understand the product operators in Besov spaces, it suffices to investigate the continuity properties of the operators $\dot{T}$ and $\dot{R}$.

For simplicity, it will be understood that whenever the expressions $\dot{T}_f g$ or $\dot{R}(f,g)$ appear in the text, the series with general terms

$$\dot{S}_{j-1} f \dot{\Delta}_j g, \text{ or } \sum_{|\nu|\leqslant 1} \dot{\Delta}_{j-\nu} f \dot{\Delta}_j g$$

converges to some tempered distribution that belongs to $\mathscr{S}'_h$.

Similarly, we can define nonhomogeneous paraproducts as follows.

Definition 10.25. The nonhomogeneous paraproduct of g by f is defined by

$$T_f g := \sum_j S_{j-1} f \Delta_j g.$$

The nonhomogeneous remainder of f and g is defined by

$$R(f,g) = \sum_{|k-j|\leqslant 1} \Delta_k f \Delta_j g.$$

At least formally, the operators T and R are bilinear, and we have the following ***Bony's paraproduct decomposition***

$$fg = T_f g + T_g f + R(f,g). \tag{10.15}$$

We shall sometimes also use the following simplified decomposition

$$fg = T_f g + T'_g f, \quad \text{with } T'_g f := \sum_j S_{j+2} g \Delta_j f. \tag{10.16}$$

We can now state our main result concerning continuity of the homogeneous paraproduct operator $\dot{T}$.

Theorem 10.26. *Let $s \in \mathbb{R}$ and $1 \leqslant p, r \leqslant \infty$. Then, for any $(f,g) \in L^\infty \times \dot{\mathcal{B}}^s_{p,r}$, we have*

$$\|\dot{T}_f g\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C^{1+|s|} \|f\|_\infty \|g\|_{\dot{\mathcal{B}}^s_{p,r}}.$$

Moreover, let $s_1 < 0$, $s_2 \in \mathbb{R}$ and $1 \leqslant p, r_1, r_2 \leqslant \infty$, then, we have, for any $(f,g) \in \dot{\mathcal{B}}^{s_1}_{\infty,r_1} \times \dot{\mathcal{B}}^{s_2}_{p,r_2}$,

$$\|\dot{T}_f g\|_{\dot{\mathcal{B}}^{s_1+s_2}_{p,r}} \leqslant \frac{C^{1+|s_1+s_2|}}{|s_1|} \|f\|_{\dot{\mathcal{B}}^{s_1}_{\infty,r_1}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{p,r_2}}, \quad \text{with } \frac{1}{r} := \min\left(1, \frac{1}{r_1} + \frac{1}{r_2}\right).$$

Proof. From Proposition 9.6 and Proposition 9.8, we have

$$\begin{aligned}
\|\dot{T}_f g\|_{\dot{\mathcal{B}}^s_{p,r}}^r =& \sum_{k\in\mathbb{Z}} \left(2^{sk} \left\|\dot{\Delta}_k\left(\sum_{j\in\mathbb{Z}} \dot{S}_{j-1} f \dot{\Delta}_j g\right)\right\|_p\right)^r \\
\leqslant& \sum_{k\in\mathbb{Z}} \left(\sum_{-2\leqslant k-j\leqslant 1} 2^{sk} \|\dot{\Delta}_k (\dot{S}_{j-1} f \dot{\Delta}_j g)\|_p\right)^r \\
\leqslant& C^r 4^{r-1} \sum_{k\in\mathbb{Z}} \sum_{-2\leqslant k-j\leqslant 1} \left(2^{s(k-j)} 2^{sj} \|f\|_\infty \|\dot{\Delta}_j g\|_p\right)^r \\
=& C^r 4^{r-1} \|f\|_\infty^r \sum_{j\in\mathbb{Z}} \sum_{-2\leqslant k-j\leqslant 1} 2^{s(k-j)r} \left(2^{sj} \|\dot{\Delta}_j g\|_p\right)^r \\
\leqslant& C^r 4^r 2^{2|s|r} \|f\|_\infty^r \|g\|_{\dot{\mathcal{B}}^s_{p,r}}^r,
\end{aligned}$$

which yields

$$\|\dot{T}_f g\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant 4C \cdot 2^{2|s|} \|f\|_\infty \|g\|_{\dot{\mathcal{B}}^s_{p,r}}.$$

Similarly, we obtain

$$\begin{aligned}
\|\dot{T}_f g\|_{\dot{\mathcal{B}}^{s_1+s_2}_{p,r}} =& \left(\sum_{k\in\mathbb{Z}} \left(2^{k(s_1+s_2)} \left\|\dot{\Delta}_k\left(\sum_{j\in\mathbb{Z}} \dot{S}_{j-1} f \dot{\Delta}_j g\right)\right\|_p\right)^r\right)^{1/r} \\
\leqslant& \left(\sum_{k\in\mathbb{Z}} \left(\sum_{-2\leqslant k-j\leqslant 1} 2^{k(s_1+s_2)} \|\dot{\Delta}_k (\dot{S}_{j-1} f \dot{\Delta}_j g)\|_p\right)^r\right)^{1/r} \\
\leqslant& C 4^{1-1/r} \left(\sum_{k\in\mathbb{Z}} \sum_{-2\leqslant k-j\leqslant 1} \left(2^{(k-j)s_2} 2^{ks_1} 2^{js_2} \|\dot{S}_{j-1} f\|_\infty \|\dot{\Delta}_j g\|_p\right)^r\right)^{1/r} \\
=& C 4^{1-1/r} \left(\sum_{j\in\mathbb{Z}} \sum_{-2\leqslant k-j\leqslant 1} \left(2^{(k-j)s_2} 2^{ks_1} 2^{js_2} \|\dot{S}_{j-1} f\|_\infty \|\dot{\Delta}_j g\|_p\right)^r\right)^{1/r}.
\end{aligned}$$

For the case $\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} \leqslant 1$, by Hölder's inequality and Theorem 10.12, we have

$$\begin{aligned}
&\|\dot{T}_f g\|_{\dot{\mathcal{B}}^{s_1+s_2}_{p,r}} \\
\leqslant& 4^{1-1/r} \left(\sum_{j\in\mathbb{Z}} \sum_{-2\leqslant k-j\leqslant 1} \left(2^{(k-j)(s_1+s_2)+s_1} 2^{(j-1)s_1} \|\dot{S}_{j-1} f\|_\infty\right)^{r_1}\right)^{1/r_1} \\
&\cdot \left(4 \sum_{j\in\mathbb{Z}} \left(2^{js_2} \|\dot{\Delta}_j g\|_p\right)^{r_2}\right)^{1/r_2} \\
\leqslant& \frac{C}{|s_1|} 4^{1-1/r+1/r_1+1/r_2} 2^{2|s_1+s_2|} \|f\|_{\dot{\mathcal{B}}^{s_1}_{\infty,r_1}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{p,r_2}} \\
=& \frac{C}{|s_1|} 4 2^{2|s_1+s_2|} \|f\|_{\dot{\mathcal{B}}^{s_1}_{\infty,r_1}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{p,r_2}}
\end{aligned}$$

$$\leqslant \frac{C}{|s_1|} 2^{2|s_1+s_2|} \|f\|_{\dot{\mathcal{B}}^{s_1}_{\infty,r_1}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{p,r_2}}.$$

For the case $\frac{1}{r_1} + \frac{1}{r_2} > 1$ and so $r = 1$, we have $r_2 < r_1'$ with $\frac{1}{r_1} + \frac{1}{r_1'} = 1$ and so $\ell^{r_2} \subset \ell^{r_1'}$. Thus, by replacing r_2 with r_1' in the first case, we can obtain the desired result. ❑

Similarly, the main continuity properties of nonhomogeneous paraproducts are described below.

Theorem 10.27. *Let $s, s_1, s_2 \in \mathbb{R}$ and $s_1 < 0$, and $1 \leqslant p, r_1, r_2 \leqslant \infty$. Then, there exists a constant $C > 0$ such that*

$$\|T_f g\|_{B^s_{p,r}} \leqslant C^{|s|+1} \|f\|_\infty \|g\|_{B^s_{p,r}},$$

$$\|T_f g\|_{B^{s_1+s_2}_{p,r}} \leqslant \frac{C^{|s_1+s_2|+1}}{|s_1|} \|f\|_{B^{s_1}_{\infty,r_1}} \|g\|_{B^{s_2}_{p,r_2}}, \text{ with } \frac{1}{r} := \min\left(1, \frac{1}{r_1} + \frac{1}{r_2}\right).$$

Proof. The proof is analogous to that of Theorem 10.26 and is thus omitted. ❑

We now examine the behavior of the remainder operator $\dot{R}$. Here, we have to consider terms of the type $\sum_{|k-j|\leqslant 1} \dot{\Delta}_k f \dot{\Delta}_j g$, the Fourier transforms of which are not supported in the annulus but rather in balls of the type $\{\xi \in \mathbb{R}^n : |\xi| \leqslant \frac{9}{8} \cdot 2^k\}$. Thus, to prove that the remainder terms belong to certain Besov spaces, we need the following lemma.

Lemma 10.28. *Let $s > 0$, $1 \leqslant p, r \leqslant \infty$ and B be a ball in $\mathbb{R}^n$. Assume that $\{f_k\}_{k\in\mathbb{Z}}$ is a sequence of smooth functions satisfying that the series $\sum_{k\in\mathbb{Z}} f_k$ converges to f in $\mathscr{S}_h'$ and*

$$\operatorname{supp} \widehat{f_k} \subset 2^k B \quad \text{and} \quad \left\| \{2^{ks} \|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{Z})} < \infty.$$

Then, we have

$$f \in \dot{\mathcal{B}}^s_{p,r} \quad \text{and} \quad \|f\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant \frac{C^s}{s \ln 2} \left\| \{2^{ks} \|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{Z})},$$

where C is a positive constant independent of s.

Proof. Denote $\mathbb{A}_j = \{\xi \in \mathbb{R}^n : \frac{8}{9} \cdot 2^j \leqslant |\xi| \leqslant \frac{9}{8} \cdot 2^{j+1}\}$. There exists an integer N_1 such that if $j \geqslant k + N_1$, then $\mathbb{A}_j \cap 2^k B = \varnothing$ and so $\widehat{\dot{\Delta}_j f_k} = 0$. Hence, we have, by Young's inequality for series and Proposition 9.8, that

$$\|f\|_{\dot{\mathcal{B}}^s_{p,r}} = \left\| \{2^{js} \|\dot{\Delta}_j f\|_p\}_j \right\|_{\ell^r(\mathbb{Z})} = \left\| \left\{ 2^{js} \left\| \sum_{k>j-N_1} \dot{\Delta}_j f_k \right\|_p \right\}_j \right\|_{\ell^r(\mathbb{Z})}$$

$$\leqslant \left\| \left\{ \sum_{k>j-N_1} 2^{(j-k)s} 2^{ks} \|\dot{\Delta}_j f_k\|_p \right\}_j \right\|_{\ell^r(\mathbb{Z})}$$

$$\leqslant \sum_{j<N_1} 2^{js} \left\| \{2^{ks}\|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{Z})} = \frac{2^{N_1 s}2^{-s}}{1-2^{-s}} \left\| \{2^{ks}\|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{Z})}$$

$$\leqslant \frac{2^{N_1 s}}{s\ln 2} \left\| \{2^{ks}\|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{Z})},$$

since $s > 0$. ❑

Now, we can state a result concerning the continuity of the remainder operator.

Theorem 10.29. *Let $s_1, s_2 \in \mathbb{R}$ and $1 \leqslant p_1, p_2, r_1, r_2 \leqslant \infty$. Assume that*

$$\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2} \leqslant 1, \text{ and } \frac{1}{r} := \frac{1}{r_1} + \frac{1}{r_2} \leqslant 1.$$

If $s_1 + s_2 > 0$, then we have, for any $(f,g) \in \dot{\mathcal{B}}^{s_1}_{p_1,r_1} \times \dot{\mathcal{B}}^{s_2}_{p_2,r_2}$,

$$\|\dot{R}(f,g)\|_{\dot{\mathcal{B}}^{s_1+s_2}_{p,r}} \leqslant \frac{C^{1+|s_1|+s_1+s_2}}{s_1+s_2} \|f\|_{\dot{\mathcal{B}}^{s_1}_{p_1,r_1}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{p_2,r_2}}.$$

If $r = 1$ and $s_1 + s_2 \geqslant 0$, then we have, for any $(f,g) \in \dot{\mathcal{B}}^{s_1}_{p_1,r_1} \times \dot{\mathcal{B}}^{s_2}_{p_2,r_2}$,

$$\|\dot{R}(f,g)\|_{\dot{\mathcal{B}}^{s_1+s_2}_{p,\infty}} \leqslant C^{1+|s_1|+s_1+s_2} \|f\|_{\dot{\mathcal{B}}^{s_1}_{p_1,r_1}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{p_2,r_2}}.$$

Proof. By the definition of the homogeneous remainder operator,

$$\dot{R}(f,g) = \sum_k R_k, \quad \text{with } R_k = \sum_{|\nu|\leqslant 1} \dot{\Delta}_{k-\nu} f \dot{\Delta}_k g.$$

Because $\operatorname{supp} \widehat{R_k} \subset \{\xi : |\xi| \leqslant 2^{k+1}\cdot \frac{27}{8}\}$ and $\operatorname{supp} \varphi_j \subset \{\xi : \frac{8}{9}\cdot 2^j \leqslant |\xi| \leqslant \frac{9}{8}\cdot 2^{j+1}\}$, we have

$$\dot{\Delta}_j R_k = 0, \quad \forall j > k+2.$$

Thus, we obtain

$$\dot{\Delta}_j \dot{R}(f,g) = \sum_{k\geqslant j-2} \dot{\Delta}_j R_k.$$

By Hölder's inequality, we infer that

$$\begin{aligned}
&2^{j(s_1+s_2)}\|\dot{\Delta}_j \dot{R}(f,g)\|_p \\
\lesssim & 2^{j(s_1+s_2)} \sum_{\substack{|\nu|\leqslant 1\\ k\geqslant j-2}} \|\dot{\Delta}_{k-\nu} f \dot{\Delta}_k g\|_p \\
\lesssim & 2^{j(s_1+s_2)} \sum_{\substack{|\nu|\leqslant 1\\ k\geqslant j-2}} \|\dot{\Delta}_{k-\nu} f\|_{p_1} \|\dot{\Delta}_k g\|_{p_2} \\
\lesssim & \sum_{\substack{|\nu|\leqslant 1\\ k\geqslant j-2}} 2^{\nu s_1} 2^{(j-k)(s_1+s_2)} 2^{(k-\nu)s_1} \|\dot{\Delta}_{k-\nu} f\|_{p_1} 2^{ks_2} \|\dot{\Delta}_k g\|_{p_2}.
\end{aligned}$$

Using Hölder's and Young's inequalities for series, we obtain the conclusion in the case where $s_1 + s_2$ is positive.

In the case where $r = 1$ and $s_1 + s_2 \geqslant 0$, it follows immediately from the fact that

$$2^{j(s_1+s_2)}\|\dot{\Delta}_j \dot{R}(f,g)\|_p \lesssim \sum_{|\nu|\leqslant 1} 2^{\nu s_1} 2^{2(s_1+s_2)} \sum_k 2^{(k-\nu)s_1}\|\dot{\Delta}_{k-\nu} f\|_{p_1} 2^{ks_2}\|\dot{\Delta}_k g\|_{p_2},$$

take the supremum over j, and use Hölder's inequality for series. ❑

By taking advantage of Bony's paraproduct decomposition (10.14), many results on continuities may be deduced from Theorems 10.26 and 10.29. As an initial example, we derive the following so-called ***tame estimates*** for the product of two functions in Besov spaces.

Corollary 10.30. *Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$. If (s, p, r) satisfies condition (10.1), then $L^\infty \cap \dot{\mathcal{B}}^s_{p,r}$ is an algebra under pointwise multiplication. Moreover, there exists a constant C, depending only on n, such that*

$$\|fg\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant \frac{C^{s+1}}{s}\left(\|f\|_\infty \|g\|_{\dot{\mathcal{B}}^s_{p,r}} + \|f\|_{\dot{\mathcal{B}}^s_{p,r}}\|g\|_\infty\right).$$

Proof. Using Bony's paraproduct decomposition, we have

$$fg = \dot{T}_f g + \dot{T}_g f + \dot{R}(f,g).$$

According to Theorem 10.26, we have

$$\|\dot{T}_f g\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C^{s+1}\|f\|_\infty \|g\|_{\dot{\mathcal{B}}^s_{p,r}}, \quad \text{and} \quad \|\dot{T}_g f\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C^{s+1}\|f\|_{\dot{\mathcal{B}}^s_{p,r}}\|g\|_\infty.$$

Now, using Theorem 10.29, we obtain

$$\|\dot{R}(f,g)\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant \frac{C^{s+1}}{s}\|f\|_{\dot{\mathcal{B}}^0_{\infty,\infty}}\|g\|_{\dot{\mathcal{B}}^s_{p,r}}.$$

Since $\|f\|_{\dot{\mathcal{B}}^0_{\infty,\infty}} \lesssim \|f\|_\infty$, we obtain the desired inequality. ❑

Our second example addresses the product of two functions in homogeneous Sobolev spaces.

Corollary 10.31. *For any $s_1, s_2 \in (-n/2, n/2)$, if $s_1 + s_2 > 0$, then we have*

$$\|fg\|_{\dot{\mathcal{B}}^{s_1+s_2-n/2}_{2,1}} \leqslant C\|f\|_{\dot{H}^{s_1}}\|g\|_{\dot{H}^{s_2}},$$

where the constant C is bounded by

$$C_n^{s_1+s_2}\max\left(\frac{1}{n-2s_1}, \frac{1}{n-2s_2}, \frac{1}{s_1+s_2}\right)$$

with C_n depending only on the dimension n.

Proof. We again use Bony's paraproduct decomposition. First, for $k = 1, 2$, as $\dot{H}^{s_k} = \dot{\mathcal{B}}^{s_k}_{2,2} \subset \dot{\mathcal{B}}^{s_k - n/2}_{\infty,2}$ with $s_k - n/2 = s_k - n/2 - n/\infty$, and $s_k - n/2 < 0$, Theorem 10.26 implies that

$$\begin{aligned}
&\|\dot{T}_f g + \dot{T}_g f\|_{\dot{\mathcal{B}}^{s_1+s_2-n/2}_{2,1}} \\
\leqslant & \frac{C^{|s_1+s_2-n/2|+1}}{|s_1 - n/2|} \|f\|_{\dot{\mathcal{B}}^{s_1-n/2}_{\infty,2}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{2,2}} + \frac{C^{|s_1+s_2-n/2|+1}}{|s_2 - n/2|} \|f\|_{\dot{\mathcal{B}}^{s_1}_{2,2}} \|g\|_{\dot{\mathcal{B}}^{s_2-n/2}_{\infty,2}} \\
\leqslant & C\|f\|_{\dot{H}^{s_1}} \|g\|_{\dot{H}^{s_2}}.
\end{aligned}$$

Second, as $s_1 + s_2 > 0$, Theorem 10.29 guarantees that

$$\|\dot{R}(f,g)\|_{\dot{\mathcal{B}}^{s_1+s_2}_{1,1}} \leqslant \frac{C^{s_1+s_2+1}}{s_1+s_2} \|f\|_{\dot{\mathcal{B}}^{s_1}_{2,2}} \|g\|_{\dot{\mathcal{B}}^{s_2}_{2,2}} \leqslant C\|f\|_{\dot{H}_{s_1}} \|g\|_{\dot{H}_{s_2}}.$$

As the embedding $\dot{\mathcal{B}}^{s_1+s_2}_{1,1} \subset \dot{\mathcal{B}}^{s_1+s_2-n/2}_{2,1}$, the corollary is proved. ❑

For the continuity properties of the remainder operator R in the nonhomogeneous case, we also need the following nonhomogeneous version of Lemma 10.28.

Lemma 10.32. *Let $s > 0$, $1 \leqslant p, r \leqslant \infty$ and B be a ball in $\mathbb{R}^n$. Assume that $\{f_k\}_{k\in\mathbb{N}_0}$ is a sequence of smooth functions satisfying*

$$\operatorname{supp} \widehat{f_k} \subset 2^k B \quad \textit{and} \quad \left\| \{2^{ks} \|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{N}_0)} < \infty.$$

Then, we have

$$f := \sum_{k\in\mathbb{N}_0} f_k \in B^s_{p,r}, \quad \textit{and} \quad \|f\|_{B^s_{p,r}} \leqslant C_s \left\| \{2^{ks} \|f_k\|_p\}_k \right\|_{\ell^r(\mathbb{N}_0)}.$$

Then, we have, via a similar proof to Theorem 10.29:

Theorem 10.33. *Let $s_1, s_2 \in \mathbb{R}$ and $1 \leqslant p_1, p_2, r_1, r_2 \leqslant \infty$. Assume that*

$$\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2} \leqslant 1, \textit{ and } \frac{1}{r} := \frac{1}{r_1} + \frac{1}{r_2} \leqslant 1.$$

If $s_1 + s_2 > 0$, then for any $(f,g) \in B^{s_1}_{p_1,r_1} \times B^{s_2}_{p_2,r_2}$,

$$\|R(f,g)\|_{B^{s_1+s_2}_{p,r}} \leqslant \frac{C^{1+|s_1|+s_1+s_2}}{s_1+s_2} \|f\|_{B^{s_1}_{p_1,r_1}} \|g\|_{B^{s_2}_{p_2,r_2}}.$$

If $r = 1$ and $s_1 + s_2 = 0$, then we have, for any $(f,g) \in B^{s_1}_{p_1,r_1} \times B^{s_2}_{p_2,r_2}$,

$$\|R(f,g)\|_{B^0_{p,\infty}} \leqslant C^{1+|s_1|} \|f\|_{B^{s_1}_{p_1,r_1}} \|g\|_{B^{s_2}_{p_2,r_2}}.$$

From this theorem and Theorem 10.27, we infer the following tame estimate.

Corollary 10.34. *If $s > 0$ and $1 \leqslant p, r \leqslant \infty$, then $L^\infty \cap B^s_{p,r}$ is an algebra. Moreover, there exists a constant C, depending only on n, such that*

$$\|fg\|_{B^s_{p,r}} \leqslant \frac{C^{s+1}}{s}\left(\|f\|_\infty\|g\|_{B^s_{p,r}} + \|f\|_{B^s_{p,r}}\|g\|_\infty\right).$$

§ 10.4 The paralinearization theorems

In this section, we first consider the action of smooth functions on the space $\dot{\mathcal{B}}^s_{p,r}$. More precisely, if f is a smooth function vanishing at 0, and u is a function of $\dot{\mathcal{B}}^s_{p,r}$, does $f(u)$ belong to $\dot{\mathcal{B}}^s_{p,r}$? The answer is given by the following theorem which is based on three lemmas.

Theorem 10.35. *Let $s_1, s_2 > 0$ and $1 \leqslant p_1, p_2, r_1, r_2 \leqslant \infty$. Let f be a smooth function on $\mathbb{R}$ satisfying $f(0) = 0$. Assume that (s_1, p_1, r_1) satisfies condition (10.1). Then, for any real-valued function $u \in \dot{\mathcal{B}}^{s_1}_{p_1,r_1} \cap \dot{\mathcal{B}}^{s_2}_{p_2,r_2} \cap L^\infty$, the function $f(u)$ belongs to the same space, and we have, for $k = 1, 2$, we have*

$$\|f(u)\|_{\dot{\mathcal{B}}^{s_k}_{p_k,r_k}} \leqslant C(f', \|u\|_\infty)\|u\|_{\dot{\mathcal{B}}^{s_k}_{p_k,r_k}}.$$

Proof. As u is bounded, we can assume without loss of generality that f is compactly supported. We introduce the telescopic series

$$\sum_{j\in\mathbb{Z}} f_j, \quad \text{with } f_j := f(\dot{S}_{j+1}u) - f(\dot{S}_j u).$$

The convergence of the series is ensured by the following lemma.

Lemma 10.36. *Under the hypotheses of Theorem 10.35, the series $\sum_{j\in\mathbb{Z}} f_j$ converges to $f(u) \in \mathscr{S}'_h$ in $\mathscr{S}'$, and we have*

$$f_j = m_j \dot{\Delta}_j u, \quad \text{with } m_j := \int_0^1 f'(\dot{S}_j u + t\dot{\Delta}_j u)dt. \qquad (10.17)$$

Proof. The identity (10.17) readily follows from the mean value theorem, so we will concentrate on the proof of the convergence of the series. We observe that

$$\sum_{j=-N}^{0} f_j = f(\dot{S}_1 u) - f(\dot{S}_{-N} u).$$

As $u \in \mathscr{S}'_h$ and $f(0) = 0$, we have $\|f(\dot{S}_{-N}u)\|_\infty \to 0$ as $N \to \infty$. Moreover, for all $2 \leqslant M \in \mathbb{N}$, we have

$$\sum_{j=1}^{M-1} f_j = f(\dot{S}_M u) - f(\dot{S}_1 u).$$

By virtue of the mean value theorem, we have

$$\|f(u) - f(\dot{S}_M u)\|_{p_k} \leqslant \|f'\|_\infty \|u - \dot{S}_M u\|_{p_k}.$$

Because $s_k > 0$, the function $\dot{S}_M u \to u$ in L^{p_k} as $M \to \infty$. Indeed, $u - \dot{S}_M u = \sum_{j \geqslant M} \dot{\Delta}_j u$, and then by Hölder's inequality for ℓ-spaces, we have, for $u \in \dot{\mathcal{B}}^{s_k}_{p_k,r_k}$ with $s_k > 0$,

$$\begin{aligned}
\left\| \sum_{j\geqslant M} \dot{\Delta}_j u \right\|_{p_k} &\leqslant \sum_{j\geqslant M} \|\dot{\Delta}_j u\|_{p_k} = \sum_{j\geqslant M} 2^{js_k} \|\dot{\Delta}_j u\|_{p_k} 2^{-js_k} \\
&\leqslant \begin{cases} \left(\sum_{j\geqslant M} (2^{js_k} \|\dot{\Delta}_j u\|_{p_k})^{r_k} \right)^{1/r_k} \left(\sum_{j\geqslant M} 2^{-js_k r_k'} \right)^{1/r_k'}, & \text{if } r_k \neq 1, \\ 2^{-Ms_k} \sum_{j\geqslant M} 2^{js_k} \|\dot{\Delta}_j u\|_{p_k}, & \text{if } r_k = 1 \end{cases} \\
&\leqslant \begin{cases} \dfrac{2^{-Ms_k}}{(1-2^{-s_k r_k'})^{1/r_k'}} \|u\|_{\dot{\mathcal{B}}^{s_k}_{p_k,r_k}}, & \text{if } r_k \neq 1, \\ 2^{-Ms_k} \|u\|_{\dot{\mathcal{B}}^{s_k}_{p_k,r_k}}, & \text{if } r_k = 1 \end{cases} \\
&\to 0, \quad \text{as } M \to +\infty.
\end{aligned}$$

Therefore, the series $\sum_{j\in\mathbb{Z}} f_j = f(u)$ in $L^\infty + L^{p_k}$.

Next, we prove that $f(u) \in \mathscr{S}_h'$. It suffices to show that $\|\dot{S}_j f(u)\|_\infty \to 0$ as $j \to -\infty$. For that, we use the decomposition

$$\dot{S}_j f(u) = \dot{S}_j \sum_{j' < -N} f_{j'} + \dot{S}_j \sum_{j' \geqslant -N} f_{j'}.$$

Let $\varepsilon > 0$. As the series $\sum_{j<0} f_j$ converges in L^∞, we can choose an integer N_ε such that

$$\left\| \dot{S}_j \sum_{j' < -N_\varepsilon} f_{j'} \right\|_\infty \leqslant \varepsilon/2.$$

As the f_j's are in L^{p_k} and $\sum_{j\in\mathbb{N}_0} f_j$ is convergent in L^{p_k}, we then have, using Bernstein's inequality,

$$\left\| \dot{S}_j \sum_{j' \geqslant -N_\varepsilon} f_{j'} \right\|_\infty \leqslant C 2^{jn/p_k} \left\| \dot{S}_j \sum_{j' \geqslant -N_\varepsilon} f_{j'} \right\|_{p_k} \leqslant C_\varepsilon 2^{jn/p_k}.$$

Thus, $\|\dot{S}_j f(u)\|_\infty \to 0$ as $j \to -\infty$. ❑

The terms m_j's will be handled according to the following lemma.

Lemma 10.37. *Let g be a smooth function from $\mathbb{R}^2$ to $\mathbb{R}$. For $j \in \mathbb{Z}$, we define*

$$m_j(g) := g(\dot{S}_j u, \dot{\Delta}_j u).$$

For any bounded function u, then we have

$$\|\partial^\alpha m_j(g)\|_\infty \leqslant C_\alpha(g, \|u\|_\infty) 2^{j|\alpha|}, \quad \forall \alpha \in \mathbb{N}_0^n, \ \forall j \in \mathbb{Z}.$$

Proof. For the case $\alpha = 0$, it is clear that $\|m_j(g)\|_\infty \leqslant C(g, \|u\|_\infty)$ since g is smooth and both $\dot{S}_j u$ and $\dot{\Delta}_j u$ are bounded by $\|u\|_\infty$. Thus, we need only to deal with the cases $|\alpha| \geqslant 1$. The proof relies on Faà di Bruno's formula,[a] which provides us with the formula

$$\partial^\alpha m_j(g) = \sum_{\ell_1,\ell_2,\nu} C^\nu_{\ell_1,\ell_2} \partial_1^{\ell_1} \partial_2^{\ell_2} g(\dot{S}_j u, \dot{\Delta}_j u) \left(\prod_{1 \leqslant |\beta| \leqslant |\alpha|} (\partial^\beta \dot{S}_j u)^{\nu_{\beta_1}} (\partial^\beta \dot{\Delta}_j u)^{\nu_{\beta_2}} \right),$$

where the coefficients $C^\nu_{\ell_1,\ell_2}$ are nonnegative integers, and the sum is taken over those ℓ_1, ℓ_2 and ν such that $1 \leqslant \ell_1 + \ell_2 \leqslant |\alpha|$,

$$\sum_{1 \leqslant |\beta| \leqslant |\alpha|} \nu_{\beta_k} = \ell_k \text{ for } k = 1, 2, \quad \text{and} \quad \sum_{1 \leqslant |\beta| \leqslant |\alpha|} (\nu_{\beta_1} + \nu_{\beta_2}) \beta = \alpha.$$

Note that, from Proposition 9.8, there exists a constant C such that

$$\max(\|\dot{\Delta}_j u\|_\infty, \|\dot{S}_j u\|_\infty) \leqslant C \|u\|_\infty, \quad \forall j \in \mathbb{Z}.$$

Since g and all its derivatives are bounded on $B(0, C\|u\|_\infty)$, Bernstein's inequality and the above formula thus ensure that

$$\|\partial^\alpha m_j(g)\|_\infty \leqslant C_\alpha(g, \|u\|_\infty) 2^{j|\alpha|},$$

due to $\sum_{1 \leqslant |\beta| \leqslant |\alpha|} (\nu_{\beta_1} + \nu_{\beta_2}) |\beta| = |\alpha|$. This completes the proof of the lemma. ❑

[a] Faà di Bruno's formula is an identity in mathematics generalizing the chain rule to higher derivatives, named after Francesco Faà di Bruno (1855, 1857), although he was not the first to state or prove the formula. The general case can be stated as follows.

Theorem ([BCD11, Lemma 2.3, p.54]). Let $u : \mathbb{R}^n \to \mathbb{R}^m$ and $F : \mathbb{R}^m \to \mathbb{R}$ be smooth functions. For each multi-index $\alpha \in \mathbb{N}_0^n$, we have

$$\partial^\alpha (F(u)) = \sum_{\mu,\nu} C_{\mu,\nu} \partial^\mu F \prod_{\substack{1 \leqslant |\beta| \leqslant |\alpha| \\ 1 \leqslant j \leqslant m}} (\partial^\beta u_j)^{\nu_{\beta_j}},$$

where the coefficients $C_{\mu,\nu}$ are nonnegative integers, and the sum is taken over those μ and ν such that $1 \leqslant |\mu| \leqslant |\alpha|$, $\nu_{\beta_j} \in \mathbb{N}$,

$$\sum_{1 \leqslant |\beta| \leqslant |\alpha|} \nu_{\beta_j} = \mu_j, \quad \text{for } 1 \leqslant j \leqslant m, \quad \text{and} \quad \sum_{\substack{1 \leqslant |\beta| \leqslant |\alpha| \\ 1 \leqslant j \leqslant m}} \nu_{\beta_j} \beta = \alpha.$$

In contrast with the situation that was encountered when proving Theorems 10.26 and 10.29, here, the elements f_j's of the approximating series $\sum f_j$ are not compactly supported in the frequency space. This difficulty is overcome by the following lemma.

Lemma 10.38. *Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$. Assume that $\{u_j\}_{j\in\mathbb{Z}}$ is a sequence of smooth functions where $\sum u_j$ converges to some $u \in \mathscr{S}'_h$ and*

$$N_{s,p,r}(\{u_j\}_{j\in\mathbb{Z}}) := \left\| \left\{ \sup_{|\alpha|\in\{0,[s]+1\}} 2^{j(s-|\alpha|)} \|\partial^\alpha u_j\|_p \right\}_j \right\|_{\ell^r(\mathbb{Z})} < \infty,$$

then there exists a constant C_s such that $u \in \dot{\mathcal{B}}^s_{p,r}$ and

$$\|u\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C_s N_{s,p,r}(\{u_j\}_{j\in\mathbb{Z}}).$$

Proof. As the series $\sum u_j$ converges to u in $\mathscr{S}'$, we have

$$\dot{\Delta}_k u = \sum_{j\leqslant k} \dot{\Delta}_k u_j + \sum_{j>k} \dot{\Delta}_k u_j.$$

By Proposition 9.8, we obtain

$$2^{ks} \left\| \sum_{j>k} \dot{\Delta}_k u_j \right\|_p \lesssim 2^{ks} \sum_{j>k} \|u_j\|_p \lesssim \sum_{j>k} 2^{(k-j)s} 2^{js} \|u_j\|_p. \tag{10.18}$$

By Bernstein's inequality, we may write that

$$\|\dot{\Delta}_k u_j\|_p \lesssim 2^{-k([s]+1)} \sup_{|\alpha|=[s]+1} \|\partial^\alpha u_j\|_p,$$

from which it follows that

$$2^{ks} \left\| \sum_{j\leqslant k} \dot{\Delta}_k u_j \right\|_p \lesssim \sum_{j\leqslant k} 2^{(j-k)([s]+1-s)} \sup_{|\alpha|=[s]+1} 2^{j(s-|\alpha|)} \|\partial^\alpha u_j\|_p.$$

This inequality, combined with (10.18), implies that

$$\begin{aligned}
&2^{ks}\|\dot{\Delta}_k u\|_p \\
&\lesssim \left(\sum_{j>k} 2^{(k-j)s} + \sum_{j\leqslant k} 2^{-(k-j)([s]+1-s)} \right) \sup_{|\alpha|\in\{0,[s]+1\}} 2^{j(s-|\alpha|)} \|\partial^\alpha u_j\|_p \\
&\lesssim \sum_j \left(2^{(k-j)s} \chi_{\{j-k>0\}} + 2^{-(k-j)([s]+1-s)} \chi_{\{j-k\leqslant 0\}} \right) \\
&\quad \cdot \sup_{|\alpha|\in\{0,[s]+1\}} 2^{j(s-|\alpha|)} \|\partial^\alpha u_j\|_p,
\end{aligned}$$

which proves the lemma by the Young inequality for series. ❑

Given the above three lemmas, it is now easy to prove the theorem. Note that, according to Lemma 10.38, it suffices to establish that

$$N_{s_k,p_k,r_k}(\{f_j\}_{j\in\mathbb{Z}}) < \infty. \tag{10.19}$$

Now, using Leibniz's formula, Bernstein's inequality and Lemma 10.37

with the function

$$g(x,y) = \int_0^1 f'(x+ty)dt,$$

then, from (10.17), we have

$$m_j = m_j(g) = g(\dot{S}_j u, \dot{\Delta}_j u) = \int_0^1 f'(\dot{S}_j u + t\dot{\Delta}_j u)dt.$$

Thus, we obtain, by Lemma 10.37,

$$\begin{aligned}
\|\partial^\alpha f_j\|_{p_k} =& \|\partial^\alpha (m_j \dot{\Delta}_j u)\|_{p_k} \\
\leqslant& \sum_{\beta \leqslant \alpha} C_\alpha^\beta \|\partial^\beta m_j\|_\infty \|\partial^{\alpha-\beta}\dot{\Delta}_j u\|_{p_k} \\
\lesssim& \sum_{|\beta| \leqslant |\alpha|} C_\alpha^\beta 2^{j|\beta|} C_\beta(f', \|u\|_\infty) 2^{j(|\alpha|-|\beta|)} \|\dot{\Delta}_j u\|_{p_k} \\
\leqslant& C_\alpha(f', \|u\|_\infty) 2^{j|\alpha|} \|\dot{\Delta}_j u\|_{p_k},
\end{aligned}$$

from which it follows that, for $k = 1, 2$,

$$\left\| \left\{ \sup_{|\alpha| \in \{0, [s_k]+1\}} 2^{j(s_k - |\alpha|)} \|\partial^\alpha f_j\|_{p_k} \right\}_j \right\|_{\ell^{r_k}} \tag{10.20}$$

$$\begin{aligned}
\leqslant& C_{s_k}(f', \|u\|_\infty) \left\| \left\{ 2^{js_k} \|\dot{\Delta}_j u\|_{p_k} \right\}_j \right\|_{\ell^{r_k}} \\
\leqslant& C_{s_k}(f', \|u\|_\infty) \|u\|_{\dot{\mathcal{B}}^{s_k}_{p_k, r_k}}.
\end{aligned} \tag{10.21}$$

This completes the proof of the theorem. ❑

In the case where f belongs to the space $\mathcal{C}_b^\infty(\mathbb{R})$ of smooth bounded functions with bounded derivatives of all orders and satisfies $f(0) = 0$, a slightly more accurate estimate may be obtained. Indeed, for $\alpha = 0$, the bound of m_j just follows from the boundedness of f'; in addition, for $|\alpha_k| \geqslant 1$, there exists a $\beta_k \in \mathbb{N}_0^n$ such that $\beta_k \leqslant \alpha_k$ and $|\beta_k| = 1$, then for any $j \in \mathbb{Z}$, we have, from Bernstein's inequality and Theorem 10.12,

$$\begin{aligned}
\|\partial^{\alpha_k} \dot{S}_j u\|_\infty =& \|\partial^{\alpha_k - \beta_k} \dot{S}_j \partial^{\beta_k} u\|_\infty \lesssim 2^{j(|\alpha_k|-1)} \|\dot{S}_j \partial^{\beta_k} u\|_\infty \\
\lesssim& 2^{j|\alpha_k|} \|\nabla u\|_{\dot{\mathcal{B}}^{-1}_{\infty,\infty}} \lesssim 2^{j|\alpha_k|} \|u\|_{\dot{\mathcal{B}}^0_{\infty,\infty}}.
\end{aligned}$$

Thus, for $|\alpha_k| \geqslant 1$ and any $j \in \mathbb{Z}$, we have

$$\max(\|\partial^{\alpha_k} \dot{S}_j u\|_\infty, \|\partial^{\alpha_k} \dot{\Delta}_j u\|_\infty) \lesssim 2^{j|\alpha_k|} \|u\|_{\dot{\mathcal{B}}^0_{\infty,\infty}}.$$

Arguing as in the proof of Lemma 10.37, we thus obtain

$$\|\partial^\alpha m_j\|_\infty \leqslant C_\alpha(f, \|u\|_{\dot{\mathcal{B}}^0_{\infty,\infty}}) 2^{j|\alpha|}, \quad \forall \alpha \in \mathbb{N}_0^n. \tag{10.22}$$

We now state the result we have just proven.

Corollary 10.39. *Let $f \in \mathcal{C}_b^\infty(\mathbb{R})$ satisfy $f(0) = 0$. Let $s_1, s_2 > 0$ and $1 \leqslant p_1, p_2, r_1, r_2 \leqslant \infty$. Assume that (s_1, p_1, r_1) satisfies condition (10.1). Then,* *for any real-valued function $u \in \dot{\mathcal{B}}^{s_1}_{p_1,r_1} \cap \dot{\mathcal{B}}^{s_2}_{p_2,r_2} \cap \dot{\mathcal{B}}^{0}_{\infty,\infty}$, we have $f(u) \in \dot{\mathcal{B}}^{s_1}_{p_1,r_1} \cap \dot{\mathcal{B}}^{s_2}_{p_2,r_2}$, and*

$$\|f(u)\|_{\dot{\mathcal{B}}^{s_k}_{p_k,r_k}} \leqslant C(f, \|u\|_{\dot{\mathcal{B}}^{0}_{\infty,\infty}})\|u\|_{\dot{\mathcal{B}}^{s_k}_{p_k,r_k}}, \quad k = 1, 2.$$

Finally, by combining Corollary 10.30 and Theorem 10.35 with the equality

$$f(u) - f(v) = (u - v) \int_0^1 f'(v + \tau(u - v))d\tau,$$

we readily obtain the following corollary.

Corollary 10.40. *Let f be a smooth function such that $f'(0) = 0$. Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$ and (s, p, r) satisfy the condition (10.1). For any pair (u, v) of functions in $\dot{\mathcal{B}}^s_{p,r} \cap L^\infty$, the function $f(u) - f(v)$ then belongs to $\dot{\mathcal{B}}^s_{p,r} \cap L^\infty$ and*

$$\begin{aligned}\|f(u) - f(v)\|_{\dot{\mathcal{B}}^s_{p,r}} \leqslant C\Big(&\|u - v\|_{\dot{\mathcal{B}}^s_{p,r}} \sup_{\tau\in[0,1]} \|\tau u + (1 - \tau)v\|_\infty \\ &+ \|u - v\|_\infty \sup_{\tau\in[0,1]} \|\tau u + (1 - \tau)v\|_{\dot{\mathcal{B}}^s_{p,r}}\Big),\end{aligned}$$

where C depends on f'', $\|u\|_\infty$ and $\|v\|_\infty$.

Next, we investigate the effect of left composition by smooth functions on Besov spaces $B^s_{p,r}$. We state an initial result.

Theorem 10.41. *Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$. Let f be a smooth function on $\mathbb{R}$ satisfying $f(0) = 0$. If $u \in B^s_{p,r} \cap L^\infty$, then so does $f(u)$, and we have*

$$\|f(u)\|_{B^s_{p,r}} \leqslant C(s, f', \|u\|_\infty)\|u\|_{B^s_{p,r}}.$$

This theorem can be proved along the same lines as that of Theorem 10.35. We note that it is based on the following lemma, the proof of which is left to the interested reader.

Lemma 10.42. *Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$. Assume that $(u_j)_{j\in\mathbb{N}_0}$ is a sequence of smooth functions satisfying*

$$N_{s,p,r}(\{u_j\}_{j\in\mathbb{N}_0}) := \left\| \left\{ \sup_{|\alpha|\leqslant[s]+1} 2^{j(s-|\alpha|)} \|\partial^\alpha u_j\|_p \right\}_j \right\|_{\ell^r(\mathbb{N}_0)} < \infty,$$

then there exists a constant C_s such that $u := \sum_{j\in\mathbb{N}_0} u_j \in B^s_{p,r}$ and $\|u\|_{B^s_{p,r}} \leqslant C_s N_{s,p,r}(\{u_j\}_{j\in\mathbb{N}_0})$.

In the case where the function $f \in \mathcal{C}_b^\infty(\mathbb{R})$, Theorem 10.41 may be slightly improved.

Theorem 10.43. *Let $f \in \mathcal{C}_b^\infty(\mathbb{R})$ satisfy $f(0) = 0$. Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$. If $u \in B^s_{p,r}$ and the first derivatives of u belongs to $B^{-1}_{\infty,\infty}$, then $f(u) \in B^s_{p,r}$ and we have*

$$\|f(u)\|_{B^s_{p,r}} \leqslant C(s, f, \|\nabla u\|_{B^{-1}_{\infty,\infty}})\|u\|_{B^s_{p,r}}.$$

Remark 10.44. If $u \in B^{n/p}_{p,r}$, then $\nabla u \in B^{-1}_{\infty,\infty}$. Thus, the space $B^{n/p}_{p,r}$ is stable under left composition by functions of C_b^∞ vanishing at 0. This result applies in particular to the Sobolev space $H^{n/2} = B^{n/2}_{2,2}$.

Finally, we state the nonhomogeneous counterpart of Corollary 10.40.

Corollary 10.45. *Let f be a smooth function such that $f'(0) = 0$. Let $s > 0$ and $1 \leqslant p, r \leqslant \infty$. For any couple (u, v) of functions in $B^s_{p,r} \cap L^\infty$, the function $f(u) - f(v)$ then belongs to $B^s_{p,r} \cap L^\infty$ and*

$$\begin{aligned} \|f(u) - f(v)\|_{B^s_{p,r}} \leqslant C\Big(&\|u - v\|_{B^s_{p,r}} \sup_{\tau\in[0,1]} \|\tau u + (1-\tau)v\|_\infty \\ &+ \|u - v\|_\infty \sup_{\tau\in[0,1]} \|\tau u + (1-\tau)v\|_{B^s_{p,r}}\Big), \end{aligned}$$

where C depends on f'', $\|u\|_\infty$ and $\|v\|_\infty$.

When the function u has enough regularity, we can obtain more information on $f(u)$. In the following theorem, we state that, up to an error term that proves to be more regular than u, $f(u)$ may be written as a paraproduct involving u and $f'(u)$.

Theorem 10.46. *Let $s_1, s_2 > 0$ and $s_2 \notin \mathbb{N}$, $1 \leqslant p, r_1, r_2 \leqslant \infty$ with $r_1 \leqslant r_2$, and f be a smooth function satisfying $f'(0) = 0$. Let $1 \leqslant r \leqslant \infty$ be defined by $1/r = \min(1, 1/r_1 + 1/r_2)$. For any $u \in B^{s_1}_{p,r_1} \cap B^{s_2}_{\infty,r_2}$, we then have*

$$\|f(u) - T_{f'(u)}u\|_{B^{s_1+s_2}_{p,r}} \leqslant C(f'', \|u\|_\infty)\|u\|_{B^{s_1}_{p,r_1}} \|u\|_{B^{s_2}_{\infty,r_2}}.$$

Proof. To prove this theorem, we again write that

$$f(u) = \sum_j f_j, \quad \text{with } f_j := f(S_{j+1}u) - f(S_j u).$$

According to the second order Taylor formula,[a] we have

$$f_j = f'(S_j u)\Delta_j u + M_j(\Delta_j u)^2, \quad \text{with } M_j := \int_0^1 (1-t) f''(S_j u + t\Delta_j u)dt.$$

Since $T_{f'(u)}u = \sum_j S_{j-1}f'(u)\Delta_j u$, we have

$$f(u) - T_{f'(u)}u = \sum_j [(f'(S_j u) - S_{j-1}f'(u))\Delta_j u + M_j(\Delta_j u)^2].$$

Let $\mu_j := f'(S_j u) - S_{j-1}(f'(u))$. Obviously, we have

$$f_j - S_{j-1}(f'(u))\Delta_j u = \mu_j \Delta_j u + M_j(\Delta_j u)^2.$$

Applying Lemma 10.37 with $g(x,y) = \int_0^1 (1-t)f''(x+ty)dt$ gives

$$\|\partial^\alpha M_j\|_\infty \leqslant C_\alpha(f'', \|u\|_\infty)2^{j|\alpha|}, \quad \forall \alpha \in \mathbb{N}_0^n. \tag{10.23}$$

Using Leibniz's formula, we can write

$$\partial^\alpha(M_j(\Delta_j u)^2) = \sum_{\gamma \leqslant \beta \leqslant \alpha} C_{\alpha,\beta}C_{\beta,\gamma}\partial^{\alpha-\beta}M_j \partial^{\beta-\gamma}\Delta_j u \partial^\gamma \Delta_j u.$$

Using Bernstein's inequality and (10.23), we obtain

$$\|\partial^{\alpha-\beta}M_j\partial^{\beta-\gamma}\Delta_j u \partial^\gamma \Delta_j u\|_p \leqslant C_\alpha(f'', \|u\|_\infty)2^{j|\alpha|}\|\Delta_j u\|_\infty \|\Delta_j u\|_p.$$

Thus, according to the definition of Besov spaces, we have

$$\left\| \left\{ 2^{j(s_1+s_2-|\alpha|)}\|\partial^\alpha(M_j(\Delta_j u)^2)\|_p \right\}_{j \geqslant -1} \right\|_{\ell^r} \leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_1}_{p,r_1}}\|u\|_{B^{s_2}_{\infty,r_2}}. \tag{10.24}$$

We temporarily assume that

$$\left\| \left\{ 2^{j(s_2-|\alpha|)}\|\partial^\alpha \mu_j\|_\infty \right\}_{j \geqslant -1} \right\|_{\ell^{r_2}} \leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}. \tag{10.25}$$

Using (10.24), we have

$$\begin{aligned}
&\left\| \left\{ 2^{j(s_1+s_2-|\alpha|)}\|\partial^\alpha(f_j - S_{j-1}(f'(u))\Delta_j u)\|_p \right\}_{j \geqslant -1} \right\|_{\ell^r} \\
=&\left\| \left\{ 2^{j(s_1+s_2-|\alpha|)}\|\partial^\alpha(\mu_j \Delta_j u + M_j(\Delta_j u)^2)\|_p \right\}_{j \geqslant -1} \right\|_{\ell^r} \\
\leqslant& C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_1}_{p,r_1}}\|u\|_{B^{s_2}_{\infty,r_2}}.
\end{aligned}$$

Applying Lemma 10.42 then yields the desired result.

To complete the proof of the theorem, we must justify the inequality (10.25). First, we investigate the case where $|\alpha| < s_2$. We have

$$\mu_j = \mu_j^{(1)} + \mu_j^{(2)},$$

where $\mu_j^{(1)} := f'(S_j u) - f'(u)$ and $\mu_j^{(2)} := f'(u) - S_{j-1}(f'(u)) = \sum_{k \geqslant j-1} \dot{\Delta}_k f'(u)$. Using the fact that $S_j u$ converges to $u \in L^\infty$ in $\mathscr{S}'$ as $j \to \infty$, we obtain

$$f'(u) - f'(S_j u) = \sum_{k \geqslant j} \tilde{f}_k, \quad \text{with } \tilde{f}_k := f'(S_{k+1}u) - f'(S_k u). \tag{10.26}$$

Applying (10.20) yields

$$\left\| \left\{ 2^{k(s_2-|\alpha|)} \|\partial^\alpha \tilde{f}_k\|_\infty \right\}_{k\geqslant j} \right\|_{\ell^{r_2}} \leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}. \tag{10.27}$$

Then, by Young's inequality, we have, for $|\alpha| < s_2$,

$$\begin{aligned} &\left\| \left\{ 2^{j(s_2-|\alpha|)} \|\partial^\alpha (\mu_j^{(1)})\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \\ \leqslant &\left\| \left\{ \sum_{k\geqslant j} 2^{(j-k)(s_2-|\alpha|)} 2^{k(s_2-|\alpha|)} \|\partial^\alpha \tilde{f}_k\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \\ \leqslant &C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}. \end{aligned}$$

By Bernstein's inequality and Theorem 10.41 (need $f'(0) = 0$), we have

$$\partial^\alpha f'(u) \in B^{s_2-|\alpha|}_{\infty,r_2} \quad \text{and} \quad \|\partial^\alpha f'(u)\|_{B^{s_2-|\alpha|}_{\infty,r_2}} \leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}.$$

Thus, in view of Young's inequality, we can write that

$$\begin{aligned} &\left\| \left\{ 2^{j(s_2-|\alpha|)} \|\partial^\alpha \mu_j^{(2)}\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \\ \leqslant &\left\| \left\{ 2^{j(s_2-|\alpha|)} \sum_{k\geqslant j-1} \|\Delta_k \partial^\alpha f'(u)\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \\ \leqslant &C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}} \sum_{k\leqslant 1} 2^{k(s_2-|\alpha|)} \\ \leqslant &C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}. \end{aligned}$$

This completes the proof of (10.25) when $|\alpha| < s_2$.

Since $s_2 \notin \mathbb{N}$, we only need to consider the remainder case when $|\alpha| > s_2$ which is treated differently. As $\partial^\alpha f'(u) \in B^{s_2-|\alpha|}_{\infty,r_2}$, we have, using Theorems 10.19 and 10.41,

$$\left\| \left\{ 2^{j(s_2-|\alpha|)} \|\partial^\alpha S_{j-1} f'(u)\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}.$$

We now estimate $\partial^\alpha f'(S_j u)$. Because $S_j u$ converges to 0 in L^∞ as $j \to -\infty$, we can write that

$$f'(S_j u) = \sum_{k\leqslant j-1} \tilde{f}_k, \quad \text{with} \quad \tilde{f}_k := f'(S_{k+1}u) - f'(S_k u).$$

Using (10.27) and Young's inequality, we obtain

$$\begin{aligned} &\left\| \left\{ 2^{j(s_2-|\alpha|)} \|\partial^\alpha f'(S_j u)\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \\ \leqslant &\left\| \left\{ 2^{j(s_2-|\alpha|)} \sum_{k\leqslant j-1} \|\partial^\alpha \tilde{f}_k\|_\infty \right\}_{j\geqslant -1} \right\|_{\ell^{r_2}} \end{aligned}$$

$$\leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}} \sum_{k\geqslant 1} 2^{k(s_2-|\alpha|)}$$
$$\leqslant C_\alpha(f'', \|u\|_\infty)\|u\|_{B^{s_2}_{\infty,r_2}}.$$

The inequality (10.25) is proved, as is the theorem. ❑

[a] If $f : \mathbb{R}^n \to \mathbb{R}$ is $(k+1)$-times continuously differentiable in the closed ball B, then

$$f(x) = \sum_{|\alpha|=0}^{k} \frac{\partial^\alpha f(a)}{\alpha!}(x-a)^\alpha + \sum_{|\beta|=k+1} R_\beta(x)(x-a)^\beta,$$
$$R_\beta(x) = \frac{|\beta|}{\beta!}\int_0^1 (1-t)^{|\beta|-1}\partial^\beta f(a+t(x-a))dt.$$

§ 10.5 Commutator estimates

This section is devoted to various commutator estimates. The following basic lemma will be frequently used in this section.

Lemma 10.47. *Let $\theta \in C^1(\mathbb{R}^n)$ satisfy $(1+|\cdot|)\hat{\theta} \in L^1$, and $p, q, r \in [1,\infty]$. There exists a constant C such that for any $a \in Lip$ with $\nabla a \in L^p$ and any $b \in L^q$, we have, for any $\lambda > 0$,*

$$\|[\theta(\lambda^{-1}D), a]b\|_r \leqslant C\lambda^{-1}\|\nabla a\|_p\|b\|_q, \quad \text{with } \frac{1}{r} = \frac{1}{p} + \frac{1}{q}.$$

Proof. To prove this lemma, it suffices to rewrite $\theta(\lambda^{-1}D)$ as a convolution operator. Indeed,

$$\begin{aligned}
&([\theta(\lambda^{-1}D), a]b)(x) = \theta(\lambda^{-1}D)(ab)(x) - a(x)\theta(\lambda^{-1}D)b(x)\\
=&(2\pi)^{-n/2}\left[\left(\left(\theta(\lambda^{-1}\xi)\right)^\vee * (ab)\right)(x) - a(x)\left(\left(\theta(\lambda^{-1}\xi)\right)^\vee * b\right)(x)\right]\\
=&(2\pi)^{-n/2}\left[\lambda^n(\check{\theta}(\lambda\cdot) * (ab))(x) - a(x)\lambda^n(\check{\theta}(\lambda\cdot) * b)(x)\right]\\
=&(2\pi)^{-n/2}\lambda^n\left[\int_{\mathbb{R}^n}\check{\theta}(\lambda z)a(x-z)b(x-z)dz - a(x)\int_{\mathbb{R}^n}\check{\theta}(\lambda z)b(x-z)dz\right]\\
=&(2\pi)^{-n/2}\lambda^n\int_{\mathbb{R}^n}\check{\theta}(\lambda z)[a(x-z)-a(x)]b(x-z)dz.
\end{aligned}$$

Let $k_1(z) := (2\pi)^{-n/2}|z||\check{\theta}(z)|$. From the first-order Taylor formula, we deduce that

$$\left|([\theta(\lambda^{-1}D), a]b)(x)\right| \leqslant \lambda^{-1}\int_0^1\int_{\mathbb{R}^n}\lambda^n k_1(\lambda z)|\nabla a(x-tz)||b(x-z)|dzdt.$$

Now, taking the L^r norm of the above inequality, using Minkowski's inequality and Hölder's inequality, we obtain

$$\left\|[\theta(\lambda^{-1}D), a]b\right\|_r \leqslant \lambda^{-1}\int_0^1\int_{\mathbb{R}^n} \lambda^n k_1(\lambda z)\|\nabla a(\cdot - tz)\|_p\|b(\cdot - z)\|_q dzdt.$$

The translation invariance of the Lebesgue measure then ensures that

$$\left\|[\theta(\lambda^{-1}D), a]b\right\|_r \leqslant \lambda^{-1}\|k_1\|_1\|\nabla a\|_p\|b\|_q,$$

which yields the desired result since $\|k_1\|_1 = \int_{\mathbb{R}^n}|z||\check{\theta}(z)|đz = \int_{\mathbb{R}^n}|z||\widehat{\theta}(z)|đz < \infty$. ❑

Remark 10.48. If we take $\theta = \varphi$ and $\lambda = 2^j$, then this lemma can be interpreted as a gain of one derivative by commutation between the operator Δ_j and the multiplication by a function with gradient in L^p, i.e.,

$$\|[\Delta_j, a]b\|_r \lesssim 2^{-j}\|\nabla a\|_p\|b\|_q.$$

Theorem 10.49. *Let f be a smooth function on $\mathbb{R}^n$. Assume that f is homogeneous of degree m away from a neighborhood of 0. Let $0 < \rho < 1$, $s \in \mathbb{R}$ and $1 \leqslant p, p_1, p_2, r \leqslant \infty$ satisfy $1/p = 1/p_1 + 1/p_2$. Then, we have*

$$\|[T_a, f(D)]u\|_{B^{s-m+\rho}_{p,r}} \leqslant C\|\nabla a\|_{B^{\rho-1}_{p_1,\infty}}\|u\|_{B^s_{p_2,r}}. \tag{10.28}$$

In the limit case $\rho = 1$, we have

$$\|[T_a, f(D)]u\|_{B^{s-m+1}_{p,r}} \leqslant C\|\nabla a\|_{p_1}\|u\|_{B^s_{p_2,r}}. \tag{10.29}$$

The above constants $C > 0$ depend only on s, ρ and n.

Proof. We first treat the case $0 < \rho < 1$. For convenience, we redefine

$$\tilde{\Delta}_j = \sum_{-2\leqslant \ell - j\leqslant 1}\Delta_\ell, \quad \widetilde{\varphi}_j = \sum_{-2\leqslant \ell - j\leqslant 1}\varphi_\ell.$$

We have from Proposition 9.6

$$\begin{aligned}[T_a, f(D)]u &= \sum_{j\geqslant 1}\left[S_{j-1}af(D)\Delta_j u - f(D)(S_{j-1}a\Delta_j u)\right]\\ &= \sum_{j\geqslant 1}\left[S_{j-1}af(D)\tilde{\Delta}_j\Delta_j u - f(D)\tilde{\Delta}_j(S_{j-1}a\Delta_j u)\right]\\ &= \sum_{j\geqslant 1}[S_{j-1}a, f(D)\tilde{\Delta}_j]\Delta_j u.\end{aligned}$$

Note that the general term of the above series is spectrally supported in dyadic annuli. Hence, according to Lemma 10.17, it suffices to prove that

$$\left\|2^{j(s-m+\rho)}\|[S_{j-1}a, f(D)\tilde{\Delta}_j]\Delta_j u\|_p\right\|_{\ell^r} \leqslant C\|\nabla a\|_{B^{\rho-1}_{p_1,\infty}}\|u\|_{B^s_{p_2,r}}. \tag{10.30}$$

Owing to the homogeneity of f away from 0, there exists an $N_0 \in \mathbb{N}$ such that

$$f(D)\tilde{\Delta}_j = \mathscr{F}^{-1}f(\xi)\tilde{\varphi}_j(\xi)\mathscr{F} = 2^{jm}\mathscr{F}^{-1}f(2^{-j}\xi)\tilde{\varphi}(2^{-j}\xi)\mathscr{F}$$

$$=2^{jm}(f\tilde{\varphi})(2^{-j}D), \quad \forall j \geqslant N_0.$$

Taking advantage of Lemma 10.47 with $\theta = f\tilde{\varphi}$, we thus infer that for any $j \geqslant N_0$,

$$\|[S_{j-1}a, f(D)\tilde{\Delta}_j]\Delta_j u\|_p \leqslant C2^{j(m-1)}\|\nabla S_{j-1}a\|_{p_1}\|\Delta_j u\|_{p_2}.$$

Of course, if $1 \leqslant j < N_0$, we can still write, according to Lemma 10.47 with $\theta = f(2^j\cdot)\tilde{\varphi}$,

$$\begin{aligned}\|[S_{j-1}a, f(D)\tilde{\Delta}_j]\Delta_j u\|_p \leqslant& C2^{-j}\|\nabla S_{j-1}a\|_{p_1}\|\Delta_j u\|_{p_2}\\ \leqslant& C2^{N_0|m|}2^{j(m-1)}\|\nabla S_{j-1}a\|_{p_1}\|\Delta_j u\|_{p_2}.\end{aligned}$$

Because $\|\nabla S_{j-1}a\|_{p_1} \leqslant C2^{j(1-\rho)}\|\nabla a\|_{B^{\rho-1}_{p_1,\infty}}$ if $\rho < 1$ in view of Theorem 10.19, we can now conclude that (10.30) is satisfied, and complete the proof.

For the case $\rho = 1$, we only need to modify (10.30), where we replace the term $\|\nabla a\|_{B^{\rho-1}_{p_1,\infty}}$ by $\|\nabla a\|_{p_1}$. Then, by the same lines after (10.30), we can obtain the desired results. ❑

Finally, we give an important estimate for the commutators.

Theorem 10.50. *Let $\sigma \in \mathbb{R}$, $1 \leqslant r \leqslant \infty$ and $1 \leqslant p \leqslant p_1 \leqslant \infty$. Let v be a vector field over $\mathbb{R}^n$. Assume that*

$$\sigma > -n\min(1/p_1, 1/p') \quad (\text{or } \sigma > -1 - n\min(1/p_1, 1/p') \text{ if } \operatorname{div} v = 0). \tag{10.31}$$

Define $R_j := [v\cdot\nabla, \Delta_j]f$ (or $R_j := \operatorname{div}([v, \Delta_j]f)$ if $\operatorname{div} v = 0$). Then, there exists a constant $C > 0$, depending continuously on p, p_1, σ and n, such that

$$\left\|\left\{2^{j\sigma}\|R_j\|_p\right\}_j\right\|_{\ell^r} \leqslant C\|\nabla v\|_{B^{n/p_1}_{p_1,\infty}\cap L^\infty}\|f\|_{B^\sigma_{p,r}}, \quad \text{if } \sigma < 1 + n/p_1. \tag{10.32}$$

Furthermore, if $\sigma > 0$ (or $\sigma > -1$ if $\operatorname{div} v = 0$) and $1/p_2 = 1/p - 1/p_1$, then

$$\left\|\left\{2^{j\sigma}\|R_j\|_p\right\}_j\right\|_{\ell^r} \leqslant C\left(\|\nabla v\|_\infty\|f\|_{B^\sigma_{p,r}} + \|\nabla f\|_{p_2}\|\nabla v\|_{B^{\sigma-1}_{p_1,r}}\right). \tag{10.33}$$

In the limit case $\sigma = -n\min(1/p_1, 1/p')$ (or $\sigma = -1 - n\min(1/p_1, 1/p')$ if $\operatorname{div} v = 0$), we have

$$\sup_{j\geqslant -1} 2^{j\sigma}\|R_j\|_p \leqslant C\|\nabla v\|_{B^{n/p_1}_{p_1,1}}\|f\|_{B^\sigma_{p,\infty}}. \tag{10.34}$$

Proof. To show that only the gradient part of v is involved in the estimates, we split v into low and high frequencies: $v = S_0 v + \tilde{v}$. Obviously,

$$\|S_0\nabla v\|_q \lesssim \|\nabla v\|_q, \quad \|\nabla\tilde{v}\|_q \lesssim \|\nabla v\|_q, \quad \forall q \in [1,\infty]. \tag{10.35}$$

Furthermore, as $\tilde{v}$ is spectrally supported away from the origin, Bernstein's inequality ensures that

$$\|\Delta_j \nabla \tilde{v}\|_q \sim 2^j \|\Delta_j \tilde{v}\|_q, \quad \forall q \in [1, \infty], \quad \forall j \geqslant -1. \tag{10.36}$$

We now have (with the summation convention over repeated indices):

$$\begin{aligned} R_j =& v \cdot \nabla \Delta_j f - \Delta_j (v \cdot \nabla f) = v^k \Delta_j \partial_k f - \Delta_j (v^k \partial_k f) = [v^k, \Delta_j] \partial_k f \\ =& [(S_0 v^k + \tilde{v}^k), \Delta_j] \partial_k f = [S_0 v^k, \Delta_j] \partial_k f + [\tilde{v}^k, \Delta_j] \partial_k f. \end{aligned}$$

Writing Bony's paraproduct decomposition for $[\tilde{v}^k, \Delta_j]\partial_k f$, we end up with $R_j = \sum_{i=1}^{8} R_j^i$, where

$$\begin{aligned} &R_j^1 = [T_{\tilde{v}^k}, \Delta_j] \partial_k f, && R_j^2 = T_{\partial_k \Delta_j f} \tilde{v}^k, \\ &R_j^3 = -\Delta_j (T_{\partial_k f} \tilde{v}^k), && R_j^4 = \partial_k R(\tilde{v}^k, \Delta_j f), \\ &R_j^5 = -R(\operatorname{div} \tilde{v}, \Delta_j f), && R_j^6 = -\partial_k \Delta_j R(\tilde{v}^k, f), \\ &R_j^7 = \Delta_j R(\operatorname{div} \tilde{v}, f), && R_j^8 = [S_0 v^k, \Delta_j] \partial_k f. \end{aligned}$$

Bounds for R_j^1. By Proposition 9.6, we have

$$R_j^1 = \sum_{-2 \leqslant j - j' \leqslant 1} [S_{j'-1} \tilde{v}^k, \Delta_j] \partial_k \Delta_{j'} f.$$

Hence, according to Lemma 10.47 and (10.35), we have

$$\begin{aligned} &\left\| \left\{ 2^{j\sigma} \|R_j^1\|_p \right\}_j \right\|_{\ell^r} \\ \lesssim & \left\| \left\{ 2^{j\sigma} \sum_{-2 \leqslant j - j' \leqslant 1} 2^{-j} \|\nabla S_{j'-1} \tilde{v}^k\|_\infty \|\partial_k \Delta_{j'} f\|_p \right\}_j \right\|_{\ell^r} \\ \lesssim & \|\nabla v\|_\infty \|f\|_{B^\sigma_{p,r}}. \end{aligned} \tag{10.37}$$

Bounds for R_j^2. By Proposition 9.6, we have

$$R_j^2 = \sum_{j' \geqslant j+1} S_{j'-1} \partial_k \Delta_j f \Delta_{j'} \tilde{v}^k.$$

Hence, by Bernstein's inequality, Young's inequality and using (10.35) and (10.36), we have

$$\begin{aligned} \left\| \left\{ 2^{j\sigma} \|R_j^2\|_p \right\}_j \right\|_{\ell^r} \lesssim & \left\| \left\{ 2^{j\sigma} \sum_{j' \geqslant j+1} \|S_{j'-1} \partial_k \Delta_j f\|_p \|\Delta_{j'} \tilde{v}^k\|_\infty \right\}_j \right\|_{\ell^r} \\ \lesssim & \left\| \left\{ 2^{j\sigma} \sum_{j' \geqslant j+1} 2^j \|S_{j'-1} \Delta_j f\|_p \|\Delta_{j'} \tilde{v}\|_\infty \right\}_j \right\|_{\ell^r} \\ \lesssim & \left\| \left\{ \sum_{j' \geqslant j+1} 2^{j-j'} 2^{j\sigma} \|\Delta_j f\|_p 2^{j'} \|\Delta_{j'} \tilde{v}\|_\infty \right\}_j \right\|_{\ell^r} \\ \lesssim & \|\nabla v\|_\infty \|f\|_{B^\sigma_{p,r}}. \end{aligned} \tag{10.38}$$

Bounds for R_j^3. We proceed as follows:

$$R_j^3 = -\sum_{-2\leqslant j-j'\leqslant 1} \Delta_j \left(S_{j'-1}\partial_k f \Delta_{j'}\tilde{v}^k\right) \tag{10.39}$$

$$= -\sum_{\substack{-2\leqslant j-j'\leqslant 1\\ j''\leqslant j'-2}} \Delta_j \left(\Delta_{j''}\partial_k f \Delta_{j'}\tilde{v}^k\right). \tag{10.40}$$

Therefore, writing $1/p_2 = 1/p - 1/p_1$ (then $p \leqslant p_2$) and using Bernstein's inequality, Young's inequality, (10.35) and (10.36), we have, for $\sigma < 1 + n/p_1$,

$$\begin{aligned}
&\left\|\left\{2^{j\sigma}\|R_j^3\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim&\left\|\left\{\sum_{\substack{-2\leqslant j-j'\leqslant 1\\ j''\leqslant j'-2}} 2^{j\sigma}\|\Delta_{j''}\partial_k f\|_{p_2}\|\Delta_{j'}\tilde{v}^k\|_{p_1}\right\}_j\right\|_{\ell^r}\\
\lesssim&\left\|\left\{\sum_{\substack{-2\leqslant j-j'\leqslant 1\\ j''\leqslant j'-2}} 2^{j\sigma}2^{j''(1+n(1/p-1/p_2))}\|\Delta_{j''}f\|_p 2^{-j'}\|\Delta_{j'}\nabla\tilde{v}\|_{p_1}\right\}_j\right\|_{\ell^r}\\
\lesssim&\left\|\left\{\sum_{\substack{-2\leqslant j-j'\leqslant 1\\ j''\leqslant j'-2}} 2^{(j-j')\sigma}2^{(j'-j'')(\sigma-1-n/p_1)}2^{j''\sigma}\|\Delta_{j''}f\|_p 2^{j'n/p_1}\|\Delta_{j'}\nabla\tilde{v}\|_{p_1}\right\}_j\right\|_{\ell^r}\\
\lesssim&\|\nabla v\|_{B^{n/p_1}_{p_1,\infty}}\|f\|_{B^\sigma_{p,r}}.
\end{aligned} \tag{10.41}$$

Note that, starting from (10.39), we can alternatively obtain

$$\begin{aligned}
&\left\|\left\{2^{j\sigma}\|R_j^3\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim&\left\|\left\{\sum_{-2\leqslant j-j'\leqslant 1} 2^{j\sigma}\|\nabla S_{j'-1}f\|_{p_2}2^{-j'}\|\Delta_{j'}\nabla\tilde{v}\|_{p_1}\right\}_j\right\|_{\ell^r}\\
\lesssim&\|\nabla f\|_{p_2}\|\nabla v\|_{B^{\sigma-1}_{p_1,r}}.
\end{aligned} \tag{10.42}$$

Bounds for R_j^4 and R_j^5. We have

$$R_j^4 = \sum_{|j-j'|\leqslant 2} \partial_k(\Delta_{j'}\tilde{v}^k\tilde{\Delta}_{j'}\Delta_j f).$$

Hence, by (10.36) and (10.35), we obtain

$$\begin{aligned}
\left\|\left\{2^{j\sigma}\|R_j^4\|_p\right\}_j\right\|_{\ell^r} \lesssim&\left\|\left\{\sum_{|j-j'|\leqslant 2} 2^{j\sigma}2^{j'}\|\Delta_{j'}\tilde{v}^k\|_\infty\|\Delta_j f\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim&\|\nabla v\|_\infty\|f\|_{B^s_{p,r}}.
\end{aligned} \tag{10.43}$$

A similar bound holds for R_j^5.

Bounds for R_j^6 and R_j^7. We have

$$R_j^6 = -\partial_k \Delta_j \sum_{\substack{|j'-j''|\leqslant 1\\ j'>j-4}} \Delta_{j'}\tilde{v}^k \Delta_{j''} f.$$

We first consider the case where $1/p + 1/p_1 \leqslant 1$. Let $1/p_3 = 1/p + 1/p_1$. Then, under the condition $\sigma > -1 - n/p_1$, and by embedding $B_{p_3,r}^{\sigma+n/p_1} \subset B_{p,r}^{\sigma}$ (since R_j^6 is of the form $\Delta_j g$) and Young's inequality, we have

$$\begin{aligned}
&\left\|\left\{2^{j\sigma}\|R_j^6\|_p\right\}_j\right\|_{\ell^r} \lesssim \left\|\left\{2^{j(\sigma+n/p_1)}\|R_j^6\|_{p_3}\right\}_j\right\|_{\ell^r}\\
\lesssim &\left\|\left\{\sum_{\substack{|j'-j''|\leqslant 1\\ j'>j-4}} 2^{j(1+\sigma+n/p_1)}\|\Delta_{j'}\tilde{v}\|_{p_1}\|\Delta_{j''}f\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim &\left\|\left\{\sum_{\substack{|j'-j''|\leqslant 1\\ j'>j-4}} 2^{(j'-j'')\sigma}2^{(j-j')(\sigma+1+n/p_1)}2^{j'n/p_1}\|\Delta_{j'}\nabla\tilde{v}\|_{p_1}2^{j''\sigma}\|\Delta_{j''}f\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim &\|\nabla v\|_{B^{n/p_1}_{p_1,\infty}}\|f\|_{B^{\sigma}_{p,r}}.
\end{aligned} \tag{10.44}$$

Now, if $1/p + 1/p_1 > 1$, then the above argument has to be applied with p' instead of p_3, and by the embedding $B_{1,r}^{\sigma+n/p'} \subset B_{p,r}^{\sigma}$ and Bernstein's inequality (due to $p_1 < p'$), we still obtain, for $\sigma > -1 - n/p'$,

$$\begin{aligned}
&\left\|\left\{2^{j\sigma}\|R_j^6\|_p\right\}_j\right\|_{\ell^r} \lesssim \left\|\left\{2^{j(\sigma+n/p')}\|R_j^6\|_{1}\right\}_j\right\|_{\ell^r}\\
\lesssim &\left\|\left\{\sum_{\substack{|j'-j''|\leqslant 1\\ j'>j-4}} 2^{j(1+\sigma+n/p')}\|\Delta_{j'}\tilde{v}\|_{p'}\|\Delta_{j''}f\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim &\left\|\left\{\sum_{\substack{|j'-j''|\leqslant 1\\ j'>j-4}} 2^{(j'-j'')\sigma}2^{(j-j')(\sigma+1+n/p')}2^{j'n/p_1}\|\Delta_{j'}\nabla\tilde{v}\|_{p_1}2^{j''\sigma}\|\Delta_{j''}f\|_p\right\}_j\right\|_{\ell^r}\\
\lesssim &\|\nabla v\|_{B^{n/p_1}_{p_1,\infty}}\|f\|_{B^{\sigma}_{p,r}}.
\end{aligned} \tag{10.45}$$

Note that in the limit case $\sigma = -1 - n\min(1/p_1, 1/p')$, a similar argument yields

$$\sup_j 2^{j\sigma}\|R_j^6\|_p \lesssim \|\nabla v\|_{B^{n/p_1}_{p_1,1}}\|f\|_{B^{\sigma}_{p,\infty}}. \tag{10.46}$$

Similar arguments lead to

$$\left\|\left\{2^{j\sigma}\|R_j^7\|_p\right\}_j\right\|_{\ell^r} \lesssim \|\nabla v\|_{B^{n/p_1}_{p_1,\infty}}\|f\|_{B^{\sigma}_{p,r}}, \text{ if } \sigma > -n\min(1/p_1, 1/p'), \tag{10.47}$$

$$\left\|\left\{2^{j\sigma}\|R_j^7\|_p\right\}_j\right\|_{\ell^\infty}\lesssim\|\nabla v\|_{B^{n/p_1}_{p_1,1}}\|f\|_{B^\sigma_{p,\infty}}, \text{ if } \sigma=-n\min(1/p_1,1/p'), r=\infty. \tag{10.48}$$

Finally, we stress that if $\sigma>-1$, then taking $p_1=\infty$ in (10.44), combined with the embedding $L^\infty\subset B^0_{\infty,\infty}$, yields

$$\left\|\left\{2^{j\sigma}\|R_j^6\|_p\right\}_j\right\|_{\ell^r}\lesssim\|\nabla v\|_\infty\|f\|_{B^\sigma_{p,r}}. \tag{10.49}$$

Of course, the same inequality holds true for R_j^7 if $\sigma>0$.

Bounds for R_j^8. As $R_j^8=-\sum_{|j-j'|\leqslant 1}[\Delta_j,\Delta_{-1}v]\cdot\nabla\Delta_{j'}f$, Lemma 10.47 yields

$$\begin{aligned}\left\|\left\{2^{j\sigma}\|R_j^8\|_p\right\}_j\right\|_{\ell^r}\lesssim&\left\|\left\{\sum_{|j-j'|\leqslant 1}2^{-j}\|\nabla\Delta_{-1}v\|_\infty 2^{j'\sigma}2^{j'}\|\Delta_{j'}f\|_p\right\}_j\right\|_{\ell^r}\\ \lesssim&\|\nabla v\|_\infty\|f\|_{B^\sigma_{p,r}}.\end{aligned} \tag{10.50}$$

Combining the above inequalities yields the desired results. ❑

Remark 10.51. There are a number of variations on the statement of Theorem 10.50. For instance, inequalities (10.32), (10.33) and (10.34) are also valid in the homogeneous framework (i.e., with $\dot\Delta_j$ instead of Δ_j and with homogeneous Besov norms instead of nonhomogeneous ones), provided that (σ,p,r) satisfies condition (10.1). The proof follows along the lines of the proof of Theorem 10.50. It is simply a matter of replacing the nonhomogeneous blocks by homogeneous ones.

Remark 10.52. The inequalities (10.32), (10.33) and (10.34) are still true for the commutator

$$\dot S_{j+N_0}v\cdot\nabla\dot\Delta_j f-\dot\Delta_j(v\cdot\nabla f),$$

where N_0 is any fixed integer. Indeed, for all $j\geqslant -1$, we have

$$\begin{aligned}\left\|(\dot S_{j+N_0}v-v)\cdot\nabla\dot\Delta_j f\right\|_p\lesssim&2^j\|\dot S_{j+N_0}v-v\|_\infty\|\dot\Delta_j f\|_p\\ \lesssim&\sum_{j'\geqslant j+N_0}2^{j-j'}\|\nabla\dot\Delta_{j'}v\|_\infty\|\dot\Delta_j f\|_p\\ \lesssim&\|\nabla v\|_{\dot B^0_{\infty,\infty}}\|\dot\Delta_j f\|_p.\end{aligned}$$

§ 10.6 Time-space Besov spaces

One of the fundamental ideas, in view of Littlewood-Paley theory, is that nonlinear evolution PDEs may be treated very efficiently after localization by means of Littlewood-Paley decomposition. Indeed, it is often easier to bound each dyadic block in $L^\rho([0,T];L^p)$ than to directly estimate the solution of the whole PDE in $L^\rho([0,T];\dot{\mathcal{B}}^s_{p,r})$.

As a final step, we must combine the estimates for each block and then perform a (weighted) ℓ^r summation. In doing so, however, we do not obtain an estimate in a space of type $L^\rho([0,T];\dot{\mathcal{B}}^s_{p,r})$ since the time integration has been performed before the summation.

This naturally leads to the following definition.

Definition 10.53. For $T>0$, $s\in\mathbb{R}$, and $1\leqslant \rho,p,r\leqslant\infty$, we set

$$\|u\|_{\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r})} := \left\| \left(2^{ks} \|\dot{\Delta}_k u\|_{L^\rho_T(L^p)} \right)_k \right\|_{\ell^r(\mathbb{Z})}.$$

We can then define the space

$$\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r}) := \Big\{ u\in \mathscr{S}'((0,T)\times\mathbb{R}^n) : \lim_{k\to-\infty} \dot{S}_k u = 0 \text{ in } L^\rho([0,T];L^\infty(\mathbb{R}^n)) \text{ and } \|u\|_{\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r})} < \infty \Big\}.$$

The space $\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r})$ may be linked with the more classical spaces

$$L^\rho_T(\dot{\mathcal{B}}^s_{p,r}) := L^\rho([0,T];\dot{\mathcal{B}}^s_{p,r})$$

via the Minkowski inequality:

$$\|u\|_{\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r})} \leqslant \|u\|_{L^\rho_T(\dot{\mathcal{B}}^s_{p,r})} \quad \text{if } \rho\leqslant r, \qquad \|u\|_{\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r})} \geqslant \|u\|_{L^\rho_T(\dot{\mathcal{B}}^s_{p,r})} \quad \text{if } \rho\geqslant r.$$

The general principle is that all the properties of continuity for the product, composition, remainder, and paraproduct remain true in these spaces. The exponent ρ must behave according to Hölder's inequality for the time variable. For instance, we have the time estimate

$$\|uv\|_{\tilde{L}^\rho_T(\dot{\mathcal{B}}^s_{p,r})} \lesssim \|u\|_{L^{\rho_1}_T(L^\infty)} \|v\|_{\tilde{L}^{\rho_2}_T(\dot{\mathcal{B}}^s_{p,r})} + \|v\|_{L^{\rho_3}_T(L^\infty)} \|u\|_{\tilde{L}^{\rho_4}_T(\dot{\mathcal{B}}^s_{p,r})},$$

whenever $s>0$, $1\leqslant p,r\leqslant\infty$, $1\leqslant\rho,\rho_1,\rho_2,\rho_3,\rho_4\leqslant\infty$, and

$$\frac{1}{\rho} = \frac{1}{\rho_1} + \frac{1}{\rho_2} = \frac{1}{\rho_3} + \frac{1}{\rho_4}.$$

This approach also works in the nonhomogeneous Besov spaces $B^s_{p,r}$ which leads to function spaces denoted by $\tilde{L}^\rho_T(B^s_{p,r})$.

Exercises

Exercise 10.1. Let $q,p,r\in[1,\infty]$ and $s\in\mathbb{R}$. Prove that $L^q\cap\dot{\mathcal{B}}^s_{p,r}$ is a Banach space.

Exercise 10.2. Prove Theorem 10.20.

Bibliography

[AB12] P. Auscher and L. Bandara. *Real Harmonic Analysis*. Australia: ANU eView, 2012. DOI: 10.22459/RHA.2012 (see p. i).

[Bab61] K. I. Babenko. "An inequality in the theory of Fourier integrals". In: *Izv. Akad. Nauk SSSR Ser. Mat.* 25 (1961), pp. 531–542 (see p. 72).

[BCD11] H. Bahouri, J. Chemin, and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*. Vol. 343. GMW. Springer-Verlag, Berlin Heidelberg, 2011 (see pp. i, 73, 290).

[Bec75a] W. Beckner. "Inequalities in Fourier analysis". In: *Ann. of Math. (2)* 102.1 (1975), pp. 159–182. DOI: 10.2307/1970980 (see pp. 22, 72).

[Bec75b] W. Beckner. "Inequalities in Fourier analysis on R^n". In: *Proc. Nat. Acad. Sci. U.S.A.* 72 (1975), pp. 638–641. DOI: 10.1073/pnas.72.2.638 (see pp. 22, 72).

[BL76a] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Vol. 223. GMW. Berlin: Springer-Verlag, 1976, pp. x+207 (see pp. i, 42, 252, 257).

[BL76b] H. J. Brascamp and E. H. Lieb. "Best constants in Young's inequality, its converse, and its generalization to more than three functions". In: *Advances in Math.* 20.2 (1976), pp. 151–173. DOI: 10.1016/0001-8708(76)90184-5 (see p. 22).

[Bro15] R. Brown. *Lecture notes: harmonic analysis*. 2015 (see p. i).

[CM85] M. Cowling and G. Mauceri. "Inequalities for some maximal functions. I". In: *Trans. Amer. Math. Soc.* 287.2 (1985), pp. 431–455. DOI: 10.2307/1999654 (see p. 45).

[CM86] M. Cowling and G. Mauceri. "Inequalities for some maximal functions. II". In: *Trans. Amer. Math. Soc.* 296.1 (1986), pp. 341–365. DOI: 10.2307/2000578 (see p. 45).

[CZ56] A. P. Calderon and A. Zygmund. "A Note on the Interpolation of Sublinear Operations". In: *American Journal of Mathematics* 78.2 (1956), pp. 282–288. DOI: 10.2307/2372516 (see p. 22).

[Dan10] R. Danchin. "On the solvability of the compressible Navier-Stokes system in bounded domains". In: *Nonlinearity* 23.2 (2010), pp. 383–407. DOI: 10.1088/0951-7715/23/2/009 (see p. 278).

[DiB16] E. DiBenedetto. *Real analysis*. Second. Birkhäuser Advanced Texts: Basel Textbooks. Birkhäuser/Springer, New York, 2016, pp. xxxii+596. DOI: 10.1007/978-1-4939-4005-9 (see p. 8).

[Din07] G. Ding. *New talk of functional analysis*. Beijing: Science press, 2007 (see p. 136).

[DS88] N. Dunford and J. Schwartz. *Linear operators, part 1: general theory*. Vol. 10. John Wiley & Sons, 1988 (see p. 13).

[Fol09] G. B. Folland. *A guide to advanced real analysis*. Vol. 37. The Dolciani Mathematical Expositions. Mathematical Association of America, Washington, DC, 2009, pp. x+107 (see p. 1).

[Fol99] G. B. Folland. *Real analysis. Modern techniques and their applications*. Second. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1999, pp. xvi+386 (see pp. i, 1, 3, 4, 6–12, 14, 34, 44, 51, 73).

[FS71] C. Fefferman and E. M. Stein. "Some maximal inequalities". In: *Amer. J. Math.* 93 (1971), pp. 107–115. DOI: 10.2307/2373450 (see p. 59).

[FS72] C. Fefferman and E. M. Stein. "H^p spaces of several variables". In: *Acta Math.* 129.3-4 (1972), pp. 137–193. DOI: 10.1007/BF02392215 (see p. 180).

[Gra14a] L. Grafakos. *Classical Fourier analysis*. Third. Vol. 249. Graduate Texts in Mathematics. Springer, New York, 2014, pp. xviii+638. DOI: 10.1007/978-1-4939-1194-3 (see pp. i, 42–44, 62, 73, 95, 105, 109, 135, 141, 142, 248).

[Gra14b] L. Grafakos. *Modern Fourier analysis*. Third. Vol. 250. Graduate Texts in Mathematics. Springer, New York, 2014, pp. xvi+624. DOI: 10.1007/978-1-4939-1230-8 (see pp. i, 62, 142, 143, 166, 191, 238, 239, 262).

[H03] L. Hörmander. *The analysis of linear partial differential operators. I*. Classics in Mathematics. Distribution theory and Fourier analysis, Reprint of the second (1990) edition. Springer-Verlag, Berlin, 2003, pp. x+440. DOI: 10.1007/978-3-642-61497-2 (see p. 197).

[Hir53] I. I. Hirschman Jr. "A convexity theorem for certain groups of transformations". In: *J. Analyse Math.* 2 (1953), pp. 209–218. DOI: 10.1007/BF02825637 (see p. 24).

[Jaw77] B. Jawerth. "Some observations on Besov and Lizorkin-Triebel spaces". In: *Math. Scand.* 40.1 (1977), pp. 94–104. DOI: 10.7146/math.scand.a-11678 (see pp. 154, 252).

[JN61] F. John and L. Nirenberg. “On functions of bounded mean oscillation”. In: *Comm. Pure Appl. Math.* 14 (1961), pp. 415–426 (see p. 166).

[Mur19] J. Murphy. *A Course in Harmonic Analysis*. Vol. 2. Course Materials. 2019. 274 pp. (see p. i).

[MWZ12] C. Miao, J. Wu, and Z. Zhang. *Littlewood-Paley theory and its application in hydrodynamic equations*. Vol. 126. Fundamental Series in Mordon Mathematics. Beijing: Science Press, 2012 (see p. i).

[Pey18] J. Peyrière. *An introduction to singular integrals*. Higher Education Press, Beijing; Society for Industrial and Applied Mathematics, Philadelphia, PA, 2018, pp. viii+115 (see pp. 14, 62, 63, 191, 192).

[Pic72] S. K. Pichorides. “On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov”. In: *Studia Math.* 44 (1972), 165–179. (errata insert). DOI: 10.4064/sm-44-2-165-179 (see p. 105).

[Rub96] L. A. Rubel. *Entire and meromorphic functions*. Universitext. With the assistance of James E. Colliander. Springer-Verlag, New York, 1996, pp. viii+187. DOI: 10.1007/978-1-4612-0735-1 (see p. 24).

[Saw18] Y. Sawano. *Theory of Besov spaces*. Vol. 56. Developments in Mathematics. Springer, Singapore, 2018, pp. xxiii+945. DOI: 10.1007/978-981-13-0836-9 (see pp. 154, 252).

[Spi74] M. R. Spiegel. *Fourier analysis with applications to boundary value problems*. McGraw-Hill Comp. Inc., 1974 (see p. 94).

[SS03] E. M. Stein and R. Shakarchi. *Complex analysis*. Princeton Lectures in Analysis, II. Princeton, NJ: Princeton University Press, 2003, pp. xviii+379 (see pp. 19, 24).

[SS05] E. M. Stein and R. Shakarchi. *Real analysis. Measure theory, integration, and Hilbert spaces*. Princeton Lectures in Analysis, III. Princeton, NJ: Princeton University Press, 2005, pp. xx+402 (see pp. 1, 5, 73).

[Ste70] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton, N.J.: Princeton University Press, 1970, pp. xiv+290 (see pp. i, 58).

[Ste93] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Vol. 43. Princeton Mathematical Series. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. Princeton, NJ: Princeton University Press, 1993, pp. xiv+695 (see p. 248).

[SW71] E. M. Stein and G. Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton Mathematical Series, No. 32. Princeton, N.J.: Princeton University Press, 1971, pp. x+297 (see pp. i, 42).

[Tao06] T. Tao. *Lecture notes for MATH 247A: Fourier analysis*. 2006 (see pp. i, 100).

[Tri83] H. Triebel. *Theory of function spaces*. Vol. 78. Monographs in Mathematics. Basel: Birkhäuser Verlag, 1983, p. 284. DOI: `10.1007/978-3-0346-0416-1` (see p. 257).

[Wan+11] B. Wang et al. *Harmonic analysis method for nonlinear evolution equations*. Vol. I. World Scientific Publishing Co. Pte. Ltd., 2011 (see p. i).

[Zho99] M. Q. Zhou. *Lectures on harmonic analysis (real variable methods)*. Peking University Press, 1999 (see pp. 44, 62, 109).

Index