Graduate Lectures in Analysis

Lecture Notes on

Harmonic Analysis

Chengchun Hao

Institute of Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences

LECTURE NOTES ON HARMONIC ANALYSIS

Chengchun Hao
Institute of Mathematics, AMSS
Chinese Academy of Sciences

Email: hcc@amss.ac.cn

Updated: April 28, 2020

Academy of Mathematics and Systems Science
Chinese Academy of Sciences

Preface

Harmonic analysis, as a subfield of analysis, is particularly interested in the study of quantitative properties on functions, and how these quantitative properties change when apply various operators. In the past two centuries, it has become a vast subject with applications in areas as diverse as signal processing, quantum mechanics, and neuroscience.

Most of the material in these notes are excerpted from the book of Stein [Ste70], the book of Stein and Weiss [SW71], the books of Grafakos [Gra14a, Gra14b] and the book of Wang-Huo-Hao-Guo [WHHG11], etc. with some necessary modification.

Please email me (hcc@amss.ac.cn) with corrections or suggested improvements of any kinds.

Chengchun Hao
Beijing
April 28, 2020

References

[Gra14a] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[Gra14b] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[SW71] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

[WHHG11] Baoxiang Wang, Zhaohui Huo, Chengchun Hao, and Zihua Guo. *Harmonic analysis method for nonlinear evolution equations*, volume I. World Scientific Publishing Co. Pte. Ltd., 2011.

Contents

I

The Weak L^p Spaces and Interpolation

§1.1 The distribution function and weak L^p

In this chapter, we will consider general measurable functions in measurable spaces instead of $\mathbb{R}^n$ only.

A ***σ-algebra*** on a set X is a collection of subsets of X that includes the empty subset, is closed under complement, countable unions and countable intersections. A ***measure space*** is a set X equipped with a σ-algebra of subsets of it and a function μ from the σ-algebra to $[0,\infty]$ that satisfies $\mu(\varnothing)=0$ and

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

for any sequence $\{B_j\}$ of pairwise disjoint elements of the σ-algebra. The function μ is called a ***positive measure*** on X and elements of the σ-algebra of X are called ***measurable sets***. Measure spaces will be assumed to be complete, i.e., subsets of the σ-algebra of measure zero also belong to the σ-algebra.

A measure space X is called ***σ-finite*** if there is a sequence of measurable subsets X_n of it such that $X=\bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n)<\infty$. A real-valued function f on a measure space is called measurable if the set $\{x\in X: f(x)>\lambda\}$ is measurable for all real numbers λ. A complex-valued function is measurable if and only if its real and imaginary parts are measurable.

We adopt the usual convention that two functions are considered equal if they agree except on a set of μ-measure zero. For $p\in[1,\infty)$, we denote by $L^p(X,d\mu)$ (or simply $L^p(d\mu)$, $L^p(X)$ or even L^p) the Lebesgue-space of (all equivalence classes of) scalar-valued μ-measurable functions f on X, such that

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

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

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

It is well-known that $L^p(X,\mu)$ is a Banach space for any $p\in[1,\infty]$ (i.e., the Riesz-Fisher theorem). For any $p\in(1,\infty)$, we define the Hölder conjugate number $p'=\frac{p}{p-1}$. Moreover, we set $1'=\infty$ and $\infty'=1$, so that $(p')'=p$ for all $p\in[1,\infty]$. Hölder's inequality says that for all $p\in[1,\infty]$ and all measurable function f, g on (X,μ), we have

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

It is also a well-known fact that the dual $(L^p)'$ of L^p is isometric to $L^{p'}$ for all $p\in(1,\infty)$ and also when $p=1$ if X is σ-finite. Furthermore, the L^p norm of a function can be obtained via duality when $p\in(1,\infty)$ as follows:

$$\|f\|_p = \sup_{\|g\|_{p'}=1}\left|\int_X fg\,d\mu\right|.$$

The endpoint cases $p=1,\infty$ also work if X is σ-finite.

It is often convenient to work with functions that are only locally in some L^p space. We give the definition in the following.

Definition 1.1.1.

For $p\in[1,\infty)$, the space $L^p_{\text{loc}}(X,\mu)$ or simply $L^p_{\text{loc}}(X)$ is the set of all Lebesgue measurable functions f on $\mathbb{R}^n$ that satisfy

$$\int_K |f|^p d\mu < \infty \tag{1.1.1}$$

for any compact subset $K\subset X$. Functions that satisfy (1.1.1) with $p=1$ are called ***locally integrable function*** on X.

The union of all $L^p(X)$ spaces for $p\in[1,\infty]$ is contained in $L^1_{\text{loc}}(X)$. More generally, for $1\leqslant p<q<\infty$, we have

$$L^q \hookrightarrow L^q_{\text{loc}} \hookrightarrow L^p_{\text{loc}}.$$

We recall that a ***simple function*** is a finite linear combination of characteristic functions of measurable subsets of X, these subsets may have infinite measure. A ***finitely simple function*** has the canonical form $\sum_{j=1}^N a_j\chi_{B_j}$ where $N<\infty$, $a_j\in\mathbb{C}$, and B_j are pairwise disjoint measurable sets with $\mu(B_j)<\infty$. If $N=\infty$, this function will be called ***countably simple***. Finitely simple functions are exactly the integrable simple functions. Every non-negative measurable function is the pointwise limit of an increasing sequence of simple functions; if the space is σ-finite, these simple functions can be chosen to be finitely simple. In particular, for $p\in[1,\infty)$, the (finitely) simple functions are dense in $L^p(X,\mu)$. In addition, the space of simple functions (not necessarily with finite measure support) is dense in $L^\infty(X,\mu)$.

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

Definition 1.1.2.

Let $f(x)$ be a measurable function on (X,μ), then the function $f_*:[0,\infty)\mapsto[0,\infty]$ defined by

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

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

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

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

Now, we give some properties of distribution functions.

Proposition 1.1.3.

Let f and g be measurable functions on (X, μ). Then for all $\alpha, \beta > 0$, we have

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

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

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

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

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

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

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

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

(vii) From (vi), it follows

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

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, by noticing that $1 \leqslant p < \infty$, that

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

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

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

For other ones, they are easy to verify. ■

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

Theorem 1.1.4: The equivalent norm of L^p

Let (X, μ) be a σ-finite measure space. Then for $f \in L^p(X, \mu)$, $p \in [1, \infty]$, we have

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

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

Moreover, for any increasing continuously differentiable function φ on $[0, \infty)$ with $\varphi(0) = 0$ and every measurable function f on X with $\varphi(|f|)$ integrable on X, we have

$$\int_X \varphi(|f|) d\mu(x) = \int_0^{\infty} \varphi'(\alpha) f_*(\alpha) d\alpha. \tag{1.1.2}$$

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

$$\int_X |f(x)|^p d\mu(x) = -\int_0^\infty \alpha^p df_*(\alpha). \tag{1.1.3}$$

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

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

then, $\mu(E_j) = f_*((j-1)\varepsilon) - f_*(j\varepsilon)$, and

$$\begin{aligned}\int_X |f(x)|^p d\mu(x) =& \lim_{\varepsilon\to 0}\sum_{j=1}^\infty (j\varepsilon)^p \mu(E_j)\\ =& -\lim_{\varepsilon\to 0}\sum_{j=1}^\infty (j\varepsilon)^p [f_*(j\varepsilon) - f_*((j-1)\varepsilon)]\\ =& -\int_0^\infty \alpha^p df_*(\alpha),\end{aligned}$$

since the measure space is σ-finite.

Now we return to prove (i). If the values of both sides are infinite, then it is clearly true. If one of the integral is finite, then it is clear that $f_*(\alpha) < +\infty$ and $f(x)$ is finite almost everywhere. Thus (1.1.3) is valid.

If either $f \in L^p(X)$ or $\int_0^\infty \alpha^{p-1} f_*(\alpha) d\alpha < \infty$ for $1 \leqslant p < \infty$, then we always have $\alpha^p f_*(\alpha) \to 0$ as $\alpha \to +\infty$ and $\alpha \to 0$ from the property (vii) and (viii) in Proposition 1.1.3.

Therefore, integrating by part, we have

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

Thus, i) is true. Identity (1.1.2) follows similarly, replacing the function α^p by the more general function $\varphi(\alpha)$ which has similar properties.

For (ii), we have

$$\begin{aligned}\inf\{\alpha : f_*(\alpha) = 0\} =& \inf\{\alpha : \mu(\{x \in X : |f(x)| > \alpha\}) = 0\}\\ =& \inf\{\alpha : |f(x)| \leqslant \alpha, \ a.e.\}\\ =& \operatorname*{ess\,sup}_{x\in X} |f(x)| = \|f\|_\infty.\end{aligned}$$

We complete the proofs. ■

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

Definition 1.1.5.

For $1 \leqslant p < \infty$, the space $L^p_*(X, \mu)$, consists of all μ-measurable functions f such that

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

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

Two functions in $L^p_*(X, \mu)$ are considered equal if they are equal μ-a.e. Now, we will show that L^p_* is a quasi-normed linear space.

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

2° From (iii) in Proposition 1.1.3, we can show that for any $k \in \mathbb{C} \setminus \{0\}$

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

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

3° By the part (iv) in Proposition 1.1.3 and the triangle inequality of L^p norms, we have

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

Thus, L^p_* is a quasi-normed linear space for $1 \leqslant p < \infty$.

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

Theorem 1.1.6.

For any $1 \leqslant p < \infty$, and any $f \in L^p(X, \mu)$, we have

$$\|f\|_{L^p_*} \leqslant \|f\|_p.$$

Hence, $L^p(X, \mu) \hookrightarrow L^p_*(X, \mu)$.

Proof. From the part (vi) in Proposition 1.1.3, 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 $L^p \hookrightarrow L^p_*$ is strict for $1 \leqslant p < \infty$. 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 S^{n-1} in $\mathbb{R}^n$, but h is in $L^p_*(\mathbb{R}^n)$ and we may check easily that

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

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 $\Re z > 0$.

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

Next, we recall the notion of convergence in measure and give the relations of some convergence notions.

Definition 1.1.7.

Let f, f_n, $n = 1, 2, \cdots$, be measurable functions on the measure space (X, μ). The sequence $\{f_n\}$ is said to convergent 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

$$n > n_0 \Longrightarrow \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) < \varepsilon. \tag{1.1.4}$$

Remark 1.1.8. The above definition is equivalent to the following statement:

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0, \quad \forall \varepsilon > 0. \tag{1.1.5}$$

Clearly, (1.1.5) implies (1.1.4). To see the converse, given $\varepsilon > 0$, pick $0 < \delta < \varepsilon$ and apply (1.1.4) 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 $n \to \infty$, we deduce that

$$\limsup_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \leqslant \delta. \tag{1.1.6}$$

Since (1.1.6) holds for all $\delta \in (0, \varepsilon)$, (1.1.5) follows by letting $\delta \to 0$.

Convergence in measure is a weaker notion than convergence in either L^p or L^p_*, $1 \leqslant p \leqslant \infty$, as the following proposition indicates:

Proposition 1.1.9.

Let $p \in [1, \infty]$ and $f_n, f \in L^p_*(X, \mu)$.

(i) If $f_n, f \in L^p$ and $f_n \to f$ in L^p, then $f_n \to f$ in L^p_*.
(ii) If $f_n \to f$ in L^p_*, then $f_n \xrightarrow{\mu} f$.

Proof. For $p \in [1, \infty)$, Proposition 1.1.6 gives that

$$\|f_n - f\|_{L^p_*} \leqslant \|f_n - f\|_p,$$

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

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

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

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

Example 1.1.10. 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

$$\begin{aligned}
\|f_{k,j}\|_{L^p_*} =& \sup_{\alpha>0} \alpha|\{x : f_{k,j}(x) > \alpha\}|^{1/p} \\
=& \sup_{\alpha>0} \alpha|\{x : k^{1/p}\chi_{(\frac{j-1}{k},\frac{j}{k})}(x) > \alpha\}|^{1/p} \\
=& \sup_{\alpha>0} \alpha\left|\left\{x \in \left(\frac{j-1}{k},\frac{j}{k}\right) : k^{1/p} > \alpha\right\}\right|^{1/p} \\
=& \sup_{0<\alpha<k^{1/p}} \alpha(1/k)^{1/p} \\
\geqslant& \sup_{k\geqslant 1} \left(1 - \frac{1}{k^2}\right)^{1/p} \qquad (\text{taking } \alpha = (k-1/k)^{1/p}) \\
=& 1,
\end{aligned}$$

which implies that $f_{k,j}$ does not converge to 0 in L^p_*.

It is 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_* can be seen from the following sharpening of this statement:

Proposition 1.1.11.

Let $1 \leqslant p < q \leqslant \infty$ and $f \in L^p_*(X,\mu) \cap L^q_*(X,\mu)$, 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

$$\|f\|_r \leqslant \left(\frac{r}{r-p} + \frac{r}{q-r}\right)^{1/r} \|f\|^{1-\theta}_{L^p_*}\|f\|^{\theta}_{L^q_*}, \tag{1.1.7}$$

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 1.1.4 and the definition of the distribution function, it follows that

$$\begin{aligned}
\|f\|^r_r =& r\int_0^\infty \alpha^{r-1} f_*(\alpha)d\alpha \qquad (1.1.8)\\
\leqslant& r\int_0^\infty \alpha^{r-1} \min\left(\frac{\|f\|^p_{L^p_*}}{\alpha^p}, \frac{\|f\|^q_{L^q_*}}{\alpha^q}\right) d\alpha.
\end{aligned}$$

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

$$\begin{aligned}
\|f\|^r_r \leqslant& r\int_0^B \alpha^{r-1-p}\|f\|^p_{L^p_*} d\alpha + r\int_B^\infty \alpha^{r-1-q}\|f\|^q_{L^q_*} d\alpha \\
=& \frac{r}{r-p}\|f\|^p_{L^p_*} B^{r-p} + \frac{r}{q-r}\|f\|^q_{L^q_*} B^{r-q} \qquad (\text{due to } p < r < q)
\end{aligned}$$

$$= \left(\frac{r}{r-p} + \frac{r}{q-r} \right) \|f\|_{L^p_*}^{r(1-\theta)} \|f\|_{L^q_*}^{r\theta}.$$

For the case $q = \infty$, due to $f_*(\alpha) = 0$ for $\alpha > \|f\|_\infty$, we only use the inequality $f_*(\alpha) \leqslant \alpha^{-p} \|f\|_{L^p_*}^p$ for $\alpha \leqslant \|f\|_\infty$ for the integral in (1.1.8) to get

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

which implies the result since $p = r(1-\theta)$ and $L^\infty_* = L^\infty$ in this case. ■

Remark 1.1.12. From the Hölder inequality, we easily know that (1.1.7) holds with constant 1 if L^p_* and L^q_* are replaced by L^p and L^q, respectively.

§1.2 Complex method: Riesz-Thorin and Stein interpolation theorems

§1.2.1 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 the mapping T.

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

Definition 1.2.1.

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

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

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

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 $f = f_1 + f_2$.

Now, we have the following well-known Riesz-Thorin interpolation theorem.

Theorem 1.2.2: The Riesz-Thorin interpolation theorem

Let (X,μ) and (Y,ν) be a pair of σ-finite measure spaces. Let T be a linear operator with domain $(L^{p_0}+L^{p_1})(X,d\mu)$, $p_0,p_1,q_0,q_1\in[1,\infty]$. Assume that

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

and

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

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

$$\frac{1}{p}=\frac{1-\theta}{p_0}+\frac{\theta}{p_1},\quad \frac{1}{q}=\frac{1-\theta}{q_0}+\frac{\theta}{q_1}. \tag{1.2.1}$$

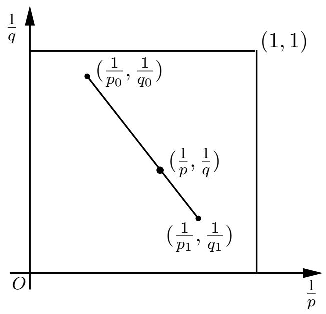
Then

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

with

$$A_\theta\leqslant A_0^{1-\theta}A_1^{\theta}. \tag{1.2.2}$$

Remark 1.2.3. 1) (1.2.2) means that A_θ is logarithmically convex, i.e., $\ln A_\theta$ is convex.
2) The geometrical meaning of (1.2.1) is that the points $(1/p,1/q)$ are the points on the line segment between $(1/p_0,1/q_0)$ and $(1/p_1,1/q_1)$.
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 (1.2.1).



In order to prove the Riesz-Thorin interpolation theorem, we first give the following three lines theorem, which is the basis for the proof and the complex interpolation method, and we will give its proof later. For convenience, let $S=\{z\in\mathbb{C}:0\leqslant\Re z\leqslant 1\}$ be the closed strip, $\mathring{S}=\{z\in\mathbb{C}:0<\Re z<1\}$ be the open strip, and $\partial S=\{z\in\mathbb{C}:\Re z\in\{0,1\}\}$. We have the following.

Theorem 1.2.4: Hadamard three lines theorem

Assume that $f(z)$ is analytic on $\mathring{S}$ and bounded and continuous on S. Then

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

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

We now prove the Riesz-Thorin interpolation theorem with the help of the Hadamard three lines theorem.

Proof of Theorem 1.2.2. 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,

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

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

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

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

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

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

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

Thus,

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

We have

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

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 $|f(x)| > \delta$, then we have

$$||f(x)|^\alpha \ln|f(x)|| \leqslant M^\alpha |\ln|f(x)|| \leqslant M^\alpha \max(|\ln M|, |\ln\delta|) < \infty.$$

Thus, $||f(x)|^\alpha \ln|f(x)|| \leqslant C$. Hence,

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

which yields

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

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

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

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

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

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

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

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

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

Using Hadamard's three line theorem, reproduced as Theorem 1.2.4, we get the conclusion

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

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

Before proving the three line theorem, we recall the following theorem.

Theorem 1.2.5: Phragmen-Lindelöf theorem/Maximum principle

Assume that $f(z)$ is analytic on $\mathring{S}$ and bounded and continuous on S. Then

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

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

$$h(z) = \frac{e^{i\pi z} - i}{e^{i\pi z} + i}, \quad z \in S. \tag{1.2.3}$$

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

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

which implies the statement in this case.

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

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

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

$$|f(z_0)| =|f_{\delta,z_0}(z_0)| \leqslant \max\left(\sup_{t\in\mathbb{R}}|f_{\delta,z_0}(it)|, \sup_{t\in\mathbb{R}}|f_{\delta,z_0}(1+it)|\right)$$

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

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

Proof of Theorem 1.2.4. Denote

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

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

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

Then by Theorem 1.2.5, it follows that $|F_\lambda(z)| \leqslant \max(A_0, e^\lambda A_1)$. Hence,

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

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

Now, we shall give a rather simple application of the Riesz-Thorin interpolation theorem.

Theorem 1.2.6: 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 1.2.7. 1) The sharp form of Young's inequality for convolutions can be found in [Bec75, Theorem 3], we just state it as follows. Under the assumption of Theorem 1.2.6, 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***, i.e., T satisfying for measurable functions f and g:

$$\begin{aligned} |T(\alpha f)| &= |\alpha||T(f)|, \quad \forall \alpha \in \mathbb{C}, \\ |T(f+g)| &\leqslant |T(f)| + |T(g)|. \end{aligned}$$

One can see [CZ56] for details.

§ 1.2.2 Stein interpolation theorem

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

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

$$\int_Y |T_z(\chi_A)\chi_B| d\nu < \infty \tag{1.2.4}$$

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

$$z \to \int_Y T_z(f) g d\nu \tag{1.2.5}$$

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

$$e^{-a|\Im z|} \ln \left| \int_Y T_z(f) g d\nu \right| \leqslant C_{f,g} < \infty \tag{1.2.6}$$

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

$$e^{-a|\Im z|} \ln \left| \int_B T_z(\chi_A) d\nu \right| \leqslant c(A, B), \tag{1.2.7}$$

then (1.2.6) holds for $f = \sum_{k=1}^M a_k \chi_{A_k}$ and $g = \sum_{j=1}^N b_j \chi_{B_j}$ and

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

In fact, by the linearity of T_z, (1.2.7) and the increasing of $\ln$, we get

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

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

Then, we have an extension of the three lines theorem.

Lemma 1.2.8.

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

$$\ln|F(z)| \leqslant Ae^{a|\Im z|} \tag{1.2.8}$$

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 1.2.8, we first recall the Poisson-Jensen formula from [Rub96, pp.21].

Theorem 1.2.9: The Poisson-Jensen formula

Suppose that f is meromorphic 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

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

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 pole at the origin.

Proof of Lemma 1.2.8. 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 on D, by the Poisson-Jensen formula, we have

$$\ln|F(h(z))| \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|F(h(Re^{i\varphi}))|\frac{R^2 - \rho^2}{|Re^{i\varphi} - \rho e^{i\theta}|^2}d\varphi \tag{1.2.9}$$

when $z = \rho e^{i\theta}$ and $|z| = \rho < R$. We observe that for $R < |\zeta| = 1$ the hypothesis on F implies that

$$\begin{aligned}\ln|F(h(Re^{i\varphi}))| \leqslant& Ae^{a\left|\Im\frac{1}{\pi i}\ln\left(i\frac{1+R\zeta}{1-R\zeta}\right)\right|} \quad (\text{let } \zeta = e^{i\varphi}, h(R\zeta) = \frac{1}{\pi i}\ln\left(i\frac{1+R\zeta}{1-R\zeta}\right)) \\ \leqslant& Ae^{\frac{a}{\pi}\left|\ln\left|\frac{1+R\zeta}{1-R\zeta}\right|\right|}\end{aligned}$$

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

(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 get

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

Now,

$$\begin{aligned}\int_{-\pi}^{\pi}|\sin\varphi|^{-\frac{a}{\pi}}d\varphi=&4\int_0^{\frac{\pi}{2}}\sin^{-\frac{a}{\pi}}\varphi d\varphi=4\int_0^{\frac{\pi}{2}}\sin^{2(\frac12-\frac{a}{2\pi})-1}\varphi\cos^{2\cdot\frac12-1}\varphi d\varphi\\ =&2B\left(\frac12,\frac12-\frac{a}{2\pi}\right)<\infty,\end{aligned}$$

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

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

converges for $\alpha,\beta>0$. Moreover, for $1>R>\frac12(\rho+1)$, it holds

$$\begin{aligned}\frac{R^2-\rho^2}{|Re^{i\varphi}-\rho e^{i\theta}|^2}=&\frac{R^2-\rho^2}{R^2-2R\rho\cos(\theta-\varphi)+\rho^2}\leqslant\frac{R^2-\rho^2}{R^2-2R\rho+\rho^2}\\ =&\frac{(R-\rho)(R+\rho)}{(R-\rho)^2}=\frac{R+\rho}{R-\rho}\leqslant\frac{2}{\frac12(\rho+1)-\rho}\leqslant\frac{4}{1-\rho}.\end{aligned}$$

Thus, (1.2.9) is uniformly bounded w.r.t. $R\in(\frac12(\rho+1),1)$.

We will now use the following consequence of Fatou's lemma: suppose that $F_R\leqslant G$, where $G\geqslant 0$ is integrable, then $\limsup_{R\to\infty}\int F_R d\varphi\leqslant\int\limsup_{R\to\infty}F_R d\varphi$. Letting $R\uparrow 1$ in (1.2.9) and using this convergence result, we obtain

$$\ln|F(h(\rho e^{i\theta}))|\leqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|F(h(e^{i\varphi}))|\frac{1-\rho^2}{1-2\rho\cos(\theta-\varphi)+\rho^2}d\varphi. \tag{1.2.10}$$

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

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

from which it follows that $\rho=(\cos\pi x)/(1+\sin\pi x)$ and $\theta=-\pi/2$ when $x\in(0,\frac12]$, while $\rho=-(\cos\pi x)/(1+\sin\pi x)$ and $\theta=\pi/2$ when $x\in[\frac12,1)$. In either case, we have $\rho=(\operatorname{sgn}(\frac12-x))(\cos\pi x)/(1+\sin\pi x)$ and $\theta=-(\operatorname{sgn}(\frac12-x))\pi/2$ for $x\in(0,1)$. We easily deduce that

$$\frac{1-\rho^2}{1-2\rho\cos(\theta-\varphi)+\rho^2}$$

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

Using this we write (1.2.10) as

$$
\ln|F(x)|\leqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin\pi x}{1+\cos\pi x\sin\varphi}\ln|F(h(e^{i\varphi}))|d\varphi. \tag{1.2.11}
$$

We now change 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

$$
\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{0}\frac{\sin\pi x}{1+\cos\pi x\sin\varphi}\ln|F(h(e^{i\varphi}))|d\varphi\\
=&\frac{1}{2}\int_{-\infty}^{\infty}\frac{\sin\pi x}{\cosh\pi t-\cos\pi x}\ln|F(it)|dt.
\end{aligned}\tag{1.2.12}
$$

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

$$
\begin{aligned}
&\frac{1}{2\pi}\int_{0}^{\pi}\frac{\sin\pi x}{1+\cos\pi x\sin\varphi}\ln|F(h(e^{i\varphi}))|d\varphi\\
=&\frac{1}{2}\int_{-\infty}^{\infty}\frac{\sin\pi x}{\cosh\pi t+\cos\pi x}\ln|F(1+it)|dt.
\end{aligned}\tag{1.2.13}
$$

Adding (1.2.12), (1.2.13) and using (1.2.11), 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 $\mathring{S}=\{z\in\mathbb{C}:0<\Re z<1\}$ and continuous on its closure S. Moreover, for some $A<\infty$ and $a\in[0,\pi)$, we have

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

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 1.2.10: 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

$$\sup_{t\in\mathbb{R}} e^{-b|t|} \ln M_j(t) < \infty \tag{1.2.14}$$

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

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \text{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}. \tag{1.2.15}$$

Suppose that

$$\|T_{it}(f)\|_{q_0} \leqslant M_0(t)\|f\|_{p_0}, \quad \|T_{1+it}(f)\|_{q_1} \leqslant M_1(t)\|f\|_{p_1} \tag{1.2.16}$$

for all finitely simple functions f on X. Then

$$\|T_\theta(f)\|_q \leqslant M(\theta)\|f\|_p, \quad \text{when } 0 < \theta < 1 \tag{1.2.17}$$

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 (1.2.15).

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

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

the Stein interpolation theorem proceeds by instead studying the expression

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

One can then repeat the proof of the Riesz-Thorin theorem more or less verbatim to obtain the Stein interpolation theorem. For convenience, we give the proof for this version of finitely simple functions.

Proof of Theorem 1.2.10. Fix $\theta \in (0,1)$ and finitely simple functions f on X and g on Y such that $\|f\|_p = \|g\|_{q'} = 1$. Note that since $\theta \in (0,1)$, we must have $p, q \in (1,\infty)$. Let

$$f = \sum_{k=1}^{m} a_k e^{i\alpha_k}\chi_{A_k} \quad \text{and} \quad g = \sum_{j=1}^{n} b_j e^{i\beta_j}\chi_{B_j},$$

where $a_k > 0$, $b_j > 0$, α_k, β_j are real, A_k are pairwise disjoint subsets of X with finite measure, and B_j are pairwise disjoint subsets of Y with finite measure for all k, j. Let

$$P(z) = \frac{p}{p_0}(1-z) + \frac{p}{p_1}z, \quad Q(z) = \frac{q'}{q_0'}(1-z) + \frac{q'}{q_1'}z. \tag{1.2.18}$$

For $z \in S$, define

$$f_z = \sum_{k=1}^{m} a_k^{P(z)} e^{i\alpha_k}\chi_{A_k}, \quad g_z = \sum_{j=1}^{n} b_j^{Q(z)} e^{i\beta_j}\chi_{B_j}, \tag{1.2.19}$$

and

$$F(z)=\int_Y T_z(f_z)g_z d\nu. \tag{1.2.20}$$

Linearity gives that

$$F(z)=\sum_{k=1}^{m}\sum_{j=1}^{n} a_k^{P(z)} b_j^{Q(z)} e^{i\alpha_k} e^{i\beta_j}\int_Y T_z(\chi_{A_k})(x)\chi_{B_j}(x)d\nu(x),$$

and the condition (1.2.4) together with the fact that $\{T_z\}_z$ is an analytic family imply that $F(z)$ is a well-defined analytic function on the unit strip that extends continuously to its boundary.

Since $\{T_z\}_z$ is a family of admissible growth, (1.2.7) holds for some $c(A_k,B_j)$ and $a\in(0,\pi)$ and this combined with the facts that

$$|a_k^{P(z)}|\leqslant (1+a_k)^{\frac{p}{p_0}+\frac{p}{p_1}},\quad |b_j^{Q(z)}|\leqslant (1+b_j)^{\frac{q'}{q'_0}+\frac{q'}{q'_1}}$$

for all $z\in\mathring{S}$, implies (1.2.8) with a as in (1.2.7) and

$$A=\ln(mn)+\sum_{k=1}^{m}\sum_{j=1}^{n}\left[c(A_k,B_j)+\left(\frac{p}{p_0}+\frac{p}{p_1}\right)\ln(1+a_k)+\left(\frac{q'}{q'_0}+\frac{q'}{q'_1}\right)\ln(1+b_j)\right].$$

Thus, F satisfies the hypotheses of Lemma 1.2.8. Moreover, the calculations in the proof of Theorem 1.2.2 show that (even when $p_0=p_1=\infty$, $q_0=q_1=1$) for $j=0,1$

$$\|f_{j+iy}\|_{p_j}=\|f\|_p^{\frac{p}{p_j}}=1=\|g\|_{q'}^{\frac{q'}{q'_j}}=\|g_{j+iy}\|_{q'_j},\quad \text{when } y\in\mathbb{R}. \tag{1.2.21}$$

Hölder's inequality, (1.2.21) and the hypothesis (1.2.16) give

$$|F(j+iy)|\leqslant \|T_{j+iy}(f_{j+iy})\|_{q_j}\|g_{j+iy}\|_{q'_j}\leqslant M_j(y)\|f_{j+iy}\|_{p_j}\|g_{j+iy}\|_{q'_j}=M_j(y)$$

for all y real and $j=0,1$. These inequalities and the conclusion of Lemma 1.2.8 yield

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

for all $\theta\in(0,1)$. But notice that

$$F(\theta)=\int_Y T_\theta(f)g d\nu. \tag{1.2.22}$$

Taking absolute values and the supremum over all finitely simple functions g on Y with $L^{q'}$ norm equal to one, we conclude the proof of (1.2.17) for finitely simple functions f with L^p norm one. Then (1.2.17) follows by replacing f by $f/\|f\|_p$. ■

§ 1.3 The decreasing rearrangement and Lorentz spaces

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

Definition 1.3.1.

If f is a measurable function on X, the ***decreasing rearrangement*** of f is the function $f^*:[0,\infty)\mapsto[0,\infty]$ defined by

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

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

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

Example 1.3.2. (Decreasing rearrangement of a simple function) Let f be a simple function of the following form

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

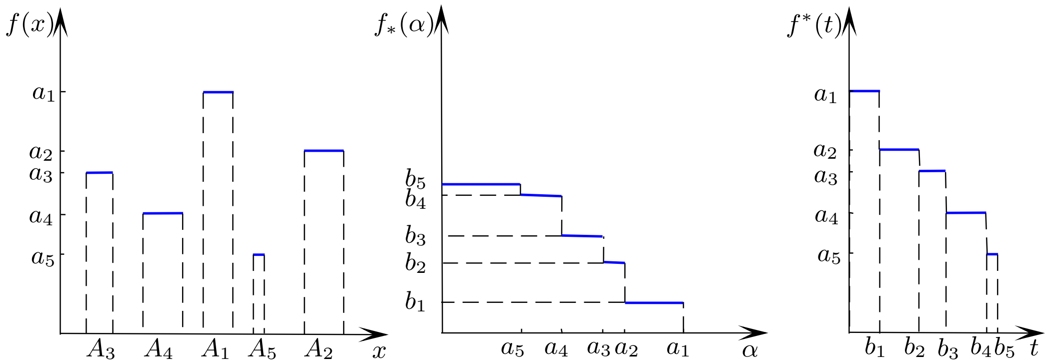
where $a_1 > a_2 > \cdots > a_k > 0$, $A_j = \{x \in \mathbb{R} : f(x) = a_j\}$ and χ_A is the characteristic function of the set A (see Figure (a)). Then

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

where $b_j = \sum_{i=1}^{j} |A_i|$, $B_j = [a_{j+1}, a_j)$ 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)). We can also find the decreasing rearrangement (by denoting $b_0 = 0$)

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

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



(a) (b) (c)

Example 1.3.3. 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,

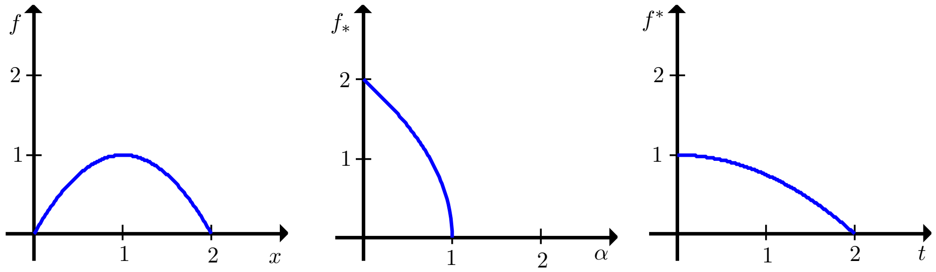
$$f_*(\alpha) = \begin{cases} 2\sqrt{1-\alpha}, & 0 \leqslant \alpha \leqslant 1, \\ 0, & \alpha > 1. \end{cases}$$

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

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

Thus,

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



(a) (b) (c)

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

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

Example 1.3.4. We define an extended function $f : [0, \infty) \mapsto [0, \infty]$ as

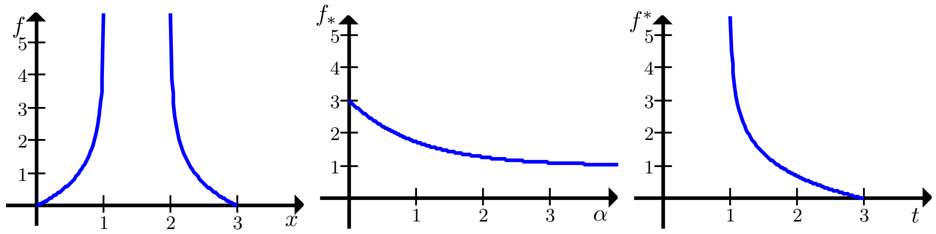
$$f(x) = \begin{cases} 0, & x = 0, \\ \ln \frac{1}{1-x}, & 0 < x < 1, \\ \infty, & 1 \leqslant x \leqslant 2, \\ \ln \frac{1}{x-2}, & 2 < x < 3, \\ 0, & x \geqslant 3. \end{cases}$$

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

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

and

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



(a) (b) (c)

Example 1.3.5. Consider the function $f(x) = x$ for all $x \in [0, \infty)$. Then $f_*(\alpha) = |\{x \in [0, \infty) : x > \alpha\}| = \infty$ for all $\alpha \geqslant 0$, which implies that $f^*(t) = \inf\{\alpha \geqslant 0 : \infty \leqslant t\} = \infty$ for all $t \geqslant 0$.

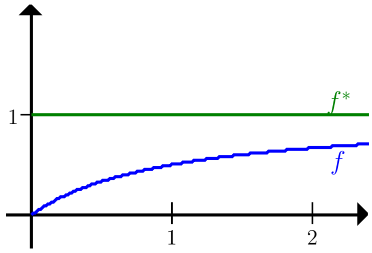
Example 1.3.6. Consider $f(x) = \frac{x}{1+x}$ for $x \geqslant 0$. It is clear that $f_*(\alpha) = 0$ for $\alpha \geqslant 1$ since $|f(x)| < 1$. For $\alpha \in [0, 1)$, we have

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

That is,

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

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



The following are some properties of the function f^*.

Proposition 1.3.7.

The decreasing rearrangement f^* of the measurable function f on (X, μ) has the following properties:

(i) $f^*(t)$ is a non-negative and non-increasing function on $[0, \infty)$.
(ii) $f^*(t)$ is right continuous on $[0, \infty)$.
(iii) $(kf)^* = |k|f^*$ for $k \in \mathbb{C}$.
(iv) $|f| \leqslant |g|$ a.e. implies that $f^* \leqslant g^*$.
(v) $(f + g)^*(t_1 + t_2) \leqslant f^*(t_1) + g^*(t_2)$.
(vi) $(fg)^*(t_1 + t_2) \leqslant f^*(t_1)g^*(t_2)$.
(vii) $|f| \leqslant \liminf_{k\to\infty} |f_k|$ a.e. implies that $f^* \leqslant \liminf_{k\to\infty} f_k^*$.
(viii) $|f_k| \uparrow |f|$ a.e. implies that $f_k^* \uparrow f^*$.
(ix) $f^*(f_*(\alpha)) \leqslant \alpha$ whenever $f_*(\alpha) < \infty$.
(x) $f_*(f^*(t)) = \mu(\{|f| > f^*(t)\}) \leqslant t \leqslant \mu(\{|f| \geqslant f^*(t)\})$ if $f^*(t) < \infty$.
(xi) $f^*(t) > \alpha$ if and only if $f_*(\alpha) > t$.
(xii) f^* is equi-measurable with f, that is, $(f^*)_*(\alpha) = f_*(\alpha)$ for any $\alpha \geqslant 0$.
(xiii) $(|f|^p)^*(t) = (f^*(t))^p$ for $1 \leqslant p < \infty$.
(xiv) $\|f^*\|_p = \|f\|_p$ for $1 \leqslant p < \infty$.
(xv) $\|f\|_\infty = f^*(0)$.
(xvi) $\sup_{t>0} t^s f^*(t) = \sup_{\alpha>0} \alpha(f_*(\alpha))^s$ for $0 < s < \infty$.

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

From (iv) in Proposition 1.1.3, it holds

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

Using the definition of the decreasing rearrangement, we have

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

(vi) Similar to (v), by (v) in Proposition 1.1.3, it holds that $(fg)_*(\alpha_1\alpha_2) \leqslant f_*(\alpha_1)+g_*(\alpha_2) \leqslant t_1+t_2$. Then, we have

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

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

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

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

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

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

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

(xiv) From Theorem 1.1.4 and (xii), we have

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

We remain the proofs of others to interested readers. ■

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

Definition 1.3.8.

Given f a measurable function on a measure space (X,μ) and $1 \leqslant p,q \leqslant \infty$, define

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

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

As in L^p and in weak L^p, two functions in $L^{p,q}$ will be considered equal if they are equal almost everywhere. Observe that the previous definition implies that $L^{p,\infty} = L^p_*$ in view of (xvi) in Proposition 1.3.7 and $L^{p,p} = L^p$ in view of (xiv) in Proposition 1.3.7 for $1 \leqslant p < \infty$. By (i) and (xv) in Proposition 1.3.7, we have $\|f\|_{L^{\infty,\infty}} = \sup_{t>0} f^*(t) = f^*(0) = \|f\|_\infty$ which implies that $L^{\infty,\infty} = L^\infty = L^\infty_*$. Thus, we have

Theorem 1.3.9.

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

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

Remark 1.3.10. For the Lorentz space $L^{p,q}$, the case when $p = \infty$ and $1 \leqslant q < \infty$ is not of any interest. The reason is that $\|f\|_{L^{\infty,q}} < \infty$ implies that $f = 0$ a.e. on X. In fact, assume that $L^{\infty,q}$ is a non-trivial space, there exists a nonzero function $f \in L^{\infty,q}$ on a nonzero measurable set, that is, there exists a constant $c > 0$ and a set E of positive measure such that $|f(x)| > c$ for all $x \in E$. Then, by (iv) in Proposition 1.3.7, we have

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

since $(f\chi_E)^*(t) = 0$ for $t > \mu(E)$. Hence, we have a contradiction. Thus, $f = 0$ a.e. on X.

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

Theorem 1.3.11.

Let $1 \leqslant p \leqslant \infty$ and $1 \leqslant q < r \leqslant \infty$. Then,

$$\|f\|_{L^{p,r}} \leqslant C_{p,q,r} \|f\|_{L^{p,q}}, \tag{1.3.1}$$

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

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

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

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

$$\|f\|_{L^{p,\infty}} \leqslant \left(\frac{q}{p} \right)^{1/q} \|f\|_{L^{p,q}}. \tag{1.3.2}$$

This establishes (1.3.1) in the case $r = \infty$. Finally, when $q < r < \infty$, we have by (1.3.2)

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

This completes the proof. ■

In general, $L^{p,q}$ is a quasi-normed space, since the functional $\|\cdot\|_{L^{p,q}}$ satisfies the conditions of normed spaces except the triangle inequality. In fact, by (v) in Proposition 1.3.7, it holds

$$\|f+g\|_{L^{p,q}} \leqslant 2^{1/p}(\|f\|_{L^{p,q}} + \|g\|_{L^{p,q}}). \tag{1.3.3}$$

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

Theorem 1.3.12.

Let (X,μ) be a measure space. Then for all $1 \leqslant p,q \leqslant \infty$, the spaces $L^{p,q}(X,\mu)$ are complete with respect to their quasi-norms, and they are therefore quasi-Banach spaces.

Proof. The proof is standard, we omit the details. One can see [Gra14, p.54, Theorem 1.4.11] for details. ■

For the dual of Lorentz spaces, we have

Theorem 1.3.13.

Suppose that (X,μ) is a non-atomic σ-finite measure space. Let $1 < p,q < \infty$, $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$. Then we have

$$(L^{p,q})' = L^{p',q'}, \quad (L^{1,1})' = (L^1)' = L^\infty, \quad (L^{1,q})' = \{0\}, \quad (L^{p,1})' = L^{p',\infty}.$$

Proof. See [Gra14, p. 57-60, Theorem 1.4.16]. ■

For more results, one can see [Gra14, Kri02].

§ 1.4 Real method: Marcinkiewicz' interpolation theorem

We first introduce the definition of quasi-linear operators.

Definition 1.4.1.

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

The idea we have used, in Definition 1.2.1, of splitting f into two parts according to their respective size, is the main idea of the proof of the theorem that follows. There, we will also use two easily proved inequalities, which are well-known results of Hardy's (see [HLP88, p. 245–246]):

Lemma 1.4.2: Hardy inequalities

If $q \geqslant 1$, $r > 0$ and g is a measurable, non-negative function on $(0,\infty)$, then

$$\left(\int_0^\infty \left(\int_0^t g(y)dy\right)^q t^{-r}\frac{dt}{t}\right)^{1/q} \leqslant \frac{q}{r}\left(\int_0^\infty (yg(y))^q y^{-r}\frac{dy}{y}\right)^{1/q}, \tag{1.4.1}$$

$$\left(\int_0^\infty \left(\int_t^\infty g(y)dy\right)^q t^r \frac{dt}{t}\right)^{1/q} \leqslant \frac{q}{r}\left(\int_0^\infty (yg(y))^q y^r \frac{dy}{y}\right)^{1/q}. \tag{1.4.2}$$

Proof. To prove (1.4.1), we use Jensen's inequality with the convex function $\varphi(x) = x^q$ on $(0, \infty)$. Then

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

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

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

which yields (1.4.1) immediately.

To prove (1.4.2), we denote $f(x) = g(1/x)/x^2$. Then by taking $t = 1/s$ and $y = 1/x$, and then applying (1.4.1) and changing variable again by $x = 1/y$, we obtain

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

Thus, we complete the proofs. ∎

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

Theorem 1.4.3: Marcinkiewicz interpolation theorem

Let (X, μ) and (Y, ν) be a pair of σ-finite measure spaces. Assume that $1 \leqslant p_i \leqslant q_i \leqslant \infty$, $p_0 < p_1$, $q_0 \neq q_1$ and T is a quasi-linear mapping, defined on $L^{p_0}(X) + L^{p_1}(X)$, which is simultaneously of weak types (p_0, q_0) and (p_1, q_1), i.e.,

$$\begin{aligned}\|Tf\|_{L^{q_0,\infty}(Y)} \leqslant & A_0\|f\|_{p_0}, \quad \forall f\in L^{p_0}(X),\\ \|Tf\|_{L^{q_1,\infty}(Y)} \leqslant & A_1\|f\|_{p_1}, \quad \forall f\in L^{p_1}(X).\end{aligned} \tag{1.4.3}$$

If $0<\theta<1$, and

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

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

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

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

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

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

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

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

$$f=f^t+f_t$$

where

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

and $f_t=f-f^t$.

Then we can verify that

$$\begin{aligned}(f^t)^*(y)&\begin{cases}\leqslant f^*(y), & 0\leqslant y\leqslant t^\sigma,\\ =0, & y>t^\sigma,\end{cases}\\ (f_t)^*(y)&\leqslant\begin{cases} f^*(t^\sigma), & 0\leqslant y\leqslant t^\sigma,\\ f^*(y), & y>t^\sigma.\end{cases}\end{aligned} \tag{1.4.4}$$

In fact, by (iv) in Proposition 1.3.7, $|f^t|\leqslant|f|$ implies $(f^t)^*(y)\leqslant f^*(y)$ for all $y\geqslant 0$. Moreover, since for $0\leqslant\alpha\leqslant f^*(t^\sigma)$

$$\begin{aligned}(f^t)_*(\alpha)=&\mu(\{x:|f^t(x)|>\alpha\})=\mu(\{x:|f(x)|>f^*(t^\sigma),\text{and }|f(x)|>\alpha\})\\ =&\mu(\{x:|f(x)|>f^*(t^\sigma)\})=f_*(f^*(t^\sigma)),\end{aligned}$$

by the definition of f^t and (x) in Proposition 1.3.7, we have

$$(f^t)_*(\alpha)\leqslant(f^t)_*(f^*(t^\sigma))=f_*(f^*(t^\sigma))\leqslant t^\sigma, \quad \forall\alpha\geqslant 0.$$

Thus, for $y>t^\sigma$, we get $(f^t)^*(y)=0$.

Similarly, by (iv) in Proposition 1.3.7, we have $(f_t)^*(y)\leqslant f^*(y)$ for any $y\geqslant 0$ since $|f_t|\leqslant|f|$. On the other hand, for $y\geqslant 0$, we have $(f_t)^*(y)\leqslant(f_t)^*(0)=\|f_t\|_\infty\leqslant f^*(t^\sigma)$ with the help of the non-increasing of $(f_t)^*(y)$ and (xv) in Proposition 1.3.7. Thus, $(f_t)^*(y)\leqslant\min(f^*(y),f^*(t^\sigma))$ for any $y\geqslant 0$ which implies (1.4.4).

Suppose $p_1<\infty$. Notice that $p\leqslant q$, because $p_i\leqslant q_i$. Denote $K_{p,q}=K(p/q)^{1/p-1/q}$. By Theorems 1.3.9 and 1.3.11, (1.4.3), (1.4.4) and then by a change of variables, Hardy's inequalities (1.4.1) and (1.4.2), and (xiv) in Proposition 1.3.7, we get

$$\|Tf\|_q\leqslant K(\|Tf^t\|_q+\|Tf_t\|_q)=K(\|Tf^t\|_{L^{q,q}}+\|Tf_t\|_{L^{q,q}})$$

$$
\begin{aligned}
&\leqslant K(p/q)^{1/p-1/q}(\|Tf^t\|_{L^{q,p}}+\|Tf_t\|_{L^{q,p}})\\
&=K_{p,q}\left\{\left(\int_0^\infty \left[t^{1/q}(Tf^t)^*(t)\right]^p\frac{dt}{t}\right)^{1/p}+\left(\int_0^\infty \left[t^{1/q}(Tf_t)^*(t)\right]^p\frac{dt}{t}\right)^{1/p}\right\}\\
&\leqslant K_{p,q}\left\{A_0\left(\int_0^\infty \left[t^{1/q-1/q_0}\|f^t\|_{p_0}\right]^p\frac{dt}{t}\right)^{1/p}\right.\\
&\qquad\qquad\left.+A_1\left(\int_0^\infty \left[t^{1/q-1/q_1}\|f_t\|_{p_1}\right]^p\frac{dt}{t}\right)^{1/p}\right\}\\
&\leqslant K_{p,q}\left\{A_0\left(\int_0^\infty \left[t^{1/q-1/q_0}\left(\frac{1}{p_0}\right)^{1-1/p_0}\|f^t\|_{L^{p_0,1}}\right]^p\frac{dt}{t}\right)^{1/p}\right.\\
&\qquad\qquad\left.+A_1\left(\int_0^\infty \left[t^{1/q-1/q_1}\left(\frac{1}{p_1}\right)^{1-1/p_1}\|f_t\|_{L^{p_1,1}}\right]^p\frac{dt}{t}\right)^{1/p}\right\}\\
&\leqslant K_{p,q}\left\{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}\left(\int_0^\infty \left[t^{1/q-1/q_0}\left(\int_0^{t^\sigma} y^{1/p_0}f^*(y)\frac{dy}{y}\right)\right]^p\frac{dt}{t}\right)^{1/p}\right.\\
&\qquad+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty \left[t^{1/q-1/q_1}\left(\int_{t^\sigma}^\infty y^{1/p_1}f^*(y)\frac{dy}{y}\right)\right]^p\frac{dt}{t}\right)^{1/p}\\
&\qquad\left.+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty \left[t^{1/q-1/q_1}\left(\int_0^{t^\sigma} y^{1/p_1}f^*(t^\sigma)\frac{dy}{y}\right)\right]^p\frac{dt}{t}\right)^{1/p}\right\}\\
&=K_{p,q}|\sigma|^{-\frac{1}{p}}\left\{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}\left(\int_0^\infty s^{-p(1/p_0-1/p)}\left(\int_0^s y^{1/p_0}f^*(y)\frac{dy}{y}\right)^p\frac{ds}{s}\right)^{1/p}\right.\\
&\qquad+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty s^{p(1/p-1/p_1)}\left(\int_s^\infty y^{1/p_1}f^*(y)\frac{dy}{y}\right)^p\frac{ds}{s}\right)^{1/p}\\
&\qquad\left.+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty s^{p(1/p-1/p_1)}\left(\int_0^s y^{1/p_1}f^*(s)\frac{dy}{y}\right)^p\frac{ds}{s}\right)^{1/p}\right\}\\
&\leqslant K_{p,q}|\sigma|^{-\frac{1}{p}}\left\{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}\frac{1}{(1/p_0-1/p)}\left(\int_0^\infty \left(y^{1/p}f^*(y)\right)^p\frac{dy}{y}\right)^{1/p}\right.\\
&\qquad+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\frac{1}{(1/p-1/p_1)}\left(\int_0^\infty \left(y^{1/p}f^*(y)\right)^p\frac{dy}{y}\right)^{1/p}\\
&\qquad\left.+A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}\left(\int_0^\infty s^{1-p/p_1}(p_1s^{1/p_1}f^*(s))^p\frac{ds}{s}\right)^{1/p}\right\}\\
&=K_{p,q}|\sigma|^{-1/p}\left\{\frac{A_0\left(\frac{1}{p_0}\right)^{1-1/p_0}}{\frac{1}{p_0}-\frac{1}{p}}+\frac{A_1\left(\frac{1}{p_1}\right)^{1-1/p_1}}{\frac{1}{p}-\frac{1}{p_1}}+A_1p_1^{1/p_1}\right\}\|f\|_p\\
&=A\|f\|_p.
\end{aligned}
$$

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

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

Thus, we complete the proof. ■

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 [BL76, Chapter 5].

References

[Bec75] William Beckner. Inequalities in Fourier analysis on R^n. *Proc. Nat. Acad. Sci. U.S.A.*, 72:638–641, 1975.

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[CZ56] A. P. Calderon and A. Zygmund. A note on the interpolation of sublinear operations. *American Journal of Mathematics*, 78(2):282–288, 1956.

[Gra14] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[HLP88] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.

[HW64] Richard A. Hunt and Guido Weiss. The Marcinkiewicz interpolation theorem. *Proc. Amer. Math. Soc.*, 15:996–998, 1964.

[Kri02] Erik Kristiansson. Decreasing rearrangement and Lorentz $L(p,q)$ spaces. Master's thesis, Luleå University of Technology, 2002.

[Rub96] Lee A. Rubel. *Entire and meromorphic functions*. Universitext. Springer-Verlag, New York, 1996. With the assistance of James E. Colliander.

[SW71] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

II

Fourier Transform and Tempered Distributions

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

§ 2.1 Fourier transform of L^1 functions

§ 2.1.1 The definition and properties

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

Definition 2.1.1.

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

$$\widehat{f}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i x \cdot \xi} f(x) dx \tag{2.1.1}$$

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

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

$$\tau^y f(x) = f(x - y), \tag{2.1.2}$$

$$\delta^a f(x) = f(ax), \tag{2.1.3}$$

$$\widetilde{f}(x) = f(-x).$$

Proposition 2.1.2.

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: $\widehat{\tau^y f}(\xi) = e^{-\omega i y\cdot\xi}\widehat{f}(\xi)$.

(iii) Modulation: $\widehat{(e^{\omega i x\cdot y} f(x))}(\xi) = \tau^y \widehat{f}(\xi)$.

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

(v) Differentiation: $\widehat{\partial^\alpha f}(\xi) = (\omega i\xi)^\alpha \widehat{f}(\xi)$, $\partial^\alpha \widehat{f}(\xi) = \widehat{((-\omega i x)^\alpha f(x))}(\xi)$.

(vi) Convolution: $\left(\frac{|\omega|}{2\pi}\right)^{n/2} \widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$ and $\widehat{fg} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \widehat{f} * \widehat{g}$.

(vii) Transformation: $\widehat{f\circ A}(\xi) = \widehat{f}(A\xi)$, where A is an orthogonal matrix and ξ is a column vector.

(viii) Conjugation: $\widehat{\overline{f}} = \overline{\widehat{\widetilde{f}}}$.

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

$$\begin{aligned}
\mathscr{F}(f\circ A)(\xi) =& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} f(Ax)dx \\
=& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i A^{-1}y\cdot\xi} f(y)dy \\
=& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i A^{\top}y\cdot\xi} f(y)dy \\
=& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i y\cdot A\xi} f(y)dy \\
=& \widehat{f}(A\xi),
\end{aligned}$$

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

Corollary 2.1.3.

(i) The Fourier transform of a radial function is radial.

(ii) Products and convolutions of radial functions are radial.

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

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

by (vii) in Proposition 2.1.2. Products and convolutions of radial functions are easily seen to be radial. ■

It is easy to establish the following results:

Theorem 2.1.4: Uniform continuity

(i) $\|\widehat{f}\|_\infty \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|f\|_1$.

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

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

$$\widehat{f}(\xi+h)-\widehat{f}(\xi)=\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}[e^{-\omega ix\cdot h}-1]f(x)dx,$$

we have

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

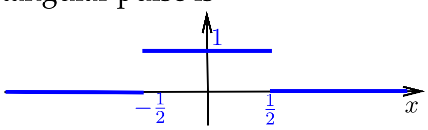
since for any $\theta \geqslant 0$

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

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

Example 2.1.5. Suppose that a signal consists of a single rectangular pulse of width 1 and height 1. Let's say that it gets turned on at $x = -\frac{1}{2}$ and turned off at $x = \frac{1}{2}$. The standard name for this "normalized" rectangular pulse is

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



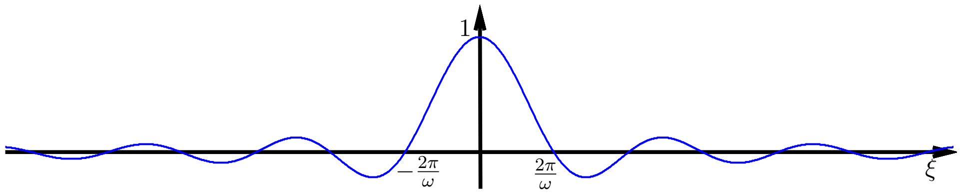
It is also called, variously, the normalized boxcar function, the top hat function, the indicator function, or the characteristic function for the interval $(-1/2, 1/2)$. The Fourier transform of this signal is

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

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

$$\lim_{\xi\to 0}\widehat{\Pi}(\xi)=\left(\frac{|\omega|}{2\pi}\right)^{1/2}\lim_{\xi\to 0}2\frac{\sin\frac{\omega\xi}{2}}{\omega\xi}=\left(\frac{|\omega|}{2\pi}\right)^{1/2}\lim_{\xi\to 0}2\frac{\frac{\omega}{2}\cos\frac{\omega\xi}{2}}{\omega}=\left(\frac{|\omega|}{2\pi}\right)^{1/2}=\widehat{\Pi}(0),$$

so $\widehat{\Pi}(\xi)$ is continuous at $\xi = 0$. There is a standard function called "sinc" that is defined by $\text{sinc}(\xi) = \frac{\sin\xi}{\xi}$ for $\xi \neq 0$ and $\text{sinc}(0) = 1$ for the unnormalized version. In this notation $\widehat{\Pi}(\xi) = \text{sinc}\frac{\omega\xi}{2}$. Here is the graph of $\widehat{\Pi}(\xi)$.



Remark 2.1.6. The above definition of the Fourier transform in (2.1.1) extends immediately to finite Borel measures: if μ is such a measure on $\mathbb{R}^n$, we define $\mathscr{F}\mu$ by letting

$$\mathscr{F}\mu(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} d\mu(x).$$

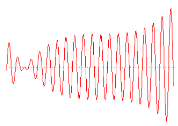
Theorem 2.1.4 is valid for this Fourier transform if we replace the L^1 norm by the total variation of μ.

§ 2.1.2 Riemann-Lebesgue lemma

The following theorem plays a central role in Fourier Analysis. It takes its name from the fact that it holds even for functions that are integrable according to the definition of Lebesgue.

Theorem 2.1.7: Riemann-Lebesgue lemma

If $f \in L^1(\mathbb{R}^n)$, then $\lim\limits_{|\xi|\to\infty} \widehat{f}(\xi) = 0$; thus, in view of the last result, we can conclude that $\widehat{f} \in \mathcal{C}_0(\mathbb{R}^n)$ of all continuous functions vanishing at infinity.



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

Proof. We first consider the case when $f \in \mathscr{D}(\mathbb{R}^n) := \mathcal{C}_c^\infty(\mathbb{R}^n)$ of all $\mathcal{C}^\infty$ functions with compact support. Integrating by parts gives $|\omega\xi|^2\widehat{f}(\xi) = -\widehat{\Delta f}(\xi)$ for each $\xi \in \mathbb{R}^n \setminus \{0\}$, where $\Delta f := \sum\limits_{j=1}^n \partial_j^2 f$. Hence,

$$|\widehat{f}(\xi)| \leqslant \frac{|\widehat{\Delta f}(\xi)|}{|\omega\xi|^2} \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} \frac{\|\Delta f\|_1}{|\omega\xi|^2}, \quad \forall \xi \in \mathbb{R}^n \setminus \{0\}, \tag{2.1.4}$$

from which it is clear that $\lim\limits_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ in this case.

Consider now the case when f is an arbitrary function in $L^1(\mathbb{R}^n)$. Since $\mathscr{D}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, for each fixed $\varepsilon > 0$, there exists a $g \in \mathscr{D}(\mathbb{R}^n)$ such that $\|f-g\|_1 < \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \frac{\varepsilon}{2}$. Then, there is an M such that $|\widehat{g}(\xi)| < \frac{\varepsilon}{2}$ for $|\xi| > M$, since $\widehat{g}(\xi) \to 0$ as $|\xi| \to \infty$. It follows that

$$|\widehat{f}(\xi)| \leqslant |\widehat{f}(\xi) - \widehat{g}(\xi)| + |\widehat{g}(\xi)| \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|f-g\|_1 + |\widehat{g}(\xi)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2},$$

provided $|\xi| > M$ by Theorem 2.1.4. This implies that $|\widehat{f}(\xi)| \to 0$ as $|\xi| \to \infty$. ∎

Theorem 2.1.7 gives a necessary condition for a function to be a Fourier transform. However, that belonging to $\mathcal{C}_0$ is not a sufficient condition for being the Fourier transform of an integrable function. See the following example.

Example 2.1.8. Suppose, for simplicity, that $n = 1$. Let

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

$$g(\xi) = -\, g(-\xi), \quad \xi < 0.$$

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

Assume that there exists an $f \in L^1(\mathbb{R})$ such that $\widehat{f}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{1/2} g(\xi)$, i.e.,

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

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

$$g(\xi) = \int_{-\infty}^{\infty} e^{-\omega i x \xi} f(x) dx = -i \int_{-\infty}^{\infty} \sin(\omega x \xi) f(x) dx = \int_{0}^{\infty} \sin(\omega x \xi) F(x) dx,$$

where $F(x) = i[f(-x) - f(x)] \in L^1(\mathbb{R})$. Integrating $\frac{g(\xi)}{\xi}$ over $(0, N)$ yields

$$\begin{aligned} \int_0^N \frac{g(\xi)}{\xi} d\xi &= \int_0^\infty F(x) \left(\int_0^N \frac{\sin(\omega x \xi)}{\xi} d\xi \right) dx \\ &= \int_0^\infty F(x) \left(\int_0^{\omega x N} \frac{\sin t}{t} dt \right) dx. \end{aligned}$$

Noticing that

$$\left| \int_a^b \frac{\sin t}{t} dt \right| \leqslant C, \qquad \int_0^\infty \frac{\sin t}{t} dt = \frac{\pi}{2} \quad \text{(i.e. Dirichlet integral)},$$

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

$$\lim_{N \to \infty} \int_0^N \frac{g(\xi)}{\xi} d\xi = \frac{\pi}{2} \int_0^\infty F(x) dx < \infty,$$

which yields $\int_e^\infty \frac{g(\xi)}{\xi} d\xi < \infty$ since $\int_0^e \frac{g(\xi)}{\xi} d\xi = 1$. However,

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

This contradiction indicates that the assumption was invalid.

§ 2.1.3 Approximate identities

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

$$C \int_{\mathbb{R}^n} e^{\omega i x \cdot \xi} \widehat{f}(\xi) d\xi. \tag{2.1.5}$$

Unfortunately, $\widehat{f}$ need not be integrable, for example, let $n=1$ and f be the characteristic function of a finite interval, as in Example 2.1.5, we have

$$\begin{aligned}\int_0^\infty |\mathrm{sinc}(x)|dx=&\int_0^\infty \left|\frac{\sin x}{x}\right|dx=\sum_{k=0}^\infty \int_{k\pi}^{(k+1)\pi}\left|\frac{\sin x}{x}\right|dx\\ =&\sum_{k=0}^\infty\int_0^\pi \frac{\sin x}{k\pi+x}dx\geqslant \sum_{k=0}^\infty\int_0^\pi\frac{\sin x}{(k+1)\pi}dx\\ =&\sum_{k=0}^\infty \frac{2}{(k+1)\pi}=\infty.\end{aligned}$$

In order to get around this difficulty, we shall use certain summability methods for integrals. We first introduce the ***Abel method of summability***, whose analog for series is very well-known. For each $\varepsilon>0$, we define the Abel means $A_\varepsilon=A_\varepsilon(f)$ to be the integral

$$A_\varepsilon(f)=A_\varepsilon=\int_{\mathbb{R}^n} e^{-\varepsilon|x|}f(x)dx. \tag{2.1.6}$$

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

$$\lim_{\varepsilon\to 0}A_\varepsilon(f)=\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}e^{-\varepsilon|x|}f(x)dx \tag{2.1.7}$$

may exist even when f is not integrable. A classical example of such a case is obtained by letting $f(x)=\mathrm{sinc}(x)$ when $n=1$, as a similar way as in Example 2.1.8. Whenever the limit in (2.1.7) exists and is finite, we say that $\int_{\mathbb{R}^n} fdx$ is Abel summable to this limit.

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

$$G_\varepsilon(f)=\int_{\mathbb{R}^n}e^{-\varepsilon|x|^2}f(x)dx. \tag{2.1.8}$$

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

$$\lim_{\varepsilon\to 0}G_\varepsilon(f)=\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}e^{-\varepsilon|x|^2}f(x)dx \tag{2.1.7$'$}$$

exists and equals the number ℓ.

We see that both (2.1.7) and (2.1.7′) can be put in the form

$$M_{\varepsilon,\Phi}(f)=M_\varepsilon(f)=\int_{\mathbb{R}^n}\Phi(\varepsilon x)f(x)dx, \tag{2.1.9}$$

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

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

Theorem 2.1.9.

For all $a>0$, we have

$$\mathscr{F}e^{-a|\omega x|^2}(\xi)=(2|\omega|a)^{-n/2}e^{-\frac{|\xi|^2}{4a}}. \tag{2.1.10}$$

Proof. The integral in question is

$$\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi}e^{-a|\omega x|^2}dx.$$

Notice that this factors as a product of one variable integrals. Thus it is sufficient to prove the case $n=1$. It is clear that

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

We observe that the function

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

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

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

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

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

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. ■

For the special cases, we have a fixed point for special definitions of the Fourier transform.

Corollary 2.1.10.

It holds

$$\widehat{e^{-\frac{|\omega||x|^2}{2}}}(\xi) = e^{-\frac{|\omega||\xi|^2}{2}}. \tag{2.1.11}$$

Proof. It is clear by taking $a=1/(2|\omega|)$ in (2.1.10). ■

The second one is somewhat harder to obtain:

Theorem 2.1.11.

For all $a>0$, we have

$$\widehat{e^{-a|\omega x|}}(\xi)=\left(\frac{|\omega|}{2\pi}\right)^{-n/2}\frac{c_n a}{(a^2+|\xi|^2)^{(n+1)/2}},\quad c_n=\frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}. \tag{2.1.12}$$

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

$$\mathscr{F}(e^{-a|\omega x|})=\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi}e^{-a|\omega x|}dx=\left(\frac{|\omega|}{2\pi}\right)^{n/2}(a|\omega|)^{-n}\int_{\mathbb{R}^n} e^{-ix\cdot\xi/a}e^{-|x|}dx,$$

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

$$e^{-\gamma} = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-\eta}}{\sqrt{\eta}} e^{-\gamma^2/4\eta} d\eta, \quad \gamma > 0. \tag{2.1.13}$$

Then, using (2.1.10) to establish the third equality,

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

Thus,

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

Consequently, the theorem will be established once we show (2.1.13). In fact, by changes of variables, we have

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

which yields the desired identity (2.1.13). ■

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

$$W(\xi, a) = (4\pi a)^{-n/2} e^{-\frac{|\xi|^2}{4a}}, \quad P(\xi, a) = \frac{c_n a}{(a^2+|\xi|^2)^{(n+1)/2}}. \tag{2.1.14}$$

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

Theorem 2.1.12: 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 2.1.13.

If f and Φ belong to $L^1(\mathbb{R}^n)$, $\varphi = \widehat{\Phi}$ and $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$, then
$$\int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}\Phi(\varepsilon\xi)\widehat{f}(\xi)d\xi = \int_{\mathbb{R}^n} \varphi_\varepsilon(y-x)f(y)dy$$
for all $\varepsilon > 0$. In particular,
$$\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}e^{-\varepsilon|\omega\xi|}\widehat{f}(\xi)d\xi = \int_{\mathbb{R}^n} P(y-x,\varepsilon)f(y)dy,$$
and
$$\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}e^{-\varepsilon|\omega\xi|^2}\widehat{f}(\xi)d\xi = \int_{\mathbb{R}^n} W(y-x,\varepsilon)f(y)dy.$$

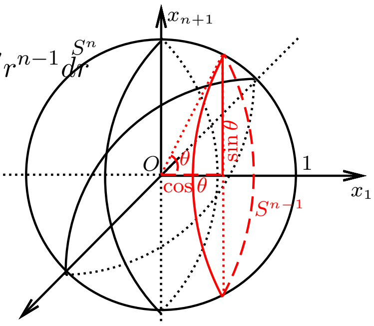
Proof. From (iii) and (iv) in Proposition 2.1.2, it implies $(\mathscr{F} e^{\omega ix\cdot\xi}\Phi(\varepsilon\xi))(y) = \varphi_\varepsilon(y-x)$. The first result holds immediately with the help of Theorem 2.1.12. The last two follow from (2.1.10), (2.1.12) and (2.1.14). ■

Lemma 2.1.14.

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

Proof. By a change of variable, we first note that
$$\int_{\mathbb{R}^n} W(x,\varepsilon)dx = \int_{\mathbb{R}^n} (4\pi\varepsilon)^{-n/2}e^{-\frac{|x|^2}{4\varepsilon}}dx = \int_{\mathbb{R}^n} W(x,1)dx,$$
and
$$\int_{\mathbb{R}^n} P(x,\varepsilon)dx = \int_{\mathbb{R}^n} \frac{c_n\varepsilon}{(\varepsilon^2+|x|^2)^{(n+1)/2}}dx = \int_{\mathbb{R}^n} P(x,1)dx.$$
Thus, it suffices to prove the lemma when $\varepsilon = 1$. For the first one, we use a change of variables and the formula for the integral of a Gaussian: $\int_{\mathbb{R}} e^{-\pi x^2}dx = 1$ to get
$$\int_{\mathbb{R}^n} W(x,1)dx = \int_{\mathbb{R}^n} (4\pi)^{-n/2}e^{-\frac{|x|^2}{4}}dx = \int_{\mathbb{R}^n} (4\pi)^{-n/2}e^{-\pi|y|^2}2^n\pi^{n/2}dy = 1.$$
For the second one, we have
$$\int_{\mathbb{R}^n} P(x,1)dx = c_n\int_{\mathbb{R}^n} \frac{1}{(1+|x|^2)^{(n+1)/2}}dx.$$
Letting $r = |x|$, $x' = x/r$ (when $x \neq 0$), $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$, dx' the element of surface area on S^{n-1} whose surface area is denoted by ω_{n-1} and, finally, putting $r = \tan\theta$, we have

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



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

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

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

Theorem 2.1.15.

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

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

In particular, the ***Poisson integral of f***:

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

and the ***Gauss-Weierstrass integral of f***:

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

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

Proof. By a change of variables, we have

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

Hence,

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

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

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

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

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

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

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

This completes the proof. ■

With the same argument, we have

Corollary 2.1.16.

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

Proof. Once we observe that

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

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

In particular, we also have

Corollary 2.1.17.

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

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

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

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

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

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

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

§ 2.1.4 Fourier inversion

From Theorems 2.1.13 and 2.1.15, we obtain the following solution to the Fourier inversion problem:

Theorem 2.1.18.

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

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

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

Theorem 2.1.19: Fourier inversion theorem

If both f and $\widehat{f}$ are integrable, then

$$f(x) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}\widehat{f}(\xi)d\xi, \quad \forall x \text{ a.e.}$$

Remark 2.1.20. We know from Theorem 2.1.4 that $\widehat{f}$ is continuous. If $\widehat{f}$ is integrable, the integral $\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}\widehat{f}(\xi)d\xi$ also defines a continuous function (in fact, it equals $\widehat{\widehat{f}}(-x)$). Thus, by changing f on a set of measure 0, we can obtain equality in Theorem 2.1.19 for all x.

It is clear from Theorem 2.1.18 that if $\widehat{f}(\xi) = 0$ for all ξ then $f(x) = 0$ a.e. Applying this to $f = f_1 - f_2$, we obtain the following uniqueness result for the Fourier transform:

Corollary 2.1.21: Uniqueness

If f_1 and f_2 belong to $L^1(\mathbb{R}^n)$ and $\widehat{f}_1(\xi) = \widehat{f}_2(\xi)$ for $\xi \in \mathbb{R}^n$, then $f_1 = f_2$ a.e.

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

$$\check{f}(x) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}f(\xi)d\xi. \tag{2.1.15}$$

We give a very useful result.

Theorem 2.1.22.

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

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

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

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

for almost every x.

Proof. By Theorem 2.1.13, we have

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

From Lemma 2.1.14, we get, for any $\delta > 0$,

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

Since f is continuous at 0, for any given $\sigma > 0$, we can choose δ small enough such that $|f(y) - f(0)| \leqslant \sigma$ when $|y| < \delta$. Thus, $I_1 \leqslant \sigma$ by Lemma 2.1.14. For the second term, we have, by a change of variables, that

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

as $\varepsilon \to 0$. Thus, $\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\varepsilon|\omega\xi|}\widehat{f}(\xi)d\xi \to f(0)$ as $\varepsilon \to 0$. On the other hand, by Levi monotone convergence theorem due to $0 \leqslant e^{-\varepsilon|\omega\xi|}\widehat{f}(\xi) \uparrow \widehat{f}(\xi)$ as $\varepsilon \downarrow 0$, we obtain

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

which implies $\widehat{f} \in L^1(\mathbb{R}^n)$ due to $\widehat{f} \geqslant 0$. Therefore, from Theorem 2.1.19, it follows the desired result. ■

An immediate consequence is

Corollary 2.1.23.

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

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

Proof. Noticing that

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

we have the desired results by Theorem 2.1.22. ■

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

Corollary 2.1.24.

If α_1 and α_2 are positive real numbers, then

(i) $W(\xi,\alpha_1+\alpha_2)=\displaystyle\int_{\mathbb{R}^n}W(\xi-\eta,\alpha_1)W(\eta,\alpha_2)d\eta.$

(ii) $P(\xi,\alpha_1+\alpha_2)=\displaystyle\int_{\mathbb{R}^n}P(\xi-\eta,\alpha_1)P(\eta,\alpha_2)d\eta.$

Proof. It follows, from Corollary 2.1.23 and the Fubini theorem, that

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

A similar argument can give the other equality. ■

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

Example 2.1.25. Consider the Cauchy problem to the ***heat equation***

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

Taking the Fourier transform, we have

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

Thus, it follows, from Theorem 2.1.9, that

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

Then, we obtain

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

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

§ 2.2 Fourier transform on L^p for $1 < p \leqslant 2$

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 2.2.1: Plancherel theorem

If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then

$$\|\widehat{f}\|_2 = \|f\|_2.$$

Proof. Let $g(x) = \overline{f(-x)} \in L^1(\mathbb{R}^n)$. Then, $h = f * g \in L^1(\mathbb{R}^n)$ by Theorem 1.2.6 and $\widehat{h} = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{f}\widehat{g}$ by Proposition 2.1.2. But $\widehat{g} = \overline{\widehat{f}}$, thus $\widehat{h} = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} |\widehat{f}|^2 \geqslant 0$. In addition, h is continuous at $\{0\}$, since

$$\begin{aligned} |h(x) - h(0)| &= \left| \int_{\mathbb{R}^n} f(x-y)g(y)dy - \int_{\mathbb{R}^n} f(-y)g(y)dy \right| \\ &= \left| \int_{\mathbb{R}^n} (f(x-y) - f(-y))\overline{f(-y)}dy \right| \\ &\leqslant \Delta_f(x)\|f\|_2, \end{aligned}$$

by the Hölder inequality, where $\Delta_f(x) = \|f(x+\cdot) - f\|_2 \to 0$ as $x \to 0$ in view of the proof of Theorem 2.1.15. Thus, applying Theorem 2.1.22, we have $h(0) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} \widehat{h}(\xi)d\xi$ and $\widehat{h} \in L^1(\mathbb{R}^n)$. It follows that

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

which completes the proof. ■

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

A linear operator on $L^2(\mathbb{R}^n)$ that is an isometry and maps onto $L^2(\mathbb{R}^n)$ is called a ***unitary operator***. It is an immediate consequence of Theorem 2.2.1 that $\mathscr{F}$ is an isometry. Moreover, we have the additional property that $\mathscr{F}$ is onto:

Theorem 2.2.2.

$\mathscr{F}$ is a unitary operator on $L^2(\mathbb{R}^n)$.

Proof. Since $\mathscr{F}$ is an isometry, its range is a closed subspace of $L^2(\mathbb{R}^n)$. If this subspace were not all of $L^2(\mathbb{R}^n)$, we could find a function g such that $\int_{\mathbb{R}^n} \widehat{f}gdx = 0$ for all $f \in L^2$ and $\|g\|_2 \neq 0$. Theorem 2.1.12 obviously extends to L^2; consequently, $\int_{\mathbb{R}^n} f\widehat{g}dx = \int_{\mathbb{R}^n} \widehat{f}gdx = 0$ for all $f \in L^2$. But this implies that $\widehat{g}(x) = 0$ for almost every x, contradicting the fact that $\|\widehat{g}\|_2 = \|g\|_2 \neq 0$ in view of Theorem 2.2.1. ■

Theorem 2.2.2 is a major part of the basic theorem in the L^2 theory of the Fourier transform:

Theorem 2.2.3.

The inverse of the Fourier transform, $\mathscr{F}^{-1}$ or $\check{\cdot}$, can be obtained by letting

$$\check{f}(x) = \hat{f}(-x)$$

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

Having set down the basic facts concerning the action of the Fourier transform on L^1 and L^2, we extend its definition on L^p for $1 < p < 2$. Given a function $f \in L^p(\mathbb{R}^n)$ with $1 < p < 2$, we define $\hat{f} = \hat{f_1} + \hat{f_2}$, where $f_1 \in L^1(\mathbb{R}^n)$, $f_2 \in L^2(\mathbb{R}^n)$, and $f = f_1 + f_2$; we may take, for instance, $f_1 = f\chi_{|f|>1}$ and $f_2 = f\chi_{|f|\leqslant 1}$. The definition of $\hat{f}$ is independent of the choice of f_1 and f_2, since if $f_1 + f_2 = h_1 + h_2$ for $f_1, h_1 \in L^1(\mathbb{R}^n)$ and $f_2, h_2 \in L^2(\mathbb{R}^n)$, we have $f_1 - h_1 = h_2 - f_2 \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Since these functions are equal on $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, their Fourier transforms are also equal, and we obtain $\hat{f_1} - \hat{h_1} = \hat{h_2} - \hat{f_2}$, which yields $\widehat{f_1 + f_2} = \widehat{h_1 + h_2}$. We have the following result concerning the action of the Fourier transform on L^p.

Theorem 2.2.4: Hausdorff-Young inequality

Let $1 \leqslant p \leqslant 2$ and $1/p + 1/p' = 1$. Then the Fourier transform defined as in (2.1.1) satisfies

$$\|\mathscr{F} f\|_{p'} \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n(1/p-1/2)} \|f\|_p. \tag{2.2.1}$$

Proof. It follows from using the Riesz-Thorin interpolation theorem between the L^1-L^∞ result $\|\mathscr{F} f\|_\infty \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|f\|_1$ (cf. Theorem 2.1.4) and Plancherel's theorem $\|\mathscr{F} f\|_2 = \|f\|_2$ (cf. Theorem 2.2.1). ■

Remark 2.2.5. 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 [Bec75] in general.

Remark 2.2.6. The p' can not be replaced by some q in (2.2.1). Namely, if it holds

$$\|\hat{f}\|_q \leqslant C\|f\|_p, \quad \forall f \in L^p(\mathbb{R}^n), \tag{2.2.2}$$

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

$$\|f_\lambda\|_p = \lambda^{-n} \left(\int_{\mathbb{R}^n} |f(x/\lambda)|^p dx\right)^{1/p} = \lambda^{-n} \left(\int_{\mathbb{R}^n} \lambda^n |f(y)|^p dy\right)^{1/p} = \lambda^{-\frac{n}{p'}} \|f\|_p. \tag{2.2.3}$$

By the property of the Fourier transform, we have $\widehat{f_\lambda} = \lambda^{-n}\widehat{\delta^{\lambda^{-1}} f} = \delta^\lambda \hat{f}$ and

$$\|\widehat{f_\lambda}\|_q = \left(\int_{\mathbb{R}^n} |\hat{f}(\lambda\xi)|^q d\xi\right)^{1/q} = \lambda^{-\frac{n}{q}} \|\hat{f}\|_q.$$

Thus, (2.2.2) 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 ∞.

Remark 2.2.7. Except in the case $p = 2$ the inequality (2.2.1) is not reversible, in the sense that there is no constant C such that $\|\widehat{f}\|_{p'} \geqslant C\|f\|_p$ when $f \in \mathscr{D}$. Equivalently (in view of the inversion theorem) the result can not be extended to the case $p > 2$. In order to show it, we take $\omega = 2\pi$ for simplicity, and $f_\lambda(x) = \phi(x)e^{-\pi(1+i\lambda)|x|^2}$, 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 $(1+i\lambda)^{-n/2}e^{-\pi(1+i\lambda)^{-1}|x|^2}$ (cf. [Gra14, Ex.2.3.13, p.133] or [BCD11, Proposition 1.28]), which has L^∞ norm $(1+\lambda^2)^{-n/4}$. Accordingly, if $p \in [1, 2)$ then

$$\|\widehat{f_\lambda}\|_{p'} \leqslant \|\widehat{f_\lambda}\|_2^{\frac{2}{p'}}\|\widehat{f_\lambda}\|_\infty^{1-\frac{2}{p'}} \leqslant C(1+\lambda^2)^{-\frac{n}{2}(\frac{1}{2}-\frac{1}{p'})}.$$

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 $f \in \mathscr{D}$.

As an application of the Marcinkiewicz interpolation theorem, we present a generalization of the Hausdorff-Young inequality due to Paley. The main difference between the theorems being that Paley introduced a weight function into his inequality and resorted to the theorem of Marcinkiewicz. 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

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

With this notation we have the following theorem.

Theorem 2.2.8: Hardy-Littlewood-Paley theorem on $\mathbb{R}^n$

Assume $p \in [1, 2]$. Then

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

Proof. We consider the mapping $(Tf)(\xi) = |\xi|^n\widehat{f}(\xi)$. By Plancherel theorem, we have

$$\|Tf\|_{L^2_*(|\xi|^{-2n})} \leqslant \|Tf\|_{L^2(|\xi|^{-2n})} = \|\widehat{f}\|_2 = \|f\|_2,$$

which implies that T is of weak type $(2, 2)$. We now work towards showing that T is of weak type $(1, 1)$. Thus, the Marcinkiewicz interpolation theorem implies the theorem.

Now, consider the set $E_\alpha = \{\xi : |\xi|^n|\widehat{f}(\xi)| > \alpha\}$. For simplicity, we let ν denote the measure $|\xi|^{-2n}d\xi$ and assume that $\|f\|_1 = 1$. Then, $|\widehat{f}(\xi)| \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2}$. For $\xi \in E_\alpha$, we therefore have $\alpha < \left(\frac{|\omega|}{2\pi}\right)^{n/2}|\xi|^n$. Consequently,

$$(Tf)_*(\alpha) = \nu(E_\alpha) = \int_{E_\alpha} |\xi|^{-2n}d\xi \leqslant \int_{|\xi|^n > \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\alpha} |\xi|^{-2n}d\xi \leqslant C\alpha^{-1}.$$

Thus, we prove that

$$\alpha \cdot (Tf)_*(\alpha) \leqslant C\|f\|_1,$$

which implies T is of weak type $(1,1)$. Therefore, we complete the proof. ■

§2.3 The class of Schwartz functions

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 $\mathbb{R}^n$.

Distributions (generalized functions) $\mathscr{D}'(\mathbb{R}^n)$, which consists of continuous linear functionals on the space $\mathscr{D}(\mathbb{R}^n)$, aroused mostly due to Dirac and his delta function δ_0. The Dirac delta gives a description of a point of unit mass (placed at the origin). The mass density function satisfies that it vanishes if it is integrated on a set not containing the origin, but it is 1 if the set does contain the origin. No function (in the traditional sense) can have this property because we know that the value of a function at a particular point does not change the value of the integral.

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

The multiplication formula (Theorem 2.1.12) might suggest defining the Fourier transform of a distribution based on duality. However, there is a serious impediment in doing so. Specifically, while for every $\varphi \in \mathscr{D}(\mathbb{R}^n)$ we have $\widehat{\varphi} \in \mathcal{C}^\infty(\mathbb{R}^n)$ from the definition of the Fourier transform directly, and $\widehat{\varphi}$ decays at infinity. We nonetheless have

$$\mathscr{F}(\mathscr{D}(\mathbb{R}^n)) \not\subset \mathscr{D}(\mathbb{R}^n). \tag{2.3.1}$$

In fact, we claim that

$$\varphi \in \mathscr{D}(\mathbb{R}^n) \text{ and } \widehat{\varphi} \in \mathscr{D}(\mathbb{R}^n) \Longrightarrow \varphi = 0. \tag{2.3.2}$$

To prove it, suppose $\varphi \in \mathscr{D}(\mathbb{R}^n)$ is such that $\widehat{\varphi}$ has compact support in $\mathbb{R}^n$, and pick an arbitrary point $y = (y_1, \cdots, y_n) \in \mathbb{R}^n$. Define the function $\Phi : \mathbb{C} \to \mathbb{C}$ by

$$\Phi(z) := \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i(zx_1 + \sum_{j=2}^n y_j x_j)} \varphi(x_1, \cdots, x_n) dx_1 \cdots dx_n, \text{ for } z \in \mathbb{C}. \tag{2.3.3}$$

Then, Φ is analytic in $\mathbb{C}$ and $\Phi(t) = \widehat{\varphi}(t, y_2, \cdots, y_n)$ for each $t \in \mathbb{R}$. Given that $\widehat{\varphi}$ has compact support, this implies that Φ vanishes on $\mathbb{R} \setminus [-R, R]$ if $R > 0$ is suitably large. The identity theorem for ordinary analytic functions of one complex variable then forces $\Phi = 0$ everywhere in $\mathbb{C}$. In particular, $\widehat{\varphi}(y) = \Phi(y_1) = 0$. Since $y \in \mathbb{R}^n$ has been chosen arbitrary, we conclude that $\widehat{\varphi} = 0$ in $\mathbb{R}^n$.

To overcome the difficulty highlighted in (2.3.1), we introduce a new (topological vector) space of functions, $\mathscr{S}$, that contains $\mathscr{D}(\mathbb{R}^n)$, is invariant under $\mathscr{F}$, and that has a dual that is a subspace of $\mathscr{D}'(\mathbb{R}^n)$. Then, it would certainly have to consist of functions that are indefinitely differentiable; this, in view of part (v) in Proposition 2.1.2, indicates that each function in $\mathscr{S}$, after being multiplied by a polynomial, must still be in $\mathscr{S}$. We therefore make the following definition:

Definition 2.3.1.

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

$$\mathscr{S}(\mathbb{R}^n) = \Big\{\varphi \in \mathcal{C}^\infty(\mathbb{R}^n) : \rho_{\alpha,\beta}(\varphi) \equiv |\varphi|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta \varphi(x)| < \infty, \ \forall \alpha, \beta \in \mathbb{N}_0^n\Big\}, \quad (2.3.4)$$

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

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

Example 2.3.2. $\varphi(x) = e^{-\varepsilon|x|^2}$, $\varepsilon > 0$, belongs to $\mathscr{S}$; on the other hand, $\varphi(x) = e^{-\varepsilon|x|}$ fails to be differential at the origin and, therefore, does not belong to $\mathscr{S}$.

Example 2.3.3. $\varphi(x) = e^{-\varepsilon(1+|x|^2)^\gamma}$ belongs to $\mathscr{S}$ for any $\varepsilon, \gamma > 0$.

Example 2.3.4. The function $f(x) = \frac{1}{(1+|x|^2)^k} \notin \mathscr{S}$ for any $k \in \mathbb{N}$ since $|x|^{2k} f(x) \not\to 0$ as $|x| \to \infty$.

Example 2.3.5. Sometimes $\mathscr{S}(\mathbb{R}^n)$ is called the space of rapidly decaying functions. But observe that the function $f(x) = e^{-x^2}\sin(e^{x^2}) \notin \mathscr{S}(\mathbb{R})$ since $f'(x) \not\to 0$ as $|x| \to \infty$. Hence, rapid decay of the value of the function alone does not assure the membership in $\mathscr{S}(\mathbb{R})$.

Example 2.3.6. $\mathscr{S}$ contains the space $\mathscr{D}(\mathbb{R}^n)$.

Remark 2.3.7. But it is not immediately clear that $\mathscr{D}$ is nonempty. To find a function in $\mathscr{D}$, consider the function

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

Then, $f \in \mathcal{C}^\infty$, 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 $|t| < 1$, is zero otherwise. It clearly belongs to $\mathscr{D} = \mathscr{D}(\mathbb{R})$. We can easily obtain n-dimensional variants from φ. For examples,

(i) For $x \in \mathbb{R}^n$, define $\psi(x) = \varphi(x_1)\varphi(x_2)\cdots\varphi(x_n)$, then $\psi \in \mathscr{D}(\mathbb{R}^n)$;
(ii) For $x \in \mathbb{R}^n$, define $\psi(x) = e^{-2/(1-|x|^2)}$ for $|x| < 1$ and 0 otherwise, then $\psi \in \mathscr{D}(\mathbb{R}^n)$;
(iii) If $\eta \in \mathcal{C}^\infty$ 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 $\varepsilon \to 0$.

Remark 2.3.8. We observe that the order of multiplication by powers of $x_1, \cdots, x_n$ and differentiation, in (2.3.4), 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 $P(x)$).

Remark 2.3.9. The following alternative characterization of Schwartz functions is very useful. For $\varphi \in \mathscr{C}^\infty$,

$$\varphi \in \mathscr{S}(\mathbb{R}^n) \iff \sup_{x\in\mathbb{R}^n} [(1+|x|)^N |\partial^\alpha \varphi(x)|] < \infty, \quad \forall N \in \mathbb{N}_0, \forall \alpha \in \mathbb{N}_0^n, |\alpha| \leqslant N. \tag{2.3.5}$$

Indeed, it follows from the observation that for each $N \in \mathbb{N}_0$ there exists $C \in [1, \infty)$ such that

$$C^{-1}|x|^N \leqslant \sum_{|\gamma|=N} |x^\gamma| \leqslant C|x|^N, \quad \forall x \in \mathbb{R}^n. \tag{2.3.6}$$

The second inequality in (2.3.6) is seen by noting that $\forall \alpha \in \mathbb{N}_0^n$ and $\forall x \in \mathbb{R}^n$, we have

$$|x^\alpha| = |x_1|^{\alpha_1} \cdots |x_n|^{\alpha_n} \leqslant |x|^{\alpha_1} \cdots |x|^{\alpha_n} = |x|^{|\alpha|}. \tag{2.3.7}$$

To justify the first inequality in (2.3.6), we consider the function $g(x) := \sum_{|\gamma|=N} |x^\gamma|$ for $x \in \mathbb{R}^n$. Then its restriction to $\mathbb{S}^{n-1}$ attains a positive minimum since it has no zeros on $\mathbb{S}^{n-1}$, and the desired inequality follows by scaling.

For the three spaces, we have the following relations:

$$\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset \mathscr{C}^\infty(\mathbb{R}^n).$$

Definition 2.3.10.

We define convergence of sequences in these spaces. We say that

$$f_k \to f \text{ in } \mathscr{C}^\infty \iff f_k, f \in \mathscr{C}^\infty \text{ and } \lim_{k\to\infty} \sup_{|x|\leqslant N} |\partial^\alpha (f_k - f)(x)| = 0, \ \forall \alpha \in \mathbb{N}_0^n, \forall N \in \mathbb{N}.$$

$$f_k \to f \text{ in } \mathscr{S} \iff f_k, f \in \mathscr{S} \text{ and } \lim_{k\to\infty} \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta (f_k - f)(x)| = 0, \ \forall \alpha, \beta \in \mathbb{N}_0^n.$$

$$f_k \to f \text{ in } \mathscr{D} \iff f_k, f \in \mathscr{D}, \exists B \text{ compact, s.t. } \operatorname{supp} f_k \subset B \text{ for all } k, \text{ and } \lim_{k\to\infty} \|\partial^\alpha (f_k - f)\|_\infty = 0, \ \forall \alpha \in \mathbb{N}_0^n.$$

It follows that convergence in $\mathscr{D}(\mathbb{R}^n)$ implies convergence in $\mathscr{S}(\mathbb{R}^n)$, which in turn implies convergence in $\mathscr{C}^\infty(\mathbb{R}^n)$.

The space $\mathscr{C}^\infty(\mathbb{R}^n)$ is equipped with the family of seminorms

$$\rho_{\alpha,N}(f) = \sup_{|x|\leqslant N} |(\partial^\alpha f)(x)|, \tag{2.3.8}$$

where α ranges over all multi-indices in $\mathbb{N}_0^n$ and N ranges over $\mathbb{N}_0$. It can be shown that $\mathscr{C}^\infty(\mathbb{R}^n)$ is complete with respect to this countable family of seminorms, i.e., it is a Frèchet space. However, it is true that $\mathscr{D}(\mathbb{R}^n)$ is not complete with respect to the topology generated by this family of seminorms.

The topology of $\mathscr{D}$ given in Definition 2.3.10 is the inductive limit topology, and under this topology it is complete. Indeed, letting $\mathscr{D}(B(0,k))$ be the space of all smooth functions with support in $B(0,k)$, then $\mathscr{D}(\mathbb{R}^n)$ is equal to $\bigcup_{k=1}^\infty \mathscr{D}(B(0,k))$ and each space $\mathscr{D}(B(0,k))$ is complete with respect to the topology generated by the family of seminorms $\rho_{\alpha,N}$, hence so is $\mathscr{D}(\mathbb{R}^n)$. Nevertheless, $\mathscr{D}(\mathbb{R}^n)$ is not metrizable by the Baire category theorem.

Theorem 2.3.11.

$\mathscr{S}$ is contained in and dense in $\mathcal{C}_0(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$.

Proof. $\mathscr{S} \subset \mathcal{C}_0$ is obvious by (2.3.4). The L^p norm of $\varphi \in \mathscr{S}$ is bounded by a finite linear combination of L^∞ norms of terms of the form $x^\alpha \varphi(x)$. In fact, by (2.3.4), we have

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

For the proof of the density, it follows from the fact that $\mathscr{D}$ is dense in those spaces[1] and $\mathscr{D} \subset \mathscr{S}$. ■

Remark 2.3.12. 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 $\lim\limits_{|x|\to\infty} |\varphi(x)| = 0$ for all $\varphi \in \mathscr{S}$.

From part (v) in Proposition 2.1.2, we immediately have

Theorem 2.3.13.

If $\varphi \in \mathscr{S}$, then $\widehat{\varphi} \in \mathscr{S}$.

---

[1]For convenience, we review the proof of the fact that $\mathscr{D}$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$, and leave the case of $\mathcal{C}_0$ to the interested reader.

We will use the fact that the set of finite linear combinations of characteristic functions of bounded measurable sets in $\mathbb{R}^n$ is dense in $L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$. This is a well-known fact from functional analysis.

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

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

or

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

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

We can define the inverse Fourier transform for Schwartz functions as for L^2 functions. Given $f \in \mathscr{S}(\mathbb{R}^n)$, we define

$$\check{f}(x) = \widehat{f}(-x),$$

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

$$\mathscr{F}^{-1} : f \mapsto \check{f}$$

is called the inverse Fourier transform.

It is straightforward that the inverse Fourier transform shares the same properties as the Fourier transform. We now give the relation between the Fourier transform and the inverse Fourier transform and leave proofs to the reader.

Theorem 2.3.14.

Given f, g, and h in $\mathscr{S}(\mathbb{R}^n)$, we have

(1) Multiplication formula: $\displaystyle\int_{\mathbb{R}^n} \widehat{f}(x)g(x)dx = \int_{\mathbb{R}^n} f(x)\widehat{g}(x)dx$;

(2) Fourier inversion: $\check{\widehat{f}} = f = \widehat{\check{f}}$;

(3) Parseval's relation: $\displaystyle\int_{\mathbb{R}^n} f(x)\bar{h}(x)dx = \int_{\mathbb{R}^n} \widehat{f}(\xi)\overline{\widehat{h}(\xi)}d\xi$;

(4) Plancherel's identity: $\|f\|_2 = \|\widehat{f}\|_2 = \|\check{f}\|_2$;

(5) $\displaystyle\int_{\mathbb{R}^n} f(x)h(x)dx = \int_{\mathbb{R}^n} \widehat{f}(x)\check{h}(x)dx$.

If $\varphi, \psi \in \mathscr{S}$, then Theorem 2.3.13 implies that $\widehat{\varphi}, \widehat{\psi} \in \mathscr{S}$. Therefore, $\widehat{\varphi}\widehat{\psi} \in \mathscr{S}$. By part (vi) in Proposition 2.1.2, i.e., $\mathscr{F}(\varphi * \psi) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{\varphi}\widehat{\psi}$, an application of the inverse Fourier transform shows that

Theorem 2.3.15.

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

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

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

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

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

Theorem 2.3.16: Completeness

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

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

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

and then

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

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

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

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

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

Taking $k \to \infty$, we obtain

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

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

Moreover, some easily established properties of $\mathscr{S}(\mathbb{R}^n)$ and its topology, are the following:

Proposition 2.3.17.

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 $|h| \to 0$.
iv) The Fourier transform is a homeomorphism of $\mathscr{S}(\mathbb{R}^n)$ onto itself.
v) $\mathscr{S}(\mathbb{R}^n)$ is separable.

Finally, we describe and prove a fundamental result of Fourier analysis that is known as the uncertainty principle discovered by W. Heisenberg in 1927. It says that the position and the momentum of an object cannot both be measured exactly, at the same time, even in theory. In the context of harmonic analysis, the uncertainty principle implies that one cannot at the same time localize the value of a function and its Fourier transform. The exact statement is as follows.

Theorem 2.3.18: The Heisenberg uncertainty principle

Suppose ψ is a function in $\mathscr{S}(\mathbb{R})$. Then

$$\|x\psi\|_2 \|\xi\widehat{\psi}\|_2 \geqslant \frac{\|\psi\|_2^2}{2|\omega|},$$

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

$$\|(x - x_0)\psi\|_2 \|(\xi - \xi_0)\widehat{\psi}\|_2 \geqslant \frac{\|\psi\|_2^2}{2|\omega|}$$

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

Proof. The last inequality actually follows from the first one by replacing $\psi(x)$ by $e^{-\omega i x \xi_0}\psi(x+x_0)$ (whose Fourier transform is $e^{\omega i x_0(\xi+\xi_0)}\widehat{\psi}(\xi+\xi_0)$ by parts (ii) and (iii) in Proposition 2.1.2) and changing variables. To prove the first inequality, we argue as follows.

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

$$\begin{aligned}\|\psi\|_2^2 = &\int_{-\infty}^{\infty} |\psi(x)|^2 dx = -\int_{-\infty}^{\infty} x \frac{d}{dx}|\psi(x)|^2 dx \\ &= -\int_{-\infty}^{\infty} \left(x\psi'(x)\overline{\psi(x)} + x\overline{\psi'(x)}\psi(x) \right) dx.\end{aligned}$$

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

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

where we have used the Cauchy-Schwarz inequality. By part (v) in Proposition 2.1.2, we have $\mathscr{F}(\psi')(\xi) = \omega i \xi \widehat{\psi}(\xi)$. It follows, from the Plancherel theorem, that

$$\|\psi'\|_2 = \|\mathscr{F}(\psi')\|_2 = |\omega|\|\xi\widehat{\psi}\|_2.$$

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

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

§ 2.4 The class of tempered distributions

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

Definition 2.4.1.

The functional $T : \mathscr{S} \to \mathbb{C}$ is a ***tempered distribution*** if
(i) T is linear, i.e., $\langle T, \alpha\varphi + \beta\psi\rangle = \alpha\langle T, \varphi\rangle + \beta\langle T, \psi\rangle$ for all $\alpha, \beta \in \mathbb{C}$ and $\varphi, \psi \in \mathscr{S}$.
(ii) T is continuous on $\mathscr{S}$, i.e., there exist $n_0 \in \mathbb{N}_0$ and a constant $c_0 > 0$ such that

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

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

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

Remark 2.4.2. Since $\mathscr{D} \subset \mathscr{S}$, the space of tempered distributions $\mathscr{S}'$ is more narrow than the space of distributions $\mathscr{D}'$, i.e., $\mathscr{S}' \subset \mathscr{D}'$. Another more narrow distribution space $\mathscr{E}'$ which consists of continuous linear functionals on the (widest test function) space $\mathcal{C}^\infty(\mathbb{R}^n)$. In short, $\mathscr{D} \subset \mathscr{S} \subset \mathcal{C}^\infty$ implies that

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

By definition of the topologies on the dual spaces, we have

$$T_k \to T \text{ in } \mathscr{D}' \iff T_k, T \in \mathscr{D}' \text{ and } T_k(f) \to T(f) \text{ for all } f \in \mathscr{D}.$$
$$T_k \to T \text{ in } \mathscr{S}' \iff T_k, T \in \mathscr{S}' \text{ and } T_k(f) \to T(f) \text{ for all } f \in \mathscr{S}.$$
$$T_k \to T \text{ in } \mathscr{E}' \iff T_k, T \in \mathscr{E}' \text{ and } T_k(f) \to T(f) \text{ for all } f \in \mathcal{C}^\infty.$$

Definition 2.4.3.

Elements of the space $\mathscr{D}'(\mathbb{R}^n)$ are called ***distributions***. Elements of $\mathscr{S}'(\mathbb{R}^n)$ are called ***tempered distributions***. Elements of the space $\mathscr{E}'(\mathbb{R}^n)$ are called ***distributions with compact support***.

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:

$$\langle u, f\rangle = u(f).$$

There exists a simple and important characterization of distributions:

Theorem 2.4.4.

(1) A linear functional u on $\mathscr{D}(\mathbb{R}^n)$ is a distribution if and only if for every compact $K \subset \mathbb{R}^n$, there exist $C > 0$ and an integer m such that

$$|\langle u, f\rangle| \leqslant C \sum_{|\alpha|\leqslant m} \|\partial^\alpha f\|_\infty, \quad \forall f \in \mathcal{C}^\infty \text{ with support in } K. \tag{2.4.1}$$

(2) A linear functional u on $\mathscr{S}$ is a tempered distribution if and only if there exists a constant $C > 0$ and integers ℓ and m such that

$$|\langle u, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant \ell, |\beta|\leqslant m} \rho_{\alpha,\beta}(\varphi), \quad \forall \varphi \in \mathscr{S}. \tag{2.4.2}$$

(3) A linear functional u on $\mathcal{C}^\infty(\mathbb{R}^n)$ is a distribution with compact support if and only if there exist $C > 0$ and integers N, m such that

$$|\langle u, f\rangle| \leqslant C \sum_{|\alpha|\leqslant m} \rho_{\alpha,N}(f), \quad \forall f \in \mathcal{C}^\infty(\mathbb{R}^n). \tag{2.4.3}$$

The seminorms $\rho_{\alpha,\beta}$ and $\rho_{\alpha,N}$ are defined in (2.3.4) and (2.3.8), respectively.

Proof. We prove only (2), since the proofs of (1) and (3) 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 $k \to \infty$ if and only if $|\varphi_k - \varphi|_{\alpha,\beta} \to 0$ for all (α, β) in the topology induced by this system of neighborhoods and their translates. Thus, there exists such a set $N_{\varepsilon,\ell,m}$ satisfying $|\langle u, \varphi\rangle| \leqslant 1$ 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

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

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

Example 2.4.5. Let $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, and define $T = T_f$ by letting

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

for $\varphi \in \mathscr{S}$. It is clear that T_f is a linear functional on $\mathscr{S}$. To show that it is continuous, therefore, it suffices to show that it is continuous at the origin. Then, suppose $\varphi_k \to 0$ in $\mathscr{S}$ as $k \to \infty$. From the proof of Theorem 2.3.11, we have seen that for any $q \geqslant 1$, $\|\varphi_k\|_q$ is dominated by a finite linear combination of L^∞ norms of terms of the form $x^\alpha \varphi_k(x)$. That is, $\|\varphi_k\|_q$ is dominated by a finite linear combination of semi-norms $|\varphi_k|_{\alpha,0}$. Thus, $\|\varphi_k\|_q \to 0$ as $k \to \infty$. Choosing $q = p'$, i.e., $1/p+1/q = 1$, Hölder's inequality shows that $|\langle T, \varphi_k\rangle| \leqslant \|f\|_p \|\varphi_k\|_{p'} \to 0$ as $k \to \infty$. Thus, $T \in \mathscr{S}'$.

Example 2.4.6. We consider the case $n = 1$. Let $f(x) = \sum_{k=0}^m a_k x^k$ be a polynomial, then $f \in \mathscr{S}'$ since

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

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

Example 2.4.7. 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 2.4.5-2.4.7 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 as embedded in $\mathscr{S}'$. If we put on $\mathscr{S}'$ the weakest topology such that the linear functionals $T \to \langle T, \varphi\rangle$ ($\varphi \in \mathscr{S}$) are continuous, it is easy to see that the spaces $L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, are continuously embedded in $\mathscr{S}'$. The same is true for the space of all finite Borel measures on $\mathbb{R}^n$, i.e., $\mathscr{B}(\mathbb{R}^n)$.

Suppose that f and g are Schwartz functions and α a multi-index. Integrating by parts $|\alpha|$ times, we obtain

$$\int_{\mathbb{R}^n} (\partial^\alpha f)(x) g(x) dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x)(\partial^\alpha g)(x) dx. \tag{2.4.4}$$

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 (2.4.4) for $g \in \mathscr{S}'$ and $f \in \mathscr{S}$ if the integrals in (2.4.4) are interpreted as actions of distributions on functions. We simply use (2.4.4) to define the derivative of a distribution.

Definition 2.4.8.

Let $u \in \mathscr{S}'$ and α a multi-index. Define

$$\langle \partial^\alpha u, f\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha f\rangle. \tag{2.4.5}$$

If u is a function, the derivatives of u in the sense of distributions are called ***distributional derivatives***.

In view of Theorem 2.3.14, it is natural to give the following:

Definition 2.4.9.

Let $u \in \mathscr{S}'$. We define the Fourier transform $\widehat{u}$ and the inverse Fourier transform $\check{u}$ of a tempered distribution u by

$$\langle \widehat{u}, f\rangle = \langle u, \widehat{f}\rangle \quad \text{and } \langle \check{u}, f\rangle = \langle u, \check{f}\rangle, \tag{2.4.6}$$

for all f in $\mathscr{S}$, respectively.

Example 2.4.10. For $\varphi \in \mathscr{S}$, we have

$$\langle \widehat{\delta_0}, \varphi\rangle = \langle \delta_0, \mathscr{F}\varphi\rangle = \widehat{\varphi}(0) = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{-\omega i x\cdot 0}\varphi(x)dx = \langle \left(\frac{|\omega|}{2\pi}\right)^{n/2}, \varphi\rangle.$$

Thus, $\widehat{\delta_0} = \left(\frac{|\omega|}{2\pi}\right)^{n/2}$ in $\mathscr{S}'$. More generally, since

$$\begin{aligned}\langle \widehat{\partial^\alpha\delta_0}, \varphi\rangle =&\langle \partial^\alpha\delta_0, \widehat{\varphi}\rangle = (-1)^{|\alpha|}\langle \delta_0, \partial^\alpha\widehat{\varphi}\rangle = \langle \delta_0, \mathscr{F}[(\omega i\xi)^\alpha\varphi]\rangle\\ =&\langle \widehat{\delta_0}, (\omega i\xi)^\alpha\varphi\rangle = \langle \left(\frac{|\omega|}{2\pi}\right)^{n/2}(\omega i\xi)^\alpha, \varphi\rangle,\end{aligned}$$

we have $\widehat{\partial^\alpha\delta_0} = \left(\frac{|\omega|}{2\pi}\right)^{n/2}(\omega i\xi)^\alpha$. This calculation indicates that $\widehat{\partial^\alpha\delta_0}$ can be identified with the function $\left(\frac{|\omega|}{2\pi}\right)^{n/2}(\omega i\xi)^\alpha$.

Example 2.4.11. Since for any $\varphi \in \mathscr{S}$,

$$\begin{aligned}\langle \widehat{1}, \varphi\rangle =&\langle 1, \widehat{\varphi}\rangle = \int_{\mathbb{R}^n}\widehat{\varphi}(\xi)d\xi = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{\omega i 0\cdot\xi}\widehat{\varphi}(\xi)d\xi\\ =&\left(\frac{|\omega|}{2\pi}\right)^{-n/2}\check{\widehat{\varphi}}(0) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\varphi(0) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\langle \delta_0, \varphi\rangle,\end{aligned}$$

we have

$$\widehat{1} = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\delta_0, \quad \text{in } \mathscr{S}'.$$

Moreover, $\check{\delta_0} = \left(\frac{|\omega|}{2\pi}\right)^{n/2}$.

Now observe that the following are true whenever f, g are in $\mathscr{S}$.

$$\begin{aligned}\int_{\mathbb{R}^n} g(x)f(x-t)dx &= \int_{\mathbb{R}^n} g(x+t)f(x)dx,\\ \int_{\mathbb{R}^n} g(ax)f(x)dx &= \int_{\mathbb{R}^n} g(x)a^{-n}f(a^{-1}x)dx, \\ \int_{\mathbb{R}^n} \tilde{g}(x)f(x)dx &= \int_{\mathbb{R}^n} g(x)\tilde{f}(x)dx,\end{aligned} \tag{2.4.7}$$

for all $t \in \mathbb{R}^n$ and $a > 0$, where $\tilde{}$ denotes the reflection. Motivated by (2.4.7), we give the following:

Definition 2.4.12.

The translation $\tau^t u$, the dilation $\delta^a u$, and the reflection $\tilde{u}$ of a tempered distribution u are defined as follows:

$$\begin{aligned}\langle \tau^t u, f\rangle =&\langle u, \tau^{-t} f\rangle,\\ \langle \delta^a u, f\rangle =&\langle u, a^{-n}\delta^{1/a} f\rangle,\\ \langle \tilde{u}, f\rangle =&\langle u, \tilde{f}\rangle,\end{aligned}$$

for all $t \in \mathbb{R}^n$ and $a > 0$. Let A be an invertible matrix. 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 2.4.13. The Dirac mass at the origin δ_0 is equal to its reflection, while $\delta^a\delta_0 = a^{-n}\delta_0$ for $a > 0$. Also, $\tau^x\delta_0 = \delta_x$ for any $x \in \mathbb{R}^n$.

Now observe that for f, g and h in $\mathscr{S}$, we have

$$\int_{\mathbb{R}^n} (h * g)(x)f(x)dx = \int_{\mathbb{R}^n} g(x)(\tilde{h} * f)(x)dx. \tag{2.4.8}$$

Motivated by this identity, we define the convolution of a function with a tempered distribution as follows:

Definition 2.4.14.

Let $u \in \mathscr{S}'$ and $h \in \mathscr{S}$. Define the convolution $h * u$ by

$$\langle h * u, f\rangle = \langle u, \tilde{h} * f\rangle, \qquad f \in \mathscr{S}. \tag{2.4.9}$$

Example 2.4.15. 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

$$\langle f * \delta_{x_0}, h\rangle = \langle \delta_{x_0}, \tilde{f} * h\rangle = (\tilde{f} * h)(x_0) = \int_{\mathbb{R}^n} f(x-x_0)h(x)dx = \langle \tau^{x_0} f, h\rangle.$$

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 2.4.16.

Let $u \in \mathscr{S}'$ and let h be a $\mathcal{C}^\infty$ function that has at most polynomial growth at infinity and the same is true for all of its derivatives. This means that for all α it satisfies $|(\partial^\alpha h)(x)| \leqslant C_\alpha(1+|x|)^{k_\alpha}$ for some $C_\alpha, k_\alpha > 0$. Then define the product hu of h and u by

$$\langle hu, f\rangle = \langle u, hf\rangle, \qquad f \in \mathscr{S}. \tag{2.4.10}$$

Note that $hf \in \mathscr{S}$ and thus (2.4.10) is well defined. The product of an arbitrary $\mathcal{C}^\infty$ function with a tempered distribution is not defined.

Example 2.4.17. Let $T \in \mathscr{S}'$ and $\varphi \in \mathscr{D}(\mathbb{R}^n)$ with $\varphi(0) = 1$. Then the product $\varphi(x/k)T$ is well-defined in $\mathscr{S}'$ by

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

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

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

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

Next, we give a proposition that extends the properties of the Fourier transform to tempered distributions.

Proposition 2.4.18.

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) $\widehat{u+v} = \hat{u} + \hat{v}$,
(ii) $\widehat{bu} = b\hat{u}$,
(iii) $\widehat{\tilde{u}} = \tilde{\hat{u}}$,
(iv) $\widehat{\tau^y u}(\xi) = e^{-\omega i y\cdot\xi}\hat{u}(\xi)$,
(v) $\widehat{e^{\omega i x\cdot y}u(x)} = \tau^y\hat{u}$,
(vi) $\widehat{\delta^a u} = (\hat{u})_a = a^{-n}\delta^{a^{-1}}\hat{u}$,
(vii) $\widehat{\partial^\alpha u}(\xi) = (\omega i\xi)^\alpha\hat{u}(\xi)$,
(viii) $\partial^\alpha\hat{u} = \widehat{(-\omega i x)^\alpha u(x)}$,
(ix) $\check{\hat{u}} = u$,
(x) $\widehat{f*u} = \left(\frac{|\omega|}{2\pi}\right)^{-n/2}\hat{f}\hat{u}$,
(xi) $\widehat{fu} = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\hat{f}*\hat{u}$.

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 2.4.19.

If $u \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$, then $\varphi * u$ is a $\mathcal{C}^\infty$ function and

$$(\varphi * u)(x) = \langle u, \tau^x \tilde{\varphi} \rangle, \tag{2.4.11}$$

for all $x \in \mathbb{R}^n$. Moreover, for all multi-indices α there exist constants $C_\alpha, k_\alpha > 0$ such that

$$|\partial^\alpha (\varphi * u)(x)| \leqslant C_\alpha (1 + |x|)^{k_\alpha}. \tag{2.4.12}$$

Proof. We first prove (2.4.11). Let $\psi \in \mathscr{S}(\mathbb{R}^n)$. We have

$$\begin{aligned} \langle \varphi * u, \psi \rangle =& \langle u, \tilde{\varphi} * \psi \rangle \\ =& u\left(\int_{\mathbb{R}^n} \tilde{\varphi}(\cdot - y) \psi(y) dy \right) \\ =& u\left(\int_{\mathbb{R}^n} (\tau^y \tilde{\varphi})(\cdot) \psi(y) dy \right) \\ =& \int_{\mathbb{R}^n} \langle u, \tau^y \tilde{\varphi} \rangle \psi(y) dy, \end{aligned} \tag{2.4.13}$$

where the last step is justified by the continuity of u and by the fact that the Riemann sums of the inner integral in (2.4.13) converge uniformly to that integral in the topology of $\mathscr{S}$, a fact that will be justified later. This calculation implies (2.4.11).

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 2.3.17,

$$\frac{\tau^{-he_j} \tau^x \tilde{\varphi} - \tau^x \tilde{\varphi}}{h} \to \partial_j \tau^x \tilde{\varphi} = \tau^x \partial_j \tilde{\varphi},$$

in $\mathscr{S}$ as $h \to 0$. Thus, since u is linear and continuous, we have

$$\frac{\tau^{-he_j}(\varphi * u)(x) - (\varphi * u)(x)}{h} = u\left(\frac{\tau^{-he_j}(\tau^x \tilde{\varphi}) - \tau^x \tilde{\varphi}}{h} \right) \to \langle u, \tau^x (\partial_j \tilde{\varphi}) \rangle$$

as $h \to 0$. The same calculation for higher-order derivatives show 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 2.4.4 that for some C, m and k we have

$$\begin{aligned} |\partial^\alpha (\varphi * u)(x)| \leqslant& C \sum_{\substack{|\gamma| \leqslant m \\ |\beta| \leqslant k}} \sup_{y \in \mathbb{R}^n} |y^\gamma \tau^x (\partial^{\alpha+\beta} \tilde{\varphi})(y)| \\ =& C \sum_{\substack{|\gamma| \leqslant m \\ |\beta| \leqslant k}} \sup_{y \in \mathbb{R}^n} |(x+y)^\gamma (\partial^{\alpha+\beta} \tilde{\varphi})(y)| \quad \text{(change variables: } y - x \to y) \\ \leqslant& C_m \sum_{|\beta| \leqslant k} \sup_{y \in \mathbb{R}^n} |(1 + |x|^m + |y|^m)(\partial^{\alpha+\beta} \tilde{\varphi})(y)| \\ \leqslant& C_{m,k,\alpha} \sup_{y \in \mathbb{R}^n} \frac{1 + |x|^m + |y|^m}{(1+|y|)^N} \quad \text{(taking } N > m) \\ \leqslant& C_{m,k,\alpha} (1 + |x|^m), \end{aligned}$$

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 (2.4.13) 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, β, we must show that

$$D_N(x) = \sum_{m=1}^{(2N^2)^n} x^\alpha \partial_x^\beta \tilde{\varphi}(x-y_m)\psi(y_m)|Q_m| - \int_{\mathbb{R}^n} x^\alpha \partial_x^\beta \tilde{\varphi}(x-y)\psi(y)dy$$

converges to zero in $L^\infty(\mathbb{R}^n)$ as $N \to \infty$. We have by the mean value theorem

$$\begin{aligned}
&x^\alpha \partial_x^\beta \tilde{\varphi}(x-y_m)\psi(y_m)|Q_m| - \int_{Q_m} x^\alpha \partial_x^\beta \tilde{\varphi}(x-y)\psi(y)dy \\
=&\int_{Q_m} x^\alpha [\partial_x^\beta \tilde{\varphi}(x-y_m)\psi(y_m) - \partial_x^\beta \tilde{\varphi}(x-y)\psi(y)]dy \\
=&\int_{Q_m} x^\alpha (y-y_m)\cdot(\nabla(\partial_x^\beta \tilde{\varphi}(x-\cdot)\psi))(\xi)dy \\
=&\int_{Q_m} x^\alpha (y-y_m)\cdot(-\nabla\partial_x^\beta \tilde{\varphi}(x-\cdot)\psi + \nabla\psi\partial_x^\beta \tilde{\varphi}(x-\cdot))(\xi)dy
\end{aligned}$$

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

$$\begin{aligned}
&|x^\alpha(y-y_m)\cdot(-\nabla\partial_x^\beta \tilde{\varphi}(x-\xi)\psi(\xi) + \nabla\psi(\xi)\partial_x^\beta \tilde{\varphi}(x-\xi))| \\
\leqslant& C|x|^{|\alpha|}\frac{\sqrt{n}}{2N}\frac{1}{(2+|\xi|)^M}\frac{1}{(1+|x-\xi|)^{M/2}} \quad \text{(for } M \text{ large)} \\
\leqslant& C|x|^{|\alpha|}\frac{\sqrt{n}}{N}\frac{1}{(2+|\xi|)^{M/2}}\frac{1}{(1+|x|)^{M/2}} \\
\leqslant& C|x|^{|\alpha|}\frac{\sqrt{n}}{N}\frac{1}{(1+|y|)^{M/2}}\frac{1}{(1+|x|)^{M/2}},
\end{aligned}$$

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

$$|D_N(x)| \leqslant \frac{C}{N}\frac{|x|^{|\alpha|}}{(1+|x|)^{M/2}}\int_{[-N,N]^n}\frac{dy}{(1+|y|)^{M/2}} + \int_{([-N,N]^n)^c}|x^\alpha \partial_x^\beta \tilde{\varphi}(x-y)\psi(y)|dy.$$

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 $M > \max(2n, 2|\alpha|)$ since $(1+|x-y|)(1+|y|) \geqslant 1+|x-y|+|y| \geqslant 1+|x|$. From these estimates, it follows that

$$\sup_{x\in\mathbb{R}^n}|D_N(x)| \leqslant C\left(\frac{1}{N} + \frac{1}{N^{\frac{M}{2}-n}}\right) \to 0, \quad \text{as } N \to \infty.$$

Therefore, $\lim\limits_{N\to\infty}\sup\limits_{x\in\mathbb{R}^n}|D_N(x)| = 0$. ■

We observe that if a function g is supported in a set K, then for all $f \in \mathscr{D}(K^c)$ we have

$$\int_{\mathbb{R}^n} f(x)g(x)dx = 0. \tag{2.4.14}$$

Moreover, the support of g is the intersection of all closed sets K with the property (2.4.14) for all f in $\mathscr{D}(K^c)$. Motivated by this observation we give the following:

Definition 2.4.20.

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

$$\varphi \in \mathscr{D}(\mathbb{R}^n), \quad \operatorname{supp} \varphi \subset \mathbb{R}^n \setminus K \Longrightarrow \langle u, \varphi \rangle = 0. \tag{2.4.15}$$

Example 2.4.21. $\operatorname{supp} \delta_{x_0} = \{x_0\}$.

Along the same lines, we give the following definition:

Definition 2.4.22.

We say that a distribution $u \in \mathscr{D}'(\mathbb{R}^n)$ coincides with the function h on an open set Ω if

$$\langle u, f \rangle = \int_{\mathbb{R}^n} f(x)h(x)dx \quad \forall f \in \mathscr{D}(\Omega). \tag{2.4.16}$$

When (2.4.16) occurs we often say that u agrees with h away from Ω^c.

This definition implies $\operatorname{supp}(u - h) \subset \Omega^c$.

Example 2.4.23. 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 2.4.24.

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 (2.4.2), we have that 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

$$(\partial^\alpha \varphi)(0) = 0 \quad \forall |\alpha| \leqslant k, \tag{2.4.17}$$

then $\langle u, \varphi \rangle = 0$. To see this, fix a φ satisfying (2.4.17) 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 (2.4.17) and the continuity of the derivatives of φ at the origin, it is not hard to show that $\rho_{\alpha,\beta}(\zeta^\varepsilon\varphi-\varphi)\to 0$ as $\varepsilon\to 0$ for all $|\alpha|\leqslant m$ and $|\beta|\leqslant k$. Then

$$\begin{aligned}|\langle u,\varphi\rangle| \leqslant& |\langle u,\zeta^\varepsilon\varphi\rangle| + |\langle u,\zeta^\varepsilon\varphi-\varphi\rangle| \\ \leqslant& 0+\sum_{\substack{|\alpha|\leqslant m\\ |\beta|\leqslant k}}\rho_{\alpha,\beta}(\zeta^\varepsilon\varphi-\varphi)\to 0\end{aligned}$$

as $\varepsilon\to 0$. This proves our assertion.

Now, let $f\in\mathscr{S}(\mathbb{R}^n)$. Let η be a $\mathscr{D}(\mathbb{R}^n)$ function on $\mathbb{R}^n$ that is equal to 1 in a neighborhood of the origin. Write

$$f(x)=\eta(x)\left(\sum_{|\alpha|\leqslant k}\frac{(\partial^\alpha f)(0)}{\alpha!}x^\alpha+h(x)\right)+(1-\eta(x))f(x), \tag{2.4.18}$$

where $h(x)=O(|x|^{k+1})$ as $|x|\to 0$. Then ηh satisfies (2.4.17) and hence $\langle u,\eta h\rangle=0$ by the claim. Also,

$$\langle u,((1-\eta)f)\rangle=0$$

by our hypothesis. Applying u to both sides of (2.4.18), we obtain

$$\langle u,f\rangle=\sum_{|\alpha|\leqslant k}\frac{(\partial^\alpha f)(0)}{\alpha!}\langle u,x^\alpha\eta(x)\rangle=\sum_{|\alpha|\leqslant k}a_\alpha\partial^\alpha(\delta_0)(f),$$

with $a_\alpha=(-1)^{|\alpha|}\langle u,x^\alpha\eta(x)\rangle/\alpha!$. This proves the results. ■

An immediate consequence is the following result.

Corollary 2.4.25.

Let $u\in\mathscr{S}'(\mathbb{R}^n)$. If $\hat{u}$ is supported in the singleton $\{\xi_0\}$, then u is a finite linear combination of functions $(-\omega i\xi)^\alpha e^{\omega i\xi\cdot\xi_0}$, where $\alpha\in\mathbb{N}_0^n$. In particular, if $\hat{u}$ is supported at the origin, then u is a polynomial.

Proof. Proposition 2.4.24 gives that $\hat{u}$ is a linear combination of derivatives of Dirac masses at ξ_0, i.e.,

$$\hat{u}=\sum_{|\alpha|\leqslant k}a_\alpha\partial^\alpha\delta_{\xi_0}.$$

Then, Proposition 2.4.18 yields

$$\begin{aligned}u=&\sum_{|\alpha|\leqslant k}a_\alpha\widetilde{\partial^\alpha\delta_{\xi_0}}=\sum_{|\alpha|\leqslant k}a_\alpha\widetilde{\widehat{\partial^\alpha\delta_{\xi_0}}}\\ =&\sum_{|\alpha|\leqslant k}a_\alpha\widetilde{(\omega i\xi)^\alpha\widehat{\delta_{\xi_0}}}=\left(\frac{|\omega|}{2\pi}\right)^{n/2}\sum_{|\alpha|\leqslant k}a_\alpha\widetilde{(\omega i\xi)^\alpha e^{-\omega i\xi\cdot\xi_0}}\\ =&\left(\frac{|\omega|}{2\pi}\right)^{n/2}\sum_{|\alpha|\leqslant k}a_\alpha(-\omega i\xi)^\alpha e^{\omega i\xi\cdot\xi_0}.\end{aligned}$$

■

Proposition 2.4.26.

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 as defined in Definition 2.4.3. Then there exist $C, N, m > 0$ such that (2.4.3) holds. For a $\mathcal{C}^\infty$ function f whose support is contained in $B(0,N)^c$, the expression on the right in (2.4.3) 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 2.4.20 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 (2.4.1), 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 (2.4.3) for $f \in \mathcal{C}^\infty$. ■

For distributions with compact support, we have the following important result.

Theorem 2.4.27.

If $u \in \mathscr{E}'(\mathbb{R}^n)$, then $\hat{u}$ is a real analytic function on $\mathbb{R}^n$. In particular, $\hat{u} \in \mathcal{C}^\infty$. Furthermore, $\hat{u}$ and all of its derivatives have polynomial growth at infinity. Moreover, $\hat{u}$ has a holomorphic extension on $\mathbb{C}^n$.

Proof. Since $u \in \mathscr{E}' \subset \mathscr{S}'$, we have for $f \in \mathscr{S}$

$$\begin{aligned}\langle \hat{u}, f\rangle =&\langle u, \hat{f}\rangle = \left(\frac{|\omega|}{2\pi}\right)^{n/2} u\left(\int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} f(x)dx\right)\\ =&\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} u\left(e^{-\omega i x\cdot(\cdot)}\right) f(x)dx,\end{aligned}$$

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 $e^{-\omega i x\cdot\xi} f(x)$ 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 it, we argue as in the proof of Theorem 2.4.19. 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

$$D_N(\xi) = \sum_{m=1}^{(2N^2)^n} e^{-\omega i y_m\cdot\xi}(-\omega i y_m)^\alpha f(y_m)|Q_m| - \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi}(-\omega i x)^\alpha f(x)dx.$$

We must show that for every $M > 0$, $\sup\limits_{|\xi|\leqslant M} |D_N(\xi)|$ converges to zero as $N \to \infty$.

Setting $g(x)=(-\omega ix)^\alpha f(x)\in\mathscr{S}$, we write

$$D_N(\xi)=\sum_{m=1}^{(2N^2)^n}\int_{Q_m}[e^{-\omega iy_m\cdot\xi}g(y_m)-e^{-\omega ix\cdot\xi}g(x)]dx-\int_{([-N,N]^n)^c}e^{-\omega ix\cdot\xi}g(x)dx.$$

Using the mean value theorem, we bound the absolute value of the expression inside the square brackets by

$$(|\nabla g(z_m)|+|\omega||\xi||g(z_m)|)\frac{\sqrt{n}}{2N}\leqslant\frac{C_K(1+|\xi|)}{(2+|z_m|)^K}\frac{\sqrt{n}}{N},$$

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 $L>n$,

$$\begin{aligned}&\int_{([-N,N]^n)^c}\frac{dy}{(1+|y|)^L}\\ \leqslant&\omega_{n-1}\int_N^\infty\frac{r^{n-1}dr}{(1+r)^L}\leqslant\frac{\omega_{n-1}}{L-n}N^{n-L},\end{aligned}$$

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 $\xi\mapsto p(\xi)e^{-\omega ix\cdot\xi}$ is a well-defined function of x, which we denote by $u(p(\cdot)e^{-\omega ix\cdot(\cdot)})$. 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 $\xi\cdot z=\sum\limits_{k=1}^n\xi_kz_k$.

It is straightforward to verify that the function of z

$$F(z)=\left(\frac{|\omega|}{2\pi}\right)^{n/2}u(e^{-\omega iz\cdot(\cdot)})$$

defined on $\mathbb{C}^n$ is holomorphic, in fact entire. Indeed, the continuity and linearity of u and the fact that $(e^{-\omega i\xi_jh}-1)/h\to-\omega i\xi_j$ in $\mathcal{C}^\infty(\mathbb{R}^n)$ 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 to the $\mathcal{C}^\infty$ function

$$\xi\mapsto(-\omega i\xi_j)e^{-\omega i\sum_{j=1}^n z_j\xi_j}.$$

By induction, it follows that for all $\alpha\in\mathbb{N}_0^n$, we have

$$\partial_{z_1}^{\alpha_1}\cdots\partial_{z_n}^{\alpha_n}F=u\left((-\omega i(\cdot))^\alpha e^{-\omega i\sum_{j=1}^n z_j(\cdot)_j}\right).$$

Since F is entire, its restriction on $\mathbb{R}^n$, i.e., $F(x_1,\cdots,x_n)$, where $x_j=\Re z_j$, is real analytic. Also, an easy calculation using (2.4.3) 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 tempered distribution $\widehat{u}(x)$ can be identified with the real analytic function $F(x)$ whose derivatives have polynomial growth at infinity. ■

Finally, we finish this section by giving a density result.

Theorem 2.4.28: Density

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

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

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

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

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

$$\begin{aligned}\langle T_k, \varphi \rangle \equiv& \langle \psi_{k^{-1}} * T'_k, \varphi \rangle = \langle T'_k, \widetilde{\psi_{k^{-1}}} * \varphi \rangle = \langle \zeta(x/k)T, \psi_{k^{-1}} * \varphi \rangle \\ =& \langle T, \zeta(x/k)(\psi_{k^{-1}} * \varphi) \rangle \to \langle T, \varphi \rangle, \quad \text{as } k \to \infty,\end{aligned}$$

by the fact $\psi_{k^{-1}} * \varphi \to \varphi$ in $\mathscr{S}$ as $k \to \infty$ in view of Theorem 2.1.15, and the fact $\zeta(x/k) \to 1$ pointwise as $k \to \infty$ since $\zeta(0) = 1$ and $\zeta(x/k)\varphi \to \varphi$ in $\mathscr{S}$ as $k \to \infty$.

Finally, since $\zeta \in \mathscr{D}(\mathbb{R}^n)$, it follows from Proposition 2.4.26 that T'_k is a tempered distribution with compact support, then due to $\psi_{k^{-1}}$, T_k is $\mathcal{C}^\infty$ function with compact support by Theorem 2.4.19, namely, $T_k \in \mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$. ■

Remark 2.4.29. From the proof, it follows that $\mathscr{D}(\mathbb{R}^n)$ is also dense in $\mathscr{S}'(\mathbb{R}^n)$ with respect to the topology on $\mathscr{S}'(\mathbb{R}^n)$.

§2.5 Characterization of operators commuting with translations

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

Definition 2.5.1.

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

$$T(\tau^y f) = \tau^y(Tf)$$

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

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

Theorem 2.5.2.

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

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

The theorem will be a consequence of the following lemma.

Lemma 2.5.3.

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

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

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

Proof. Let $\xi \in \mathbb{R}^n$. Then there exists a C'_n such that (cf. (2.3.6))

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

Let us first suppose $p=1$, we shall show $\widehat{f} \in L^1$. By part (v) in Proposition 2.1.2 and part (i) in Theorem 2.1.4, we have

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

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

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

Thus, by Theorem 2.1.19, f equals almost everywhere a continuous function g and by Theorem 2.1.4,

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

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, $f\varphi$ equals almost everywhere a continuous function h such that

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

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

$$\|\partial^\alpha(f\varphi)\|_1 \leqslant \int_{|x| \leqslant 2} \sum_{\mu+\nu=\alpha} \frac{\alpha!}{\mu!\nu!} |\partial^\mu f||\partial^\nu \varphi| dx$$

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

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,

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

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

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

Proof of Theorem 2.5.2. We first prove that

$$\partial^\beta T f = T \partial^\beta f, \quad \forall f \in \mathscr{S}(\mathbb{R}^n). \tag{2.5.1}$$

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 2.3.17, $\frac{\tau^h f - f}{h_j} \to -\frac{\partial f}{\partial x_j}$ in $\mathscr{S}$ as $|h| \to 0$ and also in L^p norm due to the density of $\mathscr{S}$ in L^p. Since T is bounded operator from L^p to L^q, it follows that $\frac{\tau^h(Tf)-Tf}{h_j} \to -\frac{\partial Tf}{\partial x_j}$ in L^q as $|h| \to 0$. By induction, we get (2.5.1). By Lemma 2.5.3, Tf equals almost everywhere a continuous function g_f satisfying

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

From the proof of Theorem 2.3.11, we know that the L^p norm of $f \in \mathscr{S}$ is bounded by a finite linear combination of L^∞ norms of terms of the form $x^\alpha f(x)$. Thus, there exists an $m \in \mathbb{N}$ such that

$$|g_f(0)| \leqslant C \sum_{|\alpha|\leqslant m, |\beta|\leqslant n+1} \|x^\alpha \partial^\beta f\|_\infty = C \sum_{|\alpha|\leqslant m, |\beta|\leqslant n+1} |f|_{\alpha,\beta}.$$

Then, by Theorem 2.4.4, 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 2.4.19, we obtain

$$(u * f)(x) = \langle u, \tau^x \widetilde{f} \rangle = \langle u, \widetilde{\tau^{-x} f} \rangle = \langle \widetilde{u}, \tau^{-x} f \rangle = \langle u_1, \tau^{-x} f \rangle$$
$$= (T(\tau^{-x} f))(0) = (\tau^{-x} T f)(0) = Tf(x).$$

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

Combining this result with Theorem 2.4.19, we obtain the fact that Tf, for $f \in \mathscr{S}$, is almost everywhere equal to a $\mathcal{C}^\infty$ function which, together with all its derivatives, is slowly increasing.

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

Theorem 2.5.4.

Let T be a bounded linear operator mapping $L^1(\mathbb{R}^n)$ to itself. Then T commutes with translations if and only if there exists a finite Borel measure $\mu \in \mathscr{B}(\mathbb{R}^n)$ such that $Tf = \mu * f$, for all $f \in L^1(\mathbb{R}^n)$. We also have $\|T\| = \|\mu\|$, where $\|\mu\|$ is the total variation of the measure μ.

Proof. We first prove the sufficiency. Suppose that $Tf = \mu * f$ for a measure $\mu \in \mathscr{B}(\mathbb{R}^n)$ and all $f \in L^1(\mathbb{R}^n)$. Since $\mathscr{B} \subset \mathscr{S}'$, by Theorem 2.4.19, we have

$$\begin{aligned}\tau^h(Tf)(x) =&(Tf)(x-h) = \langle \mu, \tau^{x-h}\tilde{f}\rangle = \langle \mu(y), f(-y-x+h)\rangle\\ =&\langle \mu, \tau^x\widetilde{\tau^h f}\rangle = (\mu * \tau^h f)(x) = T(\tau^h f)(x),\end{aligned}$$

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

Now, we prove the necessariness. Suppose that T commutes with translations and $\|Tf\|_1 \leqslant \|T\|\|f\|_1$ for all $f \in L^1(\mathbb{R}^n)$. Then, by Theorem 2.5.2, there exists a unique tempered distribution μ such that $Tf = \mu * f$ for all $f \in \mathscr{S}$. The remainder is to prove $\mu \in \mathscr{B}(\mathbb{R}^n)$.

We consider the family of L^1 functions $\mu_\varepsilon = \mu * W(\cdot, \varepsilon) = TW(\cdot, \varepsilon)$, $\varepsilon > 0$. Then by assumption and Lemma 2.1.14, we get

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

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

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi(x)\mu_{\varepsilon_k}(x)dx = \int_{\mathbb{R}^n} \varphi(x)d\nu(x). \tag{2.5.2}$$

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

Therefore, we must show that $\langle \mu, \psi\rangle = \int_{\mathbb{R}^n} \psi(x)d\nu(x)$ for all $\psi \in \mathscr{S}$. Let $\psi_\varepsilon = W(\cdot, \varepsilon) * \psi$. Then, for all $\alpha \in \mathbb{N}_0^n$, we have $\partial^\alpha \psi_\varepsilon = W(\cdot, \varepsilon) * \partial^\alpha \psi$. It follows from Theorem 2.1.15 that $\partial^\alpha \psi_\varepsilon(x)$ converges to $\partial^\alpha \psi(x)$ uniformly in x. Thus, $\psi_\varepsilon \to \psi$ in $\mathscr{S}$ as $\varepsilon \to 0$ and this implies that $\langle \mu, \psi_\varepsilon\rangle \to \langle \mu, \psi\rangle$. But, since $W(\cdot, \varepsilon) = \widetilde{W(\cdot, \varepsilon)}$,

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

Thus, putting $\varepsilon = \varepsilon_k$, letting $k \to \infty$ and applying (2.5.2) with $\varphi = \psi$, we obtain the desired equality $\langle \mu, \psi\rangle = \int_{\mathbb{R}^n} \psi(x)d\nu(x)$. Hence, $\mu \in \mathscr{B}$.

Next, (2.5.2) implies that for all $\varphi \in C_0$, it holds

$$\left| \int_{\mathbb{R}^n} \varphi(x) d\mu(x) \right| \leqslant \|\varphi\|_\infty \sup_k \|\mu_{\varepsilon_k}\|_1 \leqslant \|\varphi\|_\infty \|T\|. \tag{2.5.3}$$

The Riesz representation theorem gives that the norm of the functional

$$g \mapsto \int_{\mathbb{R}^n} \varphi(x) d\mu(x)$$

on C_0 is exactly $\|\mu\|$. It follows from (2.5.3) that $\|T\| \geqslant \|\mu\|$. Thus, combining with the previous reverse inequality, we conclude that $\|T\| = \|\mu\|$. This completes the proof. ∎

Let μ be a finite Borel measure. The operator $h \mapsto h * u$ maps $L^p(\mathbb{R}^n)$ to itself for all $p \in [1, \infty]$. But there exists bounded linear operators T on L^∞ that commute with translations for which there does not exist a finite Borel measure μ such that $Th = h * \mu$ for all $h \in L^\infty(\mathbb{R}^n)$. The following example captures such a behavior, which also implies that the restriction of T on $\mathscr{S}$ does not uniquely determine T on the entire L^∞.

Example 2.5.5. Let $(X, \|\cdot\|_\infty)$ be the space of all complex-valued bounded functions on the real line such that

$$Tf = \lim_{R\to\infty} \frac{1}{R} \int_0^R f(t) dt$$

exists. Then, T is a bounded linear functional on X with norm 1 and has a bounded extension $\widetilde{T}$ on L^∞ with norm 1, by the Hahn-Banach theorem. We may think of $\widetilde{T}$ as a bounded linear operator from $L^\infty(\mathbb{R})$ to the space of constant functions, which is contained in $L^\infty(\mathbb{R})$. We note that $\widetilde{T}$ commutes with translations, since for all $f \in L^\infty(\mathbb{R})$ and $x \in \mathbb{R}$, we have

$$\widetilde{T}(\tau^x f) - \tau^x(\widetilde{T} f) = \widetilde{T}(\tau^x f) - \widetilde{T} f = \widetilde{T}(\tau^x f - f) = T(\tau^x f - f) = 0,$$

where the last two equalities follow from the fact that for L^∞ functions f and $R > |x|$,

$$\begin{aligned}
\frac{1}{R}\int_0^R (f(t-x) - f(t))\, dt =& \frac{1}{R}\left(\int_{-x}^{R-x} f(t)dt - \int_0^R f(t)dt\right) \\
=& \frac{1}{R}\left(\int_{-x}^{0} f(t)dt - \int_{R-x}^R f(t)dt\right) \\
\leqslant& \frac{2|x|}{R}\|f\|_\infty \to 0, \quad \text{as } R \to \infty.
\end{aligned}$$

If $T\varphi = \varphi * u$ for some $u \in \mathscr{S}'(\mathbb{R})$ and all $\varphi \in \mathscr{S}(\mathbb{R})$, since T vanishes on $\mathscr{S}$, i.e., $|T\varphi| \leqslant \lim\limits_{R\to\infty} \frac{\|\varphi\|_1}{R} = 0$, the uniqueness in Theorem 2.5.2 yields that $u = 0$. Hence, if there existed a finite Borel measure μ such that $\widetilde{T}h = h * \mu$ for all $h \in L^\infty$, in particular we would have $0 = T\varphi = \varphi * \mu$ for all $\varphi \in \mathscr{S}$, thus μ would be the zero measure. But obviously, this is not the case, since T is not the zero operator on X.

For the case $p = 2$, we have a very simple characterization of these operators.

Theorem 2.5.6.

Let T be a bounded linear transformation mapping $L^2(\mathbb{R}^n)$ to itself. Then T commutes with translation if and only if there exists an $m \in L^\infty(\mathbb{R}^n)$ such that $Tf = u * f$ with $\hat{u} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} m$, for all $f \in L^2(\mathbb{R}^n)$. We also have $\|T\| = \|m\|_\infty$.

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

$$\mathscr{F}(u * \varphi) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \hat{u}\hat{\varphi}. \tag{2.5.4}$$

To see this, we must show that $\langle \mathscr{F}(u * \varphi), \psi\rangle = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \langle \hat{u}\hat{\varphi}, \psi\rangle$ for all $\psi \in \mathscr{S}$. It follows immediately, from (2.4.9), part (vi) in Proposition 2.1.2 and the Fourier inversion formula, that

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

Thus, (2.5.4) is established.

Now, we prove the necessariness. Suppose that T commutes with translations and $\|Tf\|_2 \leqslant \|T\|\|f\|_2$ for all $f \in L^2(\mathbb{R}^n)$. Then, by Theorem 2.5.2, there exists a unique tempered distribution u such that $Tf = u * f$ for all $f \in \mathscr{S}$. The remainder is to prove $\hat{u} \in L^\infty(\mathbb{R}^n)$.

Let $\varphi_0 = e^{-\frac{|\omega|}{2}|x|^2}$, then, we have $\varphi_0 \in \mathscr{S}$ and $\hat{\varphi}_0 = \varphi_0$ by Corollary 2.1.10. Thus, $T\varphi_0 = u * \varphi_0 \in L^2$ and therefore $\Phi_0 := \mathscr{F}(u * \varphi_0) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \hat{u}\hat{\varphi}_0 \in L^2$ by (2.5.4) and the Plancherel theorem. Let $m(\xi) = \Phi_0(\xi)/\hat{\varphi}_0(\xi)$.

We claim that

$$\mathscr{F}(u * \varphi) = m\hat{\varphi} \tag{2.5.5}$$

for all $\varphi \in \mathscr{S}$. By (2.5.4), it suffices to show that $\langle \hat{u}\hat{\varphi}, \psi\rangle = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \langle m\hat{\varphi}, \psi\rangle$ for all $\psi \in \mathscr{D}$ since $\mathscr{D}$ is dense in $\mathscr{S}$. But, if $\psi \in \mathscr{D}$, then $(\psi/\hat{\varphi}_0)(\xi) = \psi(\xi)e^{\frac{|\omega|}{2}|\xi|^2} \in \mathscr{D}$; thus,

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

It follows immediately that $\hat{u} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} m$. In fact, we have just shown that $\langle \hat{u}, \hat{\varphi}\psi\rangle = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \langle m\hat{\varphi}, \psi\rangle = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \langle m, \hat{\varphi}\psi\rangle$ for all $\varphi \in \mathscr{S}$ and $\psi \in \mathscr{D}$. Selecting

φ such that $\widehat{\varphi}(\xi) = 1$ for $\xi \in \operatorname{supp}\psi$, this shows that $\langle \widehat{u}, \psi \rangle = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \langle m, \psi \rangle$ for all $\psi \in \mathscr{D}$. Thus, $\widehat{u} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} m$.

Due to

$$\|m\widehat{\varphi}\|_2 = \|\mathscr{F}(u * \varphi)\|_2 = \|u * \varphi\|_2 \leqslant \|T\|\|\varphi\|_2 = \|T\|\|\widehat{\varphi}\|_2$$

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

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

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

Finally, we can show the sufficiency easily. If $\widehat{u} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} m \in L^\infty(\mathbb{R}^n)$, the Plancherel theorem and (2.5.4) immediately imply that

$$\|Tf\|_2 = \|u * f\|_2 = \|m\widehat{f}\|_2 \leqslant \|m\|_\infty \|f\|_2$$

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

Thus, if $m = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{u} \in L^\infty$, then $\|T\| = \|m\|_\infty$. ■

For further results, one can see [SW71, p.30] and [Gra14, p.153-155].

§ 2.6 Fourier multipliers on L^p

We have characterized all convolution operators that map L^1 to L^1 or L^2 to L^2. In this section, we introduce briefly the Fourier multipliers on L^p.

Definition 2.6.1.

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 $\widecheck{m} * f \in L^p(\mathbb{R}^n)$ for all $f \in \mathscr{S}(\mathbb{R}^n)$, and if

$$\|m\|_{\mathcal{M}_p(\mathbb{R}^n)} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \sup_{\|f\|_p=1} \|\widecheck{m} * f\|_p$$

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: $f \mapsto \widecheck{m} * f$ can be extended to a mapping from L^p to L^p with the same norm. We write $\widecheck{m} * f$ also for the values of the extended mapping.

For $p = \infty$ (as well as for $p = 2$) we can characterize $\mathcal{M}_p$. Considering the map:

$$f \mapsto \widecheck{m} * f \quad \text{for } f \in \mathscr{S},$$

we have

$$m \in M_\infty \Leftrightarrow |(\widecheck{m} * f)(0)| \leqslant C\|f\|_\infty, \quad f \in \mathscr{S}. \tag{2.6.1}$$

Indeed, if $m \in M_\infty$, we have

$$|(\widecheck{m} * f)(0)| \leqslant \frac{\|\widecheck{m} * f\|_\infty}{\|f\|_\infty}\|f\|_\infty \leqslant C\|f\|_\infty.$$

On the other hand, if $|(\widecheck{m} * f)(0)| \leqslant C\|f\|_\infty$, we can get

$$\begin{aligned}\|\widecheck{m} * f\|_\infty = &\sup_{x\in\mathbb{R}^n} |(\widecheck{m} * f)(x)| = \sup_{x\in\mathbb{R}^n} |[\widecheck{m} * (f(x + \cdot))](0)| \\ \leqslant &C\|f(x + \cdot)\|_\infty = C\|f\|_\infty,\end{aligned}$$

which yields $m \in M_\infty$.

But (2.6.1) also means that $\check{m}$ is a bounded measure on $\mathbb{R}^n$. Thus, M_∞ is equal to the space of all Fourier transforms of bounded measures. Moreover, $\|m\|_{M_\infty}$ is equal to the total mass of $\check{m}$. In view of the inequality above and the Hahn-Banach theorem, we may extend the mapping $f \mapsto \check{m} * f$ from $\mathscr{S}$ to L^∞ to a mapping from L^∞ to L^∞ without increasing its norm. We also write the extended mapping as $f \mapsto \check{m} * f$ for $f \in L^\infty$.

Theorem 2.6.2.

Let $1 \leqslant p \leqslant \infty$ and $1/p + 1/p' = 1$, then we have

$$\mathcal{M}_p(\mathbb{R}^n) = \mathcal{M}_{p'}(\mathbb{R}^n) \quad \text{(equal norms)}. \tag{2.6.2}$$

Moreover,

$$\begin{aligned} \mathcal{M}_1(\mathbb{R}^n) &= \left\{ m \in \mathscr{S}'(\mathbb{R}^n) : \check{m} \text{ is a bounded measure on } \mathbb{R}^n \right\}, \\ \|m\|_{\mathcal{M}_1(\mathbb{R}^n)} &= \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\check{m}\|_1, \end{aligned} \tag{2.6.3}$$

and

$$\mathcal{M}_2(\mathbb{R}^n) = L^\infty(\mathbb{R}^n) \quad \text{(equal norms)}. \tag{2.6.4}$$

For the norms ($1 \leqslant p_0, p_1 \leqslant \infty$),

$$\|m\|_{\mathcal{M}_p(\mathbb{R}^n)} \leqslant \|m\|^{1-\theta}_{\mathcal{M}_{p_0}(\mathbb{R}^n)} \|m\|^{\theta}_{\mathcal{M}_{p_1}(\mathbb{R}^n)}, \quad \forall m \in \mathcal{M}_{p_0}(\mathbb{R}^n) \cap \mathcal{M}_{p_1}(\mathbb{R}^n), \tag{2.6.5}$$

if $1/p = (1-\theta)/p_0 + \theta/p_1$ ($0 \leqslant \theta \leqslant 1$). In particular, the norm $\|\cdot\|_{M^p(\mathbb{R}^n)}$ decreases with p in the interval $1 \leqslant p \leqslant 2$, and

$$\mathcal{M}_1 \hookrightarrow \mathcal{M}_p \hookrightarrow \mathcal{M}_q \hookrightarrow \mathcal{M}_2, \quad (1 \leqslant p \leqslant q \leqslant 2). \tag{2.6.6}$$

Proof. Let $f \in L^p$, $g \in L^{p'}$ and $m \in \mathcal{M}_p$. Then, we have

$$\begin{aligned} \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|m\|_{\mathcal{M}_{p'}} &= \sup_{\|g\|_{p'}=1} \|\check{m} * g\|_{p'} = \sup_{\|f\|_p = \|g\|_{p'}=1} |\langle \check{m} * g, \tilde{f}\rangle| \\ &= \sup_{\|f\|_p = \|g\|_{p'}=1} |(\check{m} * g * f)(0)| = \sup_{\|f\|_p = \|g\|_{p'}=1} |(\check{m} * f * g)(0)| \\ &= \sup_{\|f\|_p = \|g\|_{p'}=1} |\langle \check{m} * f, \tilde{g}\rangle| \\ &= \sup_{\|f\|_p=1} \|\check{m} * f\|_p = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|m\|_{\mathcal{M}_p}. \end{aligned}$$

The assertion (2.6.3) has already been established because of $\mathcal{M}_1 = M_\infty$. The Plancherel theorem immediately gives (2.6.4). In fact,

$$\begin{aligned} \|m\|_{\mathcal{M}_2} &= \left(\frac{|\omega|}{2\pi}\right)^{n/2} \sup_{\|f\|_2=1} \|\check{m} * f\|_2 \\ &= \left(\frac{|\omega|}{2\pi}\right)^{n/2} \sup_{\|f\|_2=1} \|\widehat{\check{m} * f}\|_2 \\ &= \sup_{\|f\|_2=1} \|m\hat{f}\|_2 \\ &\leqslant \|m\|_\infty. \end{aligned}$$

On the other hand, for any given $\varepsilon > 0$, let

$$E_\varepsilon = \{\xi : |\xi| \leqslant 1/\varepsilon \text{ and } |m(\xi)| > \|m\|_\infty - \varepsilon\}.$$

Then E_ε has positive and finite measure, and let $f \in L^2$ be such that $\operatorname{supp} \widehat{f} \subset E_\varepsilon$. Hence, we can obtain

$$\begin{aligned}\left(\frac{|\omega|}{2\pi}\right)^n \|\widecheck{m} * f\|_2^2 =&\|m\widehat{f}\|_2^2 = \int_{E_\varepsilon} |m(\xi)\widehat{f}(\xi)|^2 d\xi \\ \geqslant&(\|m\|_\infty - \varepsilon)^2 \int |\widehat{f}(\xi)|^2 d\xi \\ =&(\|m\|_\infty - \varepsilon)^2 \|f\|_2^2.\end{aligned}$$

It follows that $\|m\|_{\mathcal{M}_2} \geqslant \|m\|_\infty$, and then the equality holds.

Invoking the Riesz-Thorin theorem, (2.6.5) follows, since the mapping $f \mapsto \widecheck{m} * f$ 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 $\|m\|_{\mathcal{M}_{p_1}}$.

Since $1/q = (1-\theta)/p + \theta/p'$ for some θ and $p \leqslant q \leqslant 2 \leqslant p'$, by using (2.6.5) with $p_0 = p$, $p_1 = p'$, we see that

$$\|m\|_{\mathcal{M}_q} \leqslant \|m\|_{\mathcal{M}_p},$$

from which (2.6.6) follows. ■

Proposition 2.6.3.

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 $\widecheck{m}_k * f \to \widecheck{m} * f$ for any $f \in \mathscr{S}$ in sense of the strong topology on $\mathscr{S}'$. On the other hand, $\{\widecheck{m}_k * f\}$ 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 $g = \widecheck{m} * f$. 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

$$\begin{aligned}\left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\widecheck{(m_1 m_2)} * f\|_p =& \left(\frac{|\omega|}{2\pi}\right)^n \|\widecheck{m_1} * \widecheck{m_2} * f\|_p \\ \leqslant& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|m_1\|_{\mathcal{M}_p} \|\widecheck{m_2} * f\|_p \\ \leqslant& \|m_1\|_{\mathcal{M}_p} \|m_2\|_{\mathcal{M}_p} \|f\|_p,\end{aligned}$$

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 says that $\mathcal{M}_p(\mathbb{R}^n)$ is isometrically invariant under affine transforms[2] of $\mathbb{R}^n$.

Theorem 2.6.4.

Let $a : \mathbb{R}^n \to \mathbb{R}^k$ be a surjective affine transform with $n \geqslant k$, and $m \in \mathcal{M}_p(\mathbb{R}^k)$.

[2] An affine transform of $\mathbb{R}^n$ is a map $F : \mathbb{R}^n \to \mathbb{R}^n$ of the form $F(\mathrm{p}) = A\mathrm{p} + \mathrm{q}$ for all $\mathrm{p} \in \mathbb{R}^n$, where A is a linear transform of $\mathbb{R}^n$ and $\mathrm{q} \in \mathbb{R}^n$.

Then

$$\|m\circ a\|_{\mathcal{M}_p(\mathbb{R}^n)}=\|m\|_{\mathcal{M}_p(\mathbb{R}^k)}.$$

In particular, we have

$$\|\delta^c m\|_{\mathcal{M}_p(\mathbb{R}^n)}=\|m\|_{\mathcal{M}_p(\mathbb{R}^n)},\quad \forall c>0, \tag{2.6.7}$$

$$\|\tilde{m}\|_{\mathcal{M}_p(\mathbb{R}^n)}=\|m\|_{\mathcal{M}_p(\mathbb{R}^n)}, \tag{2.6.8}$$

$$\|m(\langle x,\cdot\rangle)\|_{\mathcal{M}_p(\mathbb{R}^n)}=\|m\|_{\mathcal{M}_p(\mathbb{R})},\quad \forall x\neq 0, \tag{2.6.9}$$

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

$$\eta_i=a_i(\xi),\ 1\leqslant i\leqslant k;\quad \eta_j=\xi_j,\ k+1\leqslant j\leqslant n, \tag{2.6.10}$$

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

$$\begin{aligned}
&\mathscr{F}^{-1}(m(a(\xi))\widehat{f})(x)=\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}e^{\omega i x\cdot\xi}m(a(\xi))\widehat{f}(\xi)d\xi\\
=&|\det A|\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}e^{\omega i x\cdot A\eta}m(\eta')\widehat{f}(A\eta)d\eta\\
=&|\det A|\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}e^{\omega i A^\top x\cdot\eta}m(\eta')\widehat{f}(A\eta)d\eta\\
=&\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}e^{\omega i A^\top x\cdot\eta}m(\eta')\widehat{f((A^\top)^{-1}\cdot)}(\eta)d\eta\\
=&\left(\frac{|\omega|}{2\pi}\right)^{k/2}\int_{\mathbb{R}^k}e^{\omega i(A^\top x)'\cdot\eta'}m(\eta')\\
&\left(\left(\frac{|\omega|}{2\pi}\right)^{(n-k)/2}\int_{\mathbb{R}^{n-k}}e^{\omega i(A^\top x)''\cdot\eta''}\widehat{f((A^\top)^{-1}\cdot)}(\eta',\eta'')d\eta''\right)d\eta'\\
=&\left(\frac{|\omega|}{2\pi}\right)^{k/2}\int_{\mathbb{R}^k}e^{\omega i(A^\top x)'\cdot\eta'}m(\eta')\left(\mathscr{F}^{-1}_{\eta''}[\widehat{f((A^\top)^{-1}\cdot)}]\right)(\eta',(A^\top x)'')d\eta'\\
=&\left(\frac{|\omega|}{2\pi}\right)^{k/2}\int_{\mathbb{R}^k}e^{\omega i(A^\top x)'\cdot\eta'}m(\eta')\left([\mathscr{F}_{x'}(f((A^\top)^{-1}\cdot))]\right)(\eta',(A^\top x)'')d\eta'\\
=&\mathscr{F}^{-1}_{\eta'}\left[m(\eta')\left([\mathscr{F}_{x'}(f((A^\top)^{-1}\cdot))]\right)(\eta',(A^\top x)'')\right]((A^\top x)')\\
=&\left(\frac{|\omega|}{2\pi}\right)^{k/2}\int_{\mathbb{R}^k}\widecheck{m}(y')f((A^\top)^{-1}((A^\top x)'-y',(A^\top x)''))dy'.
\end{aligned}$$

It follows from $m\in\mathcal{M}_p(\mathbb{R}^k)$ that for any $f\in\mathscr{S}(\mathbb{R}^n)$

$$\begin{aligned}
&\left(\frac{|\omega|}{2\pi}\right)^{np/2}\|\mathscr{F}^{-1}(m(a(\xi)))*f\|^p_{L^p(\mathbb{R}^n)}\\
=&\|\mathscr{F}^{-1}(m(a(\xi))\widehat{f})\|^p_{L^p(\mathbb{R}^n)}\\
=&\left(\frac{|\omega|}{2\pi}\right)^{kp/2}\left\|\int_{\mathbb{R}^k}\widecheck{m}(y')f((A^\top)^{-1}((A^\top x)'-y',(A^\top x)''))dy'\right\|^p_{L^p(\mathbb{R}^n)}
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{|\omega|}{2\pi}\right)^{kp/2} |\det A|^{-1} \left\| \int_{\mathbb{R}^k} \check{m}(y') f((A^\top)^{-1}(x'-y',x''))dy' \right\|_{L^p(\mathbb{R}^n)}^p \\
&\leqslant |\det A|^{-1} \|m\|_{\mathcal{M}_p(\mathbb{R}^k)}^p \left\| \|f((A^\top)^{-1}(x',x''))\|_{L^p(\mathbb{R}^k)}^p \right\|_{L^p(\mathbb{R}^{n-k})}^p \\
&= |\det A|^{-1} \|m\|_{\mathcal{M}_p(\mathbb{R}^k)}^p \|f((A^\top)^{-1}(x))\|_{L^p(\mathbb{R}^n)}^p \\
&= \|m\|_{\mathcal{M}_p(\mathbb{R}^k)}^p \|f\|_{L^p(\mathbb{R}^n)}^p.
\end{aligned}
$$

Thus, we have

$$
\|m(a(\cdot))\|_{\mathcal{M}_p(\mathbb{R}^n)} \leqslant \|m\|_{\mathcal{M}_p(\mathbb{R}^k)}. \tag{2.6.11}
$$

Taking $f((A^\top)^{-1}x) = f_1(x')f_2(x'')$, one can conclude that the reverse inequality (2.6.11) also holds. ■

Now we give a simple but very useful theorem for Fourier multipliers.

Theorem 2.6.5: Bernstein multiplier theorem

Assume that $k > n/2$ is an integer, and that $\partial_{x_j}^\alpha m \in L^2(\mathbb{R}^n)$, $j = 1, \cdots, n$ and $0 \leqslant \alpha \leqslant k$. Then we have $m \in \mathcal{M}_p(\mathbb{R}^n)$ for $1 \leqslant p \leqslant \infty$, and

$$
\|m\|_{\mathcal{M}_p} \lesssim \|m\|_2^{1-n/2k} \left(\sum_{j=1}^n \|\partial_{x_j}^k m\|_2 \right)^{n/2k}.
$$

Proof. Let $t > 0$ and $J(x) = \sum_{j=1}^n |x_j|^k$. By the Cauchy-Schwarz inequality and the Plancherel theorem, we obtain

$$
\int_{|x|>t} |\check{m}(x)|dx = \int_{|x|>t} J(x)^{-1} J(x) |\check{m}(x)| dx \lesssim t^{n/2-k} \sum_{j=1}^n \|\partial_{x_j}^k m\|_2.
$$

Similarly, we have

$$
\int_{|x|\leqslant t} |\check{m}(x)|dx \lesssim t^{n/2} \|m\|_2.
$$

Choosing t such that $\|m\|_2 = t^{-k} \sum_{j=1}^n \|\partial_{x_j}^k m\|_2$, we infer, with the help of Theorem 2.6.2, that

$$
\|m\|_{\mathcal{M}_p} \leqslant \|m\|_{\mathcal{M}_1} = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} |\check{m}(x)|dx \lesssim \|m\|_2^{1-n/2k} \left(\sum_{j=1}^n \|\partial_{x_j}^k m\|_2 \right)^{n/2k}.
$$

This completes the proof. ■

Remark 2.6.6. 1) From the proof of Theorem 2.6.5, we see that $\check{m} \in L^1$, in other words, it is equivalent to the Young inequality for convolution, i.e., $\|\check{m} * f\|_p \leqslant \|\check{m}\|_1 \|f\|_p$ 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 the rectangular pulse function and the sinc function introduced in Example 2.1.5.

References

[Abe11] Helmut Abels. Short lecture notes: Interpolation theory and function spaces. May 2011.

[Bab61] K. I. Babenko. An inequality in the theory of Fourier integrals. *Izv. Akad. Nauk SSSR Ser. Mat.*, 25:531–542, 1961.

[BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343 of *GMW*. Springer-Verlag, Berlin Heidelberg, 2011.

[Bec75] William Beckner. Inequalities in Fourier analysis on R^n. *Proc. Nat. Acad. Sci. U.S.A.*, 72:638–641, 1975.

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[Gra14] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[Mit13] Dorina Mitrea. *Distributions, partial differential equations, and harmonic analysis*. Universitext. Springer, New York, 2013.

[Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[SW71] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

[WHHG11] Baoxiang Wang, Zhaohui Huo, Chengchun Hao, and Zihua Guo. *Harmonic analysis method for nonlinear evolution equations*, volume I. World Scientific Publishing Co. Pte. Ltd., 2011.

III

The Maximal Function and Calderón-Zygmund Decomposition

§3.1 Two covering lemmas

Lemma 3.1.1: Finite version of Vitali covering lemma

Suppose $\mathcal{B} = \{B_1, B_2, \cdots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^n$. Then, there exists a disjoint sub-collection $B_{j_1}, B_{j_2}, \cdots, B_{j_k}$ of $\mathcal{B}$ such that

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

Proof. The argument we give is constructive and relies on the following simple observation:

Suppose B and B' are a pair of balls that intersect, with the radius of B' being not greater than that of B. Then B' is contained in the ball $\tilde{B}$ that is concentric with B but with 3 times its radius. (See Fig 3.1.)

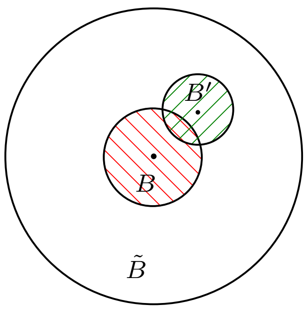


Figure 3.1: The balls B and $\tilde{B}$

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

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

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. ∎

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

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

Theorem 3.1.2: Whitney covering lemma

Let F be a non-empty closed set in $\mathbb{R}^n$ and Ω be its complement. Then there exists a 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 $\mathscr{M}_k = 2^{-k}\mathscr{M}_0$.

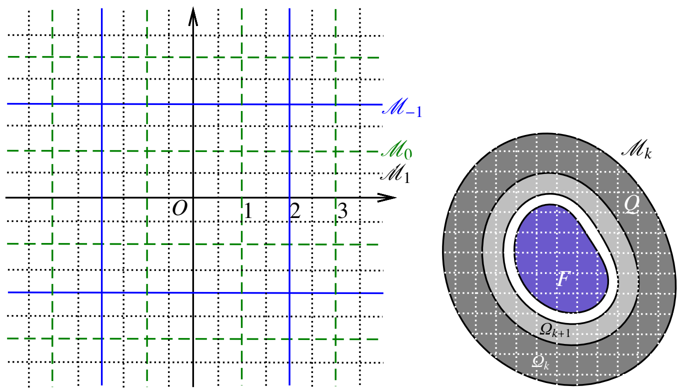


Figure 3.2: Meshes and layers: $\mathscr{M}_0$ with dashed lines; $\mathscr{M}_1$ with dotted lines; $\mathscr{M}_{-1}$ with solid lines

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

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

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

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

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

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

For appropriate choice of c, we claim that

$$\operatorname{diam}(Q) \leqslant \operatorname{dist}(Q, F) \leqslant 4 \operatorname{diam}(Q), \quad Q \in \mathcal{F}_0. \tag{3.1.1}$$

Let us prove (3.1.1) 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 get (3.1.1). Then by (3.1.1) the cubes $Q \in \mathcal{F}_0$ are disjoint from F and clearly cover Ω. Therefore, (i) is also proved.

Notice 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 which were really unnecessary.

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

Start now with any cube $Q \in \mathcal{F}_0$, and consider the maximal cube in $\mathcal{F}_0$ which contains it. In view of the inequality (3.1.1), for any cube $Q' \in \mathcal{F}_0$ which 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'' which contain Q have obviously a non-trivial intersection. Thus, by the observation made above each cube $Q \in \mathcal{F}_0$ has a unique maximal cube in $\mathcal{F}_0$ which 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 disjoint,

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

Therefore, we complete the proof. ■

§ 3.2 Hardy-Littlewood maximal function

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

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

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

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

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

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

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

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

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

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

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

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

$$\lim_{r\to 0} \frac{1}{\mu(B(x,r))}\int_{B(x,r)} f(y)dy = f(x), \quad \text{for a.e. } x? \tag{3.2.1}$$

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

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

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

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

as desired.

§ 3.2.1 Hardy-Littlewood maximal operator

In general, for this "averaging problem" (3.2.1), we shall have an affirmative answer. In order to study the limit (3.2.1), we consider its quantitative analogue, where "$\lim_{r\to 0}$" is replaced by "$\sup_{r>0}$", this is the ***(centered) maximal function***. Since the properties of this maximal function are expressed in term of relative size and do not involve any cancellation of positive and negative values, we replace f by $|f|$.

Definition 3.2.1.

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

$$Mf(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))}\int_{B(x,r)} |f(y)|dy, \quad x\in\mathbb{R}^n, \tag{3.2.2}$$

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

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

It is immediate from the definition that

Theorem 3.2.2.

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

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

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

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

Thus, we have proved

Proposition 3.2.3.

If $f \in C(\mathbb{R}^n)$, then $|f(x)| \leqslant Mf(x)$ for all $x \in \mathbb{R}^n$.

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

$$M'f(x) = \sup_{r>0} \frac{1}{(2r)^n} \int_{Q(x,r)} |f(y)|dy, \quad x \in \mathbb{R}^n. \tag{3.2.3}$$

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

$$V_n 2^{-n} Mf(x) \leqslant M'f(x) \leqslant V_n 2^{-n} n^{n/2} Mf(x). \tag{3.2.4}$$

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

$$M''f(x) = \sup_{Q \ni x} \frac{1}{\mu(Q)} \int_Q |f(y)|dy, \tag{3.2.5}$$

where the supremum is taken over all cubes containing x. Again, M'' is point-wise 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 and the latter as the non-centered maximal operator.

Alternatively, we could define the non-centered maximal function with balls instead of cubes:

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

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

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

Example 3.2.4. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \chi_{(0,1)}(x)$. Then

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

$$\widetilde{M}f(x) = M''f(x) = \begin{cases} \frac{1}{x}, & x \geqslant 1, \\ 1, & 0 < x < 1, \\ \frac{1}{1-x}, & x \leqslant 0. \end{cases}$$

In fact, for $x \geqslant 1$, we get

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

For $0 < x < 1$, it follows

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

For $x \leqslant 0$, we have

$$\begin{aligned} Mf(x) = M'f(x) =& \max\left(\sup_{0<x+h<1, h>0} \frac{x+h}{2h}, \sup_{x+h\geqslant 1} \frac{1}{2h} \right) = \frac{1}{2(1-x)}, \\ \widetilde{M}f(x) = M''f(x) =& \max\left(\sup_{h_1,h_2>0, 0<x+h_2<1} \frac{x+h_2}{h_1+h_2}, \sup_{h_1>0, x+h_2\geqslant 1} \frac{1}{h_1+h_2} \right) \\ =& \frac{1}{1-x}. \end{aligned}$$

Observe that $f \in L^1(\mathbb{R})$, but $Mf, M'f, M''f, \widetilde{M}f \notin L^1(\mathbb{R})$.

Remark 3.2.5. (i) Mf is defined at every point $x \in \mathbb{R}^n$ and if $f = g$ a.e., then $Mf(x) = Mg(x)$ at every $x \in \mathbb{R}^n$.

(ii) It may be well that $Mf = \infty$ for every $x \in \mathbb{R}^n$. For example, let $n = 1$ and $f(x) = x^2$.

(iii) There are several definitions in the literature which are often equivalent.

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

Proposition 3.2.6.

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

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

(ii) Sub-linearity: $M(f+g)(x) \leqslant Mf(x) + Mg(x)$.

(iii) Homogeneity: $M(\alpha f)(x) = |\alpha| Mf(x)$, $\alpha \in \mathbb{R}$.

(iv) Translation invariance: $M(\tau_y f) = (\tau_y Mf)(x) = Mf(x-y)$.

With the Vitali covering lemma, we can state and prove the main results for the maximal function.

Theorem 3.2.7: 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.,
$$\mu(\{x : Mf(x) > \alpha\}) \leqslant \frac{3^n}{\alpha} \|f\|_1.$$

(iii) If $f \in L^p(\mathbb{R}^n)$, $p \in (1, \infty]$, then $Mf \in L^p(\mathbb{R}^n)$ and
$$\|Mf\|_p \leqslant A_p \|f\|_p,$$
where $A_p = 3^n p/(p-1) + 1$ for $p \in (1, \infty)$ and $A_\infty = 1$.

Proof. We first prove the second one, i.e., (ii). 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 a $r > 0$ such that
$$\frac{1}{\mu(B_x)} \int_{B_x} |f(y)| dy > \widetilde{M}f(x) - \varepsilon > \alpha,$$
where we denote by B_x the open balls contains x. Then for any $z \in B_x$, we have $\widetilde{M}f(z) > \alpha$, and so $B_x \subset E_\alpha$. This implies that E_α is open.

Therefore, for each open ball B_x, we have

$$\mu(B_x) < \frac{1}{\alpha}\int_{B_x} |f(y)|dy. \tag{3.2.6}$$

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 3.1.1 guarantees the existence of a sub-collection $B_{j_1}, \cdots, B_{j_k}$ of disjoint balls with

$$\mu\left(\bigcup_{\ell=1}^N B_\ell\right) \leqslant 3^n \sum_{i=1}^k \mu(B_{j_i}). \tag{3.2.7}$$

Since the balls $B_{j_1}, \cdots, B_{j_k}$ are disjoint and satisfy (3.2.6) as well as (3.2.7), we find that

$$\begin{aligned}\mu(K) \leqslant& \mu\left(\bigcup_{\ell=1}^N B_\ell\right) \leqslant 3^n \sum_{i=1}^k \mu(B_{j_i}) \leqslant \frac{3^n}{\alpha}\sum_{i=1}^k \int_{B_{j_i}} |f(y)|dy\\ =&\frac{3^n}{\alpha}\int_{\bigcup_{i=1}^k B_{j_i}} |f(y)|dy \leqslant \frac{3^n}{\alpha}\int_{\mathbb{R}^n} |f(y)|dy.\end{aligned}$$

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

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

The above proof also gives the proof of (i) for the case when $p = 1$. For the case $p = \infty$, by Theorem 3.2.2, (i) and (iii) is true with $A_\infty = 1$.

Now, by using the Marcinkiewicz interpolation theorem between $L^1 \to L^{1,\infty}$ and $L^\infty \to L^\infty$, we can obtain simultaneously (i) and (iii) for the case $p \in (1, \infty)$. ■

Now, we make some clarifying comments.

Remark 3.2.8. (1) The weak type estimate (ii) is the ***best possible*** (as far as order of magnitude) for the distribution function of Mf, where f is an arbitrary function in $L^1(\mathbb{R}^n)$.

Indeed, we replace $|f(y)|dy$ in the definition of (3.2.2) by a Dirac measure $d\mu$ whose total measure of one is concentrated at the origin. The integral $\int_{B(x,r)} d\mu = 1$ only if the ball $B(x, r)$ contains the origin; otherwise, it will be zero. Thus,

$$M(d\mu)(x) = \sup_{r>0,\, 0\in B(x,r)} \frac{1}{|B(x, r)|} = (V_n|x|^n)^{-1},$$

i.e., it reaches the supremum when $r = |x|$. Obviously, $\|M(d\mu)\|_1 = \infty$. Moreover, the distribution function of $M(d\mu)$ is

$$\begin{aligned}(M(d\mu))_*(\alpha) =&\left| \{x : |M(d\mu)(x)| > \alpha\} \right| = \left| \left\{x : (V_n|x|^n)^{-1} > \alpha\right\} \right|\\ =&\left| \left\{x : V_n|x|^n < \alpha^{-1}\right\} \right| = |B(0, (V_n\alpha)^{-1/n})|\\ =&V_n(V_n\alpha)^{-1} = 1/\alpha,\end{aligned}$$

namely, $\|M(d\mu)\|_{L^{1,\infty}} = 1$. But we can always find a sequence $\{f_m(x)\}$ of positive integrable functions, whose L^1 norm is each 1, and which converges weakly to the measure $d\mu$. So we cannot expect an estimate essentially stronger than the estimate (ii) in Theorem 3.2.7, since, in the limit, a similar stronger version would have to hold for $M(d\mu)(x)$.

(2) It is useful, for certain applications, to observe that
$$A_p = O\left(\frac{1}{p-1}\right), \quad \text{as } p \to 1.$$
(3) It is easier to use $\widetilde{M}$ in proving (ii) than M, one can see the proof that E_α is open.

In contrast with the case $p > 1$, when $p = 1$ the mapping $f \mapsto Mf$ is not bounded on $L^1(\mathbb{R}^n)$. That is,

Theorem 3.2.9.

If $f \in L^1(\mathbb{R}^n)$ is not identically zero, then Mf is never integrable on the whole of $\mathbb{R}^n$, i.e., $Mf \notin L^1(\mathbb{R}^n)$.

Proof. We can choose an N large enough such that
$$\int_{B(0,N)} |f(x)|dx \geqslant \frac{1}{2}\|f\|_1.$$
Then, we take an $x \in \mathbb{R}^n$ such that $|x| \geqslant N$. Let $r = |x| + N$, we have
$$\begin{aligned} Mf(x) \geqslant& \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y)|dy = \frac{1}{V_n(|x|+N)^n}\int_{B(x,r)} |f(y)|dy \\ \geqslant& \frac{1}{V_n(|x|+N)^n}\int_{B(0,N)} |f(y)|dy \geqslant \frac{1}{2V_n(|x|+N)^n}\|f\|_1 \\ \geqslant& \frac{1}{2V_n(2|x|)^n}\|f\|_1. \end{aligned}$$
It follows that for sufficiently large $|x|$, we have
$$Mf(x) \geqslant c|x|^{-n}, \quad c = (V_n 2^{n+1})^{-1}\|f\|_1.$$
This implies that $Mf \notin L^1(\mathbb{R}^n)$. ■

Moreover, even if we limit our consideration to any bounded subset of $\mathbb{R}^n$, then the integrability of Mf holds only if stronger conditions than the integrability of f are required. In fact, we have the following.

Theorem 3.2.10.

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

Proof. By Theorem 1.1.4, it follows that
$$\begin{aligned} \int_E Mf(x)dx =& 2\int_0^\infty |\{x \in E : Mf(x) > 2\alpha\}|d\alpha \\ =& 2\left(\int_0^1 + \int_1^\infty\right) |\{x \in E : Mf(x) > 2\alpha\}|d\alpha \\ \leqslant& 2|E| + 2\int_1^\infty |\{x \in E : Mf(x) > 2\alpha\}|d\alpha. \end{aligned}$$
Decompose f as $f_1 + f_2$, where $f_1 = f\chi_{\{x:|f(x)|>\alpha\}}$ and $f_2 = f - f_1$. Then, by Theorem 3.2.2, it follows that
$$Mf_2(x) \leqslant \|Mf_2\|_\infty \leqslant \|f_2\|_\infty \leqslant \alpha,$$

which yields

$$\{x \in E : Mf(x) > 2\alpha\} \subset \{x \in E : Mf_1(x) > \alpha\}.$$

Hence, by Theorem 3.2.7, we have

$$\int_1^\infty |\{x \in E : Mf(x) > 2\alpha\}| d\alpha \leqslant \int_1^\infty |\{x \in E : Mf_1(x) > \alpha\}| d\alpha$$
$$\leqslant C \int_1^\infty \frac{1}{\alpha} \int_{\{x \in E : |f(x)| > \alpha\}} |f(x)| dx d\alpha \leqslant C \int_E |f(x)| \int_1^{\max(1,|f(x)|)} \frac{d\alpha}{\alpha} dx$$
$$= C \int_E |f(x)| \ln^+ |f(x)| dx.$$

This completes the proof. ■

§3.2.2 Control of other maximal operators

We now study some properties of the Hardy-Littlewood maximal function.

Definition 3.2.11.

Given a function g on $\mathbb{R}^n$ and $\varepsilon > 0$, we denote by g_ε the following function:

$$g_\varepsilon(x) = \varepsilon^{-n} g(\varepsilon^{-1} x). \tag{3.2.8}$$

If g is an integrable function with integral equal to 1, then the family defined by (3.2.8) is an approximate identity. Therefore, convolution with g_ε is an averaging operation. 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

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

Note that the function $k = V_n^{-1} \chi_{B(0,1)}$ has integral equal to 1, and convolving with k_ε is an averaging operation.

Theorem 3.2.12.

Suppose that the ***least decreasing radial majorant*** of φ is integrable, i.e., let $\psi(x) = \sup_{|y| \geqslant |x|} |\varphi(y)|$, and $\psi \in L^1(\mathbb{R}^n)$. Then for $f \in L^1_{\text{loc}}(\mathbb{R}^n)$,

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

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

$$(f * \psi_\varepsilon)(x) \leqslant AMf(x), \tag{3.2.9}$$

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

Since (3.2.9) is clearly translation invariant w.r.t f and also dilation invariant w.r.t. ψ and the maximal function, it suffices to show that

$$(f * \psi)(0) \leqslant AMf(0). \tag{3.2.10}$$

In proving (3.2.10), we may clearly assume that $Mf(0) < \infty$. Let us write $\lambda(r) = \int_{S^{n-1}} f(rx')d\sigma(x')$, and $\Lambda(r) = \int_{|x|\leqslant r} f(x)dx$, so

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

We have

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

Since $\Lambda(r) = \int_{|x|\leqslant r} f(x)dx \leqslant V_n r^n Mf(0)$, and the fact $r^n\psi(r) \to 0$ as $r \to 0$ or $r \to \infty$, we have

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

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

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

where two of the integrals are of Lebesgue-Stieltjes type, since $\psi(r)$ is decreasing which implies $\psi'(r) \leqslant 0$, and $nV_n = \omega_{n-1}$. This proves (3.2.10) and then (3.2.9). ■

§ 3.2.3 Applications to differentiation theory

We continue this section by obtaining some applications of the boundedness of the Hardy-Littlewood maximal function in differentiation theory.

We now show that the weak type $(1, 1)$ property of the Hardy-Littlewood maximal function implies almost everywhere convergence for a variety of families of functions. We deduce this from the more general fact that a certain weak type property for the supremum of a family of linear operators implies almost everywhere convergence.

Let (X, μ) and (Y, ν) be measure spaces and let $1 \leqslant p \leqslant \infty$, $1 \leqslant q < \infty$. Suppose that D is a dense subspace of $L^p(X, \mu)$. This means that for all $f \in L^p$ and all $\delta > 0$, there exists a $g \in D$ such that $\|f - g\|_p < \delta$. Suppose that for every $\varepsilon > 0$, T_ε is a linear operator that maps $L^p(X, \mu)$ into a subspace of measurable functions, which are defined everywhere on Y. For $y \in Y$, define a sublinear operator

$$T_* f(y) = \sup_{\varepsilon>0} |T_\varepsilon f(y)| \tag{3.2.11}$$

and assume that $T_* f$ is ν-measurable for any $f \in L^p(X, \mu)$. We have the following.

Theorem 3.2.13.

Let $p, q \in [1, \infty)$, T_ε and T_* be as previously stated. Suppose that for some $B > 0$ and all $f \in L^p(X, \mu)$, we have

$$\|T_* f\|_{L^{q,\infty}} \leqslant B\|f\|_p \tag{3.2.12}$$

and that for all $f \in D$,

$$\lim_{\varepsilon \to 0} T_\varepsilon f = Tf \tag{3.2.13}$$

exists and is finite ν-a.e. (and defines a linear operator on D). Then, for all $f \in L^p(X, \mu)$, the limit (3.2.13) exists and is finite ν-a.e., and defines a linear operator T on $L^p(X, \mu)$ (uniquely extending T defined on D) that satisfies

$$\|Tf\|_{L^{q,\infty}} \leqslant B\|f\|_p \tag{3.2.14}$$

for all functions $f \in L^p(X, \mu)$.

Proof. Given $f \in L^p$, we define the oscillation of f:

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

We would like to show that for all $f \in L^p$ and $\delta > 0$,

$$\nu(\{y \in Y : O_f(y) > \delta\}) = 0. \tag{3.2.15}$$

Once (3.2.15) is established, given $f \in L^p(X, \mu)$, we obtain that $O_f(y) = 0$ for ν-almost all y, which implies that $T_\varepsilon f(y)$ is Cauchy for ν-almost all y, and it therefore converges ν-a.e. to some $Tf(y)$ as $\varepsilon \to 0$. The operator T defined in this way on $L^p(X, \mu)$ is linear and extends T defined on D.

To approximate O_f, we use density. Given $\eta > 0$, find a function $g \in D$ such that $\|f - g\|_p < \eta$. Since $T_\varepsilon g \to Tg$ ν-a.e., it follows that $O_g = 0$ ν-a.e. Using this fact and the linearity of the T_ε's, we conclude that

$$O_f(y) \leqslant O_g(y) + O_{f-g}(y) = O_{f-g}(y) \quad \nu - \text{a.e.}$$

Now for any $\delta > 0$, we have by (3.2.12)

$$\begin{aligned}
\nu(\{y \in Y : O_f(y) > \delta\}) \leqslant & \nu(\{y \in Y : O_{f-g}(y) > \delta\}) \\
\leqslant & \nu(\{y \in Y : 2T_*(f-g)(y) > \delta\}) \\
\leqslant & (2\|T_*(f-g)\|_{L^{q,\infty}}/\delta)^q \\
\leqslant & (2B\|f-g\|_p/\delta)^q \\
\leqslant & (2B\eta/\delta)^q,
\end{aligned}$$

due to

$$\begin{aligned}
O_{f-g}(y) = & \limsup_{\varepsilon \to 0} \limsup_{\theta \to 0} |T_\varepsilon(f-g)(y) - T_\theta(f-g)(y)| \\
\leqslant & 2 \sup_{\varepsilon} |T_\varepsilon(f-g)(y)| = 2T_*(f-g)(y).
\end{aligned}$$

Letting $\eta \to 0$, we deduce (3.2.15). We conclude that $T_\varepsilon f$ is a Cauchy sequence, and hence it converges ν-a.e. to some Tf. Since $|Tf| \leqslant T_* f$, the conclusion (3.2.14) follows easily. ∎

As a corollary of Theorem 3.2.7 or 3.2.13, we have the differentiability almost everywhere of the integral, expressed in (3.2.1).

Theorem 3.2.14: Lebesgue differentiation theorem

If $f \in L^p(\mathbb{R}^n)$, $p \in [1, \infty]$, or more generally if f is locally integrable (i.e., $f \in L^1_{\text{loc}}(\mathbb{R}^n)$), then

$$\lim_{r \to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f(y)dy = f(x), \quad \text{for a.e. } x. \tag{3.2.16}$$

Proof. We first consider the case $p = 1$. It suffices to show that for each $\alpha > 0$, the set

$$E_\alpha = \left\{ x : \limsup_{r \to 0} \left| \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f(y)dy - f(x) \right| > 2\alpha \right\}$$

has measure zero, because this assertion then guarantees that the set $E = \bigcup_{k=1}^\infty E_{1/k}$ has measure zero, and the limit in (3.2.16) holds at all points of E^c.

Fix α, since all continuous functions of compact support (i.e., $\mathcal{C}_c(\mathbb{R}^n)$) are dense in $L^1(\mathbb{R}^n)$, for each $\varepsilon > 0$ we may select a continuous function g of compact support with $\|f - g\|_1 < \varepsilon$. As we remarked earlier, the continuity of g implies that

$$\lim_{r \to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} g(y)dy = g(x), \quad \text{for all } x.$$

Since we may write the difference $\frac{1}{\mu(B(x,r))} \int_{B(x,r)} f(y)dy - f(x)$ as

$$\begin{aligned} &\frac{1}{\mu(B(x,r))} \int_{B(x,r)} (f(y) - g(y))dy \\ &+ \frac{1}{\mu(B(x,r))} \int_{B(x,r)} g(y)dy - g(x) + g(x) - f(x), \end{aligned}$$

we find that

$$\limsup_{r \to 0} \left| \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f(y)dy - f(x) \right| \leqslant M(f-g)(x) + |g(x) - f(x)|.$$

Consequently, if

$$F_\alpha = \{x : M(f-g)(x) > \alpha\} \quad \text{and} \quad G_\alpha = \{x : |f(x) - g(x)| > \alpha\},$$

then $E_\alpha \subset F_\alpha \cup G_\alpha$, because if u_1 and u_2 are positive, then $u_1 + u_2 > 2\alpha$ only if $u_i > \alpha$ for at least one u_i.

On the one hand, Tchebychev's inequality yields

$$\mu(G_\alpha) \leqslant \frac{1}{\alpha} \|f - g\|_1,$$

and on the other hand, the weak type estimate for the maximal function gives

$$\mu(F_\alpha) \leqslant \frac{3^n}{\alpha} \|f - g\|_1.$$

Since the function g was selected so that $\|f - g\|_1 < \varepsilon$, we get

$$\mu(E_\alpha) \leqslant \frac{3^n}{\alpha}\varepsilon + \frac{1}{\alpha}\varepsilon = \frac{3^n + 1}{\alpha}\varepsilon.$$

Since ε is arbitrary, we must have $\mu(E_\alpha) = 0$, and the proof for $p = 1$ is completed.

Indeed, the limit in the theorem is taken over balls that shrink to the point x, so the behavior of f far from x is irrelevant. Thus, we expect the result to remain valid if we simply assume integrability of f on every ball. Clearly, the conclusion holds under the weaker assumption that f is locally integrable.

For the remained cases $p \in (1, \infty]$, we have by Hölder inequality, for any ball B,

$$\int_B |f(x)|dx \leqslant \|f\|_{L^p(B)} \|1\|_{L^{p'}(B)} \leqslant \mu(B)^{1/p'} \|f\|_p.$$

Thus, $f \in L^1_{loc}(\mathbb{R}^n)$ and then the conclusion is valid for $p \in (1, \infty]$. Therefore, we complete the proof of the theorem. ■

By the Lebesgue differentiation theorem, we have

Corollary 3.2.15.

Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then

$$|f(x)| \leqslant Mf(x), \quad a.e.\ x \in \mathbb{R}^n.$$

Combining with the maximal function theorem (i.e., Theorem 3.2.7), we get

Corollary 3.2.16.

If $f \in L^p(\mathbb{R}^n)$, $p \in (1, \infty]$, then we have

$$\|f\|_p \leqslant \|Mf\|_p \leqslant A_p \|f\|_p.$$

Corollary 3.2.17.

Suppose that the ***least decreasing radial majorant*** of φ is integrable, and $\int_{\mathbb{R}^n} \varphi(x)dx = 1$. Then $\lim_{\varepsilon\to 0}(f * \varphi_\varepsilon)(x) = f(x)$ a.e. for all $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$.

Proof. We can verify that if $f_1 \in \mathcal{C}_c$, then $(f_1 * \varphi_\varepsilon)(x) \to f_1(x)$ uniformly as $\varepsilon \to 0$ (cf. Theorem 2.1.15). Next we can deal with the case $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, by writing $f = f_1 + f_2$ with f_1 as described and with $\|f_2\|_p$ small. The argument then follows closely that given in the proof of Theorem 3.2.14 (the Lebesgue differentiation theorem). Thus, we get that $\lim_{\varepsilon\to 0} f * \varphi_\varepsilon(x)$ exists almost everywhere and equals $f(x)$. ■

§ 3.2.4 An application to Sobolev's inequality

As an application, we prove the (Gagliardo-Nirenberg-) Sobolev inequality by using the maximal function theorem for the case $1 < p < n$. We note that the inequality also holds for the case $p = 1$ and one can see [Eva10, p.279-281] for the proof.

Theorem 3.2.18: (Gagliardo-Nirenberg-) Sobolev inequality

Let $p \in (1, n)$ and its Sobolev conjugate $p^* = np/(n-p)$. Then for $f \in \mathscr{D}(\mathbb{R}^n)$, we have

$$\|f\|_{p^*} \leqslant C\|\nabla f\|_p,$$

where C depends only on n and p.

Proof. Since $f \in \mathscr{D}(\mathbb{R}^n)$, we have

$$f(x) = -\int_0^\infty \frac{\partial}{\partial r}(f(x + rz))dr,$$

where $z \in S^{n-1}$. Integrating this over the whole unit sphere surface S^{n-1} yields

$$\begin{aligned}\omega_{n-1} f(x) = \int_{S^{n-1}} f(x) d\sigma(z) &= -\int_{S^{n-1}} \int_0^\infty \frac{\partial}{\partial r}(f(x+rz)) dr d\sigma(z) \\ &= -\int_{S^{n-1}} \int_0^\infty \nabla f(x+rz) \cdot z dr d\sigma(z) \\ &= -\int_0^\infty \int_{S^{n-1}} \nabla f(x+rz) \cdot z d\sigma(z) dr.\end{aligned}$$

Changing variables $y = x + rz$, $d\sigma(z) = r^{-(n-1)} d\sigma(y)$, $z = (y-x)/|y-x|$ and $r = |y-x|$, we get

$$\begin{aligned}\omega_{n-1} f(x) &= -\int_0^\infty \int_{\partial B(x,r)} \nabla f(y) \cdot \frac{y-x}{|y-x|^n} d\sigma(y) dr \\ &= -\int_{\mathbb{R}^n} \nabla f(y) \cdot \frac{y-x}{|y-x|^n} dy,\end{aligned}$$

which implies that

$$|f(x)| \leqslant \frac{1}{\omega_{n-1}} \int_{\mathbb{R}^n} \frac{|\nabla f(y)|}{|y-x|^{n-1}} dy.$$

We split this integral into two parts as $\int_{\mathbb{R}^n} = \int_{B(x,r)} + \int_{\mathbb{R}^n \setminus B(x,r)}$. For the first part, we have

$$\begin{aligned}&\frac{1}{\omega_{n-1}} \int_{B(x,r)} \frac{|\nabla f(y)|}{|x-y|^{n-1}} dy \\ =&\frac{1}{\omega_{n-1}} \sum_{k=0}^\infty \int_{B(x,2^{-k}r)\setminus B(x,2^{-k-1}r)} \frac{|\nabla f(y)|}{|x-y|^{n-1}} dy \\ \leqslant&\frac{1}{\omega_{n-1}} \sum_{k=0}^\infty \int_{B(x,2^{-k}r)\setminus B(x,2^{-k-1}r)} \frac{|\nabla f(y)|}{(2^{-k-1}r)^{n-1}} dy \\ \leqslant&\sum_{k=0}^\infty \frac{2^{-k}r}{nV_n 2^{-k}r} \int_{B(x,2^{-k}r)} 2^{n-1} \frac{|\nabla f(y)|}{(2^{-k}r)^{n-1}} dy \\ \leqslant&\frac{1}{n} \sum_{k=0}^\infty 2^{-k+n-1} r \frac{1}{\mu(B(x,2^{-k}r))} \int_{B(x,2^{-k}r)} |\nabla f(y)| dy \\ \leqslant&\frac{2^{n-1}}{n} r M(\nabla f)(x) \sum_{k=0}^\infty 2^{-k} = \frac{2^n}{n} r M(\nabla f)(x).\end{aligned}$$

For the second part, by Hölder inequality, we get for $1 < p < n$

$$\begin{aligned}&\int_{\mathbb{R}^n \setminus B(x,r)} \frac{|\nabla f(y)|}{|x-y|^{n-1}} dy \\ \leqslant&\left(\int_{\mathbb{R}^n \setminus B(x,r)} |\nabla f(y)|^p dy\right)^{1/p} \left(\int_{\mathbb{R}^n \setminus B(x,r)} |x-y|^{(1-n)p'} dy\right)^{1/p'} \\ \leqslant&\left(\omega_{n-1} \int_r^\infty \rho^{(1-n)p'} \rho^{n-1} d\rho\right)^{1/p'} \|\nabla f\|_p \\ =&\left(\frac{(p-1)\omega_{n-1}}{n-p}\right)^{1/p'} r^{1-n/p} \|\nabla f\|_p.\end{aligned}$$

Choose $r = \frac{(p-1)^{(p-1)/n}}{(n-p)^{(p-1)/n} \omega_{n-1}^{1/n} 2^p} \left(\frac{n\|\nabla f\|_p}{M(\nabla f)(x)}\right)^{p/n}$ satisfying

$$\frac{2^n}{n} r M(\nabla f)(x) = \frac{1}{\omega_{n-1}} \left(\frac{(p-1)\omega_{n-1}}{n-p}\right)^{1/p'} r^{1-n/p} \|\nabla f\|_p,$$

then we get

$$|f(x)| \leqslant C\|\nabla f\|_p^{p/n}(M(\nabla f)(x))^{1-p/n}.$$

Thus, by part (iii) in Theorem 3.2.7, we obtain for $1 < p < n$

$$\begin{aligned}\|f\|_{p^*} \leqslant & C\|\nabla f\|_p^{p/n}\|M(\nabla f)\|_{p^*(1-p/n)}^{1-p/n} \\ = & C\|\nabla f\|_p^{p/n}\|M(\nabla f)\|_p^{1-p/n} \leqslant C\|\nabla f\|_p.\end{aligned}$$

This completes the proof. ■

§ 3.3 Calderón-Zygmund decomposition

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

Theorem 3.3.1: Calderón-Zygmund decomposition of $\mathbb{R}^n$

Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. Then there exists a decomposition of $\mathbb{R}^n$ such that
(i) $\mathbb{R}^n = F \cup \Omega$, $F \cap \Omega = \varnothing$.
(ii) $|f(x)| \leqslant \alpha$ for a.e. $x \in F$.
(iii) Ω is the union of cubes, $\Omega = \bigcup_k Q_k$, whose interiors are disjoint and edges parallel to the coordinate axes, and such that for each Q_k

$$\alpha < \frac{1}{\mu(Q_k)} \int_{Q_k} |f(x)| dx \leqslant 2^n \alpha. \tag{3.3.1}$$

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

$$\frac{1}{\mu(Q_k^{(0)})} \int_{Q_k^{(0)}} |f(x)| dx \leqslant \alpha, \tag{3.3.2}$$

since $f \in L^1$.

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

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

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

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

in view of (3.3.2) where $Q_k^{(1)}$ is split from $Q_{\widetilde{k}}^{(0)}$, 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

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

Thus, $|f(x)| \leqslant \alpha$ a.e. $x \in F = \Omega^c$ by a variation of the Lebesgue differentiation theorem. Thus, we complete the proof. ■

We now state an immediate corollary.

Corollary 3.3.2.

Suppose f, α, F, Ω and Q_k have the same meaning as in Theorem 3.3.1. Then there exists two constants A and B (depending only on the dimension n), such that (i) and (ii) of Theorem 3.3.1 hold and

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

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

Proof. In fact, by (3.3.1) we can take $B = 2^n$, and also because of (3.3.1)

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

This proves the corollary with $A = 1$ and $B = 2^n$. ■

It is possible however to give another proof of this corollary without using Theorem 3.3.1 from which it was deduced, but by using the maximal function theorem (Theorem 3.2.7) and also the theorem about the decomposition of an arbitrary open set as a union of disjoint cubes. This more indirect method of proof has the advantage of ***clarifying the roles of the sets F and Ω into which $\mathbb{R}^n$ was divided***.

Another proof of Corollary 3.3.2. We know that in F, $|f(x)| \leqslant \alpha$, but this fact does not determine F. The set F is however determined, in effect, by the fact that the maximal function satisfies $Mf(x) \leqslant \alpha$ on it. So we choose $F = \{x : Mf(x) \leqslant \alpha\}$ and $\Omega = E_\alpha = \{x : Mf(x) > \alpha\}$. Then by Theorem 3.2.7, part (ii) we know that $\mu(\Omega) \leqslant \frac{3^n}{\alpha}\|f\|_1$. Thus, we can take $A = 3^n$.

Since by definition F is closed, we can choose cubes Q_k according to Theorem 3.1.2, such that $\Omega = \bigcup_k Q_k$, and whose diameters are approximately proportional to their distances from F. Let Q_k then be one of these cubes, and p_k a point of F such that

$$\text{dist}\,(F, Q_k) = \text{dist}\,(p_k, Q_k).$$

Let B_k be the smallest ball whose center is p_k and which contains the interior of Q_k. Let us set

$$\gamma_k = \frac{\mu(B_k)}{\mu(Q_k)}.$$

We have, because $p_k \in \{x : Mf(x) \leqslant \alpha\}$, that

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

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

The elementary geometry and the inequality (iii) of Theorem 3.1.2 then show that

$$\begin{aligned}\text{radius}(B_k) \leqslant &\text{dist}\,(p_k, Q_k) + \text{diam}\,(Q_k) = \text{dist}\,(F, Q_k) + \text{diam}\,(Q_k)\\ \leqslant &(c_2 + 1)\,\text{diam}\,(Q_k),\end{aligned}$$

and so

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

since $\mu(Q_k) = (\operatorname{diam}(Q_k)/\sqrt{n})^n$. Thus, $\gamma_k \leqslant V_n(c_2+1)^n n^{n/2}$ for all k. Thus, we complete the proof with $A = 3^n$ and $B = V_n(c_2+1)^n n^{n/2}$. ■

Remark 3.3.3. Theorem 3.3.1 may be used to give another proof of the fundamental inequality for the maximal function in part (ii) of Theorem 3.2.7. (See [Ste70, §5.1, p.22–23] for more details.)

The Calderón-Zygmund decomposition is a key step in the real-variable analysis of singular integrals. The idea behind this decomposition is that it is often useful to split an arbitrary integrable function into its "small" and "large" parts, and then use different techniques to analyze each part.

The scheme is roughly as follows. Given a function f and an altitude α, we write $f = g + b$, where g is called the good function of the decomposition since it is both integrable and bounded; hence the letter g. The function b is called the bad function since it contains the singular part of f (hence the letter b), but it is carefully chosen to have mean value zero. To obtain the decomposition $f = g + b$, one might be tempted to "cut" f at the height α; however, this is not what works. Instead, one bases the decomposition on the set where the maximal function of f has height α.

Indeed, the Calderón-Zygmund decomposition on $\mathbb{R}^n$ may be used to deduce the Calderón-Zygmund decomposition on functions. The latter is a very important tool in harmonic analysis.

Theorem 3.3.4: Calderón-Zygmund decomposition for functions

Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. Then there exist functions g and b on $\mathbb{R}^n$ such that $f = g + b$ and

(i) $\|g\|_1 \leqslant \|f\|_1$ and $\|g\|_\infty \leqslant 2^n\alpha$.

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

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

Proof. Applying Corollary 3.3.2 (with $A = 1$ and $B = 2^n$), we have

1) $\mathbb{R}^n = F \cup \Omega$, $F \cap \Omega = \varnothing$;

2) $|f(x)| \leqslant \alpha$, a.e. $x \in F$;

3) $\Omega = \bigcup_{j=1}^\infty Q_j$, with the interiors of the Q_j mutually disjoint;

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

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

Now define

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

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

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

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

Next, we need to obtain the estimates on g. Write $\mathbb{R}^n = \cup_j Q_j \cup F$, where F is the closed set obtained by Corollary 3.3.2. Since $b = 0$ on F and $f - b_j = \frac{1}{\mu(Q_j)}\int_{Q_j} f(x)dx$ on Q_j, we have

$$g = \begin{cases} f, & \text{on } F, \\ \dfrac{1}{\mu(Q_j)}\displaystyle\int_{Q_j} f(x)dx, & \text{on } Q_j. \end{cases} \tag{3.3.3}$$

On the cube Q_j, g is equal to the constant $\frac{1}{\mu(Q_j)}\int_{Q_j} f(x)dx$, and this is bounded by $2^n\alpha$ by 4). Then by 2), we can get $\|g\|_\infty \leqslant 2^n\alpha$. Finally, it follows from (3.3.3) that $\|g\|_1 \leqslant \|f\|_1$. This completes the proof of (i) and then of the theorem. ■

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

Theorem 3.3.5.

For $p \in (1, \infty)$, there exists a constant $C = C_{n,p}$ such that, for any non-negative 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

$$\int_{\mathbb{R}^n} (Mf(x))^p\varphi(x)dx \leqslant C\int_{\mathbb{R}^n} |f(x)|^p M\varphi(x)dx. \tag{3.3.4}$$

Proof. Except when $M\varphi(x) = \infty$ a.e., in which case (3.3.4) 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 in order to get (3.3.4), it suffices to prove that M is both of 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 type $(L^\infty(\sigma), L^\infty(\nu))$. In fact, 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 $\mu(\{x \in \mathbb{R}^n : |f(x)| > \alpha\}) = 0$, 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 $\mu(\{x : Mf(x) > \alpha\}) = 0$ 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, $\|Mf\|_{L^\infty(\nu)} \leqslant \|f\|_{L^\infty(\sigma)}$.

Before proving that M is also of weak type $(L^1(\sigma), L^1(\nu))$, we give the following lemma.

Lemma 3.3.6.

Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. If the sequence $\{Q_k\}$ of cubes is chosen from the Calderón-Zygmund decomposition of $\mathbb{R}^n$ for f and $\alpha > 0$, then

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

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

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

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

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

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

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

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

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

We complete the proof of the lemma. ∎

Let us return to the proof of 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 $f \in L^1(\sigma)$

$$\begin{aligned}
\int_{\{x\in\mathbb{R}^n : Mf(x)>\alpha\}} \varphi(x)dx =& \nu(\{x \in \mathbb{R}^n : Mf(x) > \alpha\}) \\
\leqslant & \frac{C}{\alpha} \int_{\mathbb{R}^n} |f(x)|M\varphi(x)dx.
\end{aligned} \tag{3.3.5}$$

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 $\ell = 1, 2, \cdots$. Moreover, $\lim_{\ell\to\infty} f_\ell(x) = |f(x)|$ and

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

By the point-wise equivalence of M and M', there exists $c_n > 0$ such that $Mf(x) \leqslant c_n M'f(x)$ for all $x \in \mathbb{R}^n$. Applying the Calderón-Zygmund decomposition on $\mathbb{R}^n$ for f and $\alpha' = \alpha/(c_n 7^n)$, we get a sequence $\{Q_k\}$ of cubes satisfying

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

By Lemma 3.3.6 and the point-wise equivalence of M and M'', we have that

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

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

References

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[FS71] C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[Ste93] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

IV

Singular Integrals of Convolution Type

§ 4.1 Poisson kernel and Hilbert transform

§ 4.1.1 Poisson kernel and the conjugate

We shall now introduce a notation that will be indispensable in much of our further work. Indeed, we have shown some properties of Poisson kernel in Chapter 2. The setting for the application of this theory will be as follows. We shall think of $\mathbb{R}^n$ as the boundary hyperplane of the $(n+1)$ dimensional upper-half space $\mathbb{R}^{n+1}$. In coordinate notation,

$$\mathbb{R}^{n+1}_+ = \{(x,y) : x \in \mathbb{R}^n, y > 0\}.$$

We shall consider the ***Poisson integral*** of a function f given on $\mathbb{R}^n$. This Poisson integral is effectively the solution to the Dirichlet Problem for $\mathbb{R}^{n+1}_+$: find a ***harmonic function*** $u(x,y)$ on $\mathbb{R}^{n+1}_+$, whose boundary values on $\mathbb{R}^n$ (in the appropriate sense) are $f(x)$, that is

$$\begin{cases} \Delta_{x,y} u(x,y) = 0, & (x,y) \in \mathbb{R}^{n+1}_+, \\ u(x,0) = f, & x \in \mathbb{R}^n. \end{cases} \tag{4.1.1}$$

The formal solution of this problem can be given neatly in the context of the L^2 theory.

In fact, let $f \in L^2(\mathbb{R}^n)$, and consider

$$u(x,y) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{\omega i\xi\cdot x} e^{-|\omega\xi|y} \widehat{f}(\xi)d\xi, \quad y > 0. \tag{4.1.2}$$

This integral converges absolutely (cf. Theorem 2.1.15), because $\widehat{f} \in L^2(\mathbb{R}^n)$, and $e^{-|\omega\xi|y}$ is rapidly decreasing in $|\xi|$ for $y > 0$. For the same reason, the integral above may be differentiated w.r.t. x and y any number of times by carrying out the operation under the sign of integration. This gives

$$\Delta_{x,y} u = \frac{\partial^2 u}{\partial y^2} + \sum_{k=1}^n \frac{\partial^2 u}{\partial x_k^2} = 0,$$

because the factor $e^{\omega i\xi\cdot x} e^{-|\omega\xi|y}$ satisfies this property for each fixed ξ. Thus, $u(x,y)$ is a harmonic function on $\mathbb{R}^{n+1}_+$.

By Theorem 2.1.15, we get that $u(x,y) \to f(x)$ in $L^2(\mathbb{R}^n)$ norm, as $y \to 0$. That is, $u(x,y)$ satisfies the boundary condition and so $u(x,y)$ structured above is a solution for the above Dirichlet problem.

This solution of the problem can also be written without explicit use of the Fourier transform. For this purpose, we recall the ***Poisson kernel*** $P_y(x) := P(x,y)$ by

$$P_y(x) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega i\xi\cdot x} e^{-|\omega\xi|y} d\xi = \left(\frac{|\omega|}{2\pi}\right)^{n/2} (\mathscr{F}^{-1} e^{-|\omega\xi|y})(x), \quad y > 0. \tag{4.1.3}$$

Then the function $u(x,y)$ obtained above can be written as a convolution

$$u(x,y) = \int_{\mathbb{R}^n} P_y(z) f(x-z) dz, \tag{4.1.4}$$

as the same as in Theorem 2.1.15. We shall say that u is the ***Poisson integral*** of f.

For convenience, we recall (2.1.14) and (2.1.12) as follows.

Proposition 4.1.1.

The Poisson kernel has the following explicit expression:

$$P_y(x) = \frac{c_n y}{(|x|^2 + y^2)^{\frac{n+1}{2}}}, \quad c_n = \frac{\Gamma((n+1)/2)}{\pi^{\frac{n+1}{2}}}. \tag{4.1.5}$$

Remark 4.1.2. We list the properties of the Poisson kernel that are now more or less evident:

(i) The expression in (4.1.5) is independent of the definition of the Fourier transform, and $P_y(x) > 0$ for $y > 0$.

(ii) $\int_{\mathbb{R}^n} P_y(x)dx = 1$, $y > 0$ by Lemma 2.1.14; more generally, $\widehat{P_y}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} e^{-|\omega\xi|y}$ by Corollary 2.1.23.

(iii) $P_y(x)$ is homogeneous of degree $-n$ w.r.t. (x,y); and $P_y(x) = y^{-n} P_1(x/y)$, $y > 0$.

(iv) $P_y(x)$ is a decreasing function of $|x|$, and $P_y \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$. Indeed, by changes of variables, we have for $1 \leqslant p < \infty$

$$\|P_y\|_p^p = c_n^p \int_{\mathbb{R}^n} \left(\frac{y}{(|x|^2 + y^2)^{(n+1)/2}}\right)^p dx$$

$$\overset{x=yz}{=\!=}c_n^p y^{-n(p-1)}\int_{\mathbb{R}^n}\frac{1}{(1+|z|^2)^{p(n+1)/2}}dz$$
$$\overset{z=rz'}{=\!=}c_n^p y^{-n(p-1)}\omega_{n-1}\int_0^\infty\frac{1}{(1+r^2)^{p(n+1)/2}}r^{n-1}dr$$
$$\overset{r=\tan\theta}{=\!=}c_n^p y^{-n(p-1)}\omega_{n-1}\int_0^{\pi/2}\frac{1}{(\sec\theta)^{p(n+1)}}\tan^{n-1}\theta\sec^2\theta d\theta$$
$$=c_n^p y^{-n(p-1)}\omega_{n-1}\int_0^{\pi/2}\sin^{n-1}\theta\cos^{(p-1)(n+1)}\theta d\theta$$
$$=\frac{c_n^p\omega_{n-1}}{2}B\left(\frac{p(n+1)-n}{2},\frac{n}{2}\right)y^{-n(p-1)},$$

where we recall that the Beta function

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

converges for $\Re\alpha,\Re\beta>0$. Here, it is clear that $p(n+1)-n>0$ for $p\in[1,\infty)$ and thus the Beta function converges. Therefore, we have for $p\in[1,\infty)$

$$\|P_y\|_p=c_n\left[\frac{\omega_{n-1}}{2}B\left(\frac{p(n+1)-n}{2},\frac{n}{2}\right)\right]^{1/p}y^{-n/p'}.$$

For $p=\infty$, it is clear that $\|P_y(x)\|_\infty=c_ny^{-n}$.

(v) Suppose $f\in L^p(\mathbb{R}^n)$, $1\leqslant p\leqslant\infty$, then its Poisson integral u, given by (4.1.4), is harmonic in $\mathbb{R}^{n+1}_+$. This is a simple consequence of the fact that $P_y(x)$ is harmonic in $\mathbb{R}^{n+1}_+$ which is immediately derived from (4.1.3).

(vi) We have the "semi-group property" $P_{y_1}*P_{y_2}=P_{y_1+y_2}$ if $y_1,y_2>0$ in view of Corollary 2.1.24.

The boundary behavior of Poisson integrals is already described to a significant extension by the following theorem.

Theorem 4.1.3.

Suppose $f\in L^p(\mathbb{R}^n)$, $1\leqslant p\leqslant\infty$, and let $u(x,y)$ be its Poisson integral. Then

(a) $\sup_{y>0}|u(x,y)|\leqslant Mf(x)$, where Mf is the maximal function.

(b) $\lim_{y\to0}u(x,y)=f(x)$, for almost every x.

(c) If $1\leqslant p<\infty$, $u(x,y)$ converges to $f(x)$ in $L^p(\mathbb{R}^n)$ norm, as $y\to0$.

Proof. We can prove it by applying Theorem 3.2.12 directly, because of properties (i)–(iv) of the Poisson kernel in the case $\varphi(x)=\psi(x)=P_1(x)$. ■

Now, we give the definition of harmonic conjugate functions as follows.

Definition 4.1.4.

The ***harmonic conjugate*** to a given function $u(x,y)$ is a function $v(x,y)$ such that

$$f(x,y)=u(x,y)+iv(x,y)$$

is analytic, i.e., satisfies the ***Cauchy-Riemann equations***

$$u_x = v_y, \quad u_y = -v_x,$$

where $u_x \equiv \partial u/\partial x$, $u_y \equiv \partial u/\partial y$. It is given by

$$v(x,y) = \int_{(x_0,y_0)}^{(x,y)} u_x dy - u_y dx + C,$$

along any path connecting (x_0, y_0) and (x, y) in the domain, where C is a constant of integration.

Given a function f in $\mathscr{S}(\mathbb{R})$, its harmonic extension to the upper half-plane is given by $u(x,y) = P_y * f(x)$, where P_y is the Poisson kernel. We can also write, in view of (4.1.2),

$$\begin{aligned} u(z) =& u(x,y) = \left(\frac{|\omega|}{2\pi}\right)^{1/2} \int_{\mathbb{R}} e^{\omega i\xi\cdot x} e^{-|\omega\xi| y} \widehat{f}(\xi) d\xi \\ =& \left(\frac{|\omega|}{2\pi}\right)^{1/2} \left[\int_0^\infty e^{\omega i\xi\cdot x} e^{-|\omega|\xi y} \widehat{f}(\xi) d\xi + \int_{-\infty}^0 e^{\omega i\xi\cdot x} e^{|\omega|\xi y} \widehat{f}(\xi) d\xi\right] \\ =& \left(\frac{|\omega|}{2\pi}\right)^{1/2} \left[\int_0^\infty e^{\omega i\xi\cdot(x+i\,\mathrm{sgn}\,(\omega) y)} \widehat{f}(\xi) d\xi + \int_{-\infty}^0 e^{\omega i\xi\cdot(x-i\,\mathrm{sgn}\,(\omega) y)} \widehat{f}(\xi) d\xi\right], \end{aligned}$$

where $z = x + iy$. If we now define

$$\begin{aligned} i\,\mathrm{sgn}\,(\omega) v(z) = \left(\frac{|\omega|}{2\pi}\right)^{1/2} \Big[& \int_0^\infty e^{\omega i\xi\cdot(x+i\,\mathrm{sgn}\,(\omega) y)} \widehat{f}(\xi) d\xi \\ & - \int_{-\infty}^0 e^{\omega i\xi\cdot(x-i\,\mathrm{sgn}\,(\omega) y)} \widehat{f}(\xi) d\xi\Big], \end{aligned}$$

then v is also harmonic in $\mathbb{R}^2_+$ and both u and v are real if f is. Furthermore, $u + iv$ is analytic since it satisfies the Cauchy-Riemann equations $u_x = v_y = \omega i\xi u(z)$ and $u_y = -v_x = -\omega i\xi v(z)$, so v is the harmonic conjugate of u.

Clearly, v can also be written as, by Proposition 2.4.18,

$$\begin{aligned} v(z) =& -i\,\mathrm{sgn}\,(\omega) \left(\frac{|\omega|}{2\pi}\right)^{1/2} \int_{\mathbb{R}} \mathrm{sgn}\,(\xi) e^{\omega i\xi\cdot x} e^{-|\omega\xi| y} \widehat{f}(\xi) d\xi \\ =& -i\,\mathrm{sgn}\,(\omega) \mathscr{F}^{-1}(\,\mathrm{sgn}\,(\xi) e^{-|\omega\xi| y} \widehat{f}(\xi))(x) \\ =& -i\,\mathrm{sgn}\,(\omega) \left(\frac{|\omega|}{2\pi}\right)^{1/2} [\mathscr{F}^{-1}(\,\mathrm{sgn}\,(\xi) e^{-|\omega\xi| y}) * f](x), \end{aligned}$$

which is equivalent to

$$v(x,y) = Q_y * f(x), \tag{4.1.6}$$

where

$$\widehat{Q_y}(\xi) = -i\,\mathrm{sgn}\,(\omega) \left(\frac{|\omega|}{2\pi}\right)^{1/2} \mathrm{sgn}\,(\xi) e^{-|\omega\xi| y}. \tag{4.1.7}$$

Now we invert the Fourier transform, we get, by a change of variables and integration by parts,

$$\begin{aligned} Q_y(x) =& -i\,\mathrm{sgn}\,(\omega) \frac{|\omega|}{2\pi} \int_{\mathbb{R}} e^{\omega ix\cdot\xi} \,\mathrm{sgn}\,(\xi) e^{-|\omega\xi| y} d\xi \\ =& -i\,\mathrm{sgn}\,(\omega) \frac{|\omega|}{2\pi} \left[\int_0^\infty e^{\omega ix\cdot\xi} e^{-|\omega|\xi y} d\xi - \int_{-\infty}^0 e^{\omega ix\cdot\xi} e^{|\omega|\xi y} d\xi\right] \end{aligned}$$

$$
\begin{aligned}
&= - i \operatorname{sgn}(\omega) \frac{|\omega|}{2\pi} \left[\int_0^\infty e^{\omega i x \cdot \xi} e^{-|\omega|\xi y} d\xi - \int_0^\infty e^{-\omega i x \cdot \xi} e^{-|\omega|\xi y} d\xi \right] \\
&= - i \operatorname{sgn}(\omega) \frac{|\omega|}{2\pi} \int_0^\infty \left(e^{\omega i x \cdot \xi} - e^{-\omega i x \cdot \xi} \right) \frac{\partial_\xi e^{-|\omega|\xi y}}{-|\omega| y} d\xi \\
&= i \operatorname{sgn}(\omega) \frac{1}{2\pi y} \Big[\left(e^{\omega i x \cdot \xi} - e^{-\omega i x \cdot \xi} \right) e^{-|\omega|\xi y} \Big|_{\xi=0}^\infty \\
&\qquad - \int_0^\infty \omega i x \left(e^{\omega i x \cdot \xi} + e^{-\omega i x \cdot \xi} \right) e^{-|\omega|\xi y} d\xi \Big] \\
&= \frac{|\omega| x}{2\pi y} \int_0^\infty \left(e^{\omega i x \cdot \xi} + e^{-\omega i x \cdot \xi} \right) e^{-|\omega|\xi y} d\xi \\
&= \frac{|\omega| x}{2\pi y} \int_{\mathbb{R}} e^{-\omega i x \cdot \xi} e^{-|\omega \xi| y} d\xi = \frac{x}{y} \mathscr{F} \left(\left(\frac{|\omega|}{2\pi} \right)^{1/2} e^{-|\omega \xi| y} \right) \\
&= \frac{x}{y} P_y(x) = \frac{x}{y} \frac{c_1 y}{y^2 + x^2} = \frac{c_1 x}{y^2 + x^2},
\end{aligned}
$$

where $c_1 = \Gamma(1)/\pi = 1/\pi$. That is,

$$
Q_y(x) = \frac{1}{\pi} \frac{x}{y^2 + x^2}.
$$

One can immediately verify that $Q(x, y) = Q_y(x)$ is a harmonic function in the upper half-plane and is the conjugate of the Poisson kernel $P_y(x) = P(x, y)$. More precisely, they satisfy Cauchy-Riemann equations

$$
\partial_x P = \partial_y Q = -\frac{1}{\pi} \frac{2xy}{(y^2 + x^2)^2}, \quad \partial_y P = -\partial_x Q = \frac{1}{\pi} \frac{x^2 - y^2}{(y^2 + x^2)^2}.
$$

In Theorem 4.1.3, we studied the limit of $u(x, t)$ as $y \to 0$ using the fact that $\{P_y\}$ is an approximation of the identity. We would like to do the same for $v(x, y)$, but we immediately run into an obstacle: $\{Q_y\}$ is not an approximation of the identity and, in fact, Q_y is not integrable for any $y > 0$. Formally,

$$
\lim_{y \to 0} Q_y(x) = \frac{1}{\pi x},
$$

this is not even locally integrable, so we cannot define its convolution with smooth functions.

§ 4.1.2 Hilbert transform

We define a tempered distribution called the ***principal value of*** $1/x$, abbreviated $\text{p.v.}\, 1/x$, by

$$
\left\langle \text{p.v.}\, \frac{1}{x}, \phi \right\rangle = \lim_{\varepsilon \to 0} \int_{|x| > \varepsilon} \frac{\phi(x)}{x} dx, \quad \phi \in \mathscr{S}.
$$

To see that this expression defines a tempered distribution, we rewrite it as

$$
\left\langle \text{p.v.}\, \frac{1}{x}, \phi \right\rangle = \int_{|x|<1} \frac{\phi(x) - \phi(0)}{x} dx + \int_{|x| \geqslant 1} \frac{\phi(x)}{x} dx,
$$

this holds since the integral of $1/x$ on $\varepsilon < |x| < 1$ is zero. It is now immediate that

$$
\left| \left\langle \text{p.v.}\, \frac{1}{x}, \phi \right\rangle \right| \leqslant C(\|\phi'\|_\infty + \|x\phi\|_\infty).
$$

Proposition 4.1.5.

In $\mathscr{S}'(\mathbb{R})$, we have $\lim_{y\to 0} Q_y(x) = \frac{1}{\pi}\,\text{p.v.}\,\frac{1}{x}$.

Proof. For each $\varepsilon > 0$, the functions $\psi_\varepsilon(x) = x^{-1}\chi_{|x|>\varepsilon}$ are bounded and define tempered distributions. It follows at once from the definition that in $\mathscr{S}'$,

$$\lim_{\varepsilon\to 0}\psi_\varepsilon(x) = \text{p.v.}\,\frac{1}{x}.$$

Therefore, it will suffice to prove that in $\mathscr{S}'$

$$\lim_{y\to 0}\left(Q_y - \frac{1}{\pi}\psi_y\right) = 0.$$

Fix $\phi \in \mathscr{S}$, then by a change of variables, we have

$$\begin{aligned}\langle \pi Q_y - \psi_y, \phi\rangle =& \int_{\mathbb{R}} \frac{x\phi(x)}{y^2+x^2}dx - \int_{|x|\geqslant y}\frac{\phi(x)}{x}dx\\ =& \int_{|x|<y}\frac{x\phi(x)}{y^2+x^2}dx + \int_{|x|\geqslant y}\left(\frac{x}{y^2+x^2} - \frac{1}{x}\right)\phi(x)dx\\ =& \int_{|x|<1}\frac{x\phi(yx)}{1+x^2}dx - \int_{|x|\geqslant 1}\frac{\phi(yx)}{x(1+x^2)}dx.\end{aligned}$$

If we take the limit as $y\to 0$ and apply the dominated convergence theorem, we get two integrals of odd functions on symmetric domains. Hence, the limit equals 0. ■

As a consequence of this proposition, we get that

$$\lim_{y\to 0} Q_y * f(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|t|>\varepsilon}\frac{f(x-t)}{t}dt,$$

and by the continuity of the Fourier transform on $\mathscr{S}'$ and by (4.1.7), we get

$$\mathscr{F}\left(\frac{1}{\pi}\,\text{p.v.}\,\frac{1}{x}\right)(\xi) = -i\,\text{sgn}\,(\omega)\left(\frac{|\omega|}{2\pi}\right)^{1/2}\text{sgn}\,(\xi).$$

Given a function $f\in\mathscr{S}$, we can define its ***Hilbert transform*** by any one of the following equivalent expressions:

$$\begin{aligned}Hf =& \lim_{y\to 0} Q_y * f,\\ Hf =& \frac{1}{\pi}\,\text{p.v.}\,\frac{1}{x} * f,\\ Hf =& \mathscr{F}^{-1}(-i\,\text{sgn}\,(\omega)\,\text{sgn}\,(\xi)\widehat{f}(\xi)).\end{aligned}$$

The third expression also allows us to define the Hilbert transform of functions in $L^2(\mathbb{R})$, which satisfies, with the help of Theorem 2.2.1,

$$\|Hf\|_2 = \|\widehat{Hf}\|_2 = \|\widehat{f}\|_2 = \|f\|_2, \tag{4.1.8}$$

that is, H is an isometry on $L^2(\mathbb{R})$. Moreover, H satisfies

$$H^2 f = H(Hf) = \mathscr{F}^{-1}((-i\,\text{sgn}\,(\omega)\,\text{sgn}\,(\xi))^2\widehat{f}(\xi)) = -f. \tag{4.1.9}$$

By Theorem 2.2.3, we have

$$\begin{aligned}\langle Hf, g\rangle =& \int_{\mathbb{R}} Hf\cdot g dx = \int_{\mathbb{R}}\mathscr{F}^{-1}(-i\,\text{sgn}\,(\omega)\,\text{sgn}\,(\xi)\widehat{f}(\xi))\cdot g dx\\ =& \int_{\mathbb{R}} -i\,\text{sgn}\,(\omega)\,\text{sgn}\,(\xi)\widehat{f}(\xi)\cdot\check{g}(\xi)d\xi\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}} f(x) \cdot \mathscr{F}[-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \check{g}(\xi)](x) dx \\
&= \int_{\mathbb{R}} f(x) \cdot \mathscr{F}[-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \hat{g}(-\xi)](x) dx \\
&= \int_{\mathbb{R}} f(x) \cdot \mathscr{F}^{-1}[i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \hat{g}(\xi)](x) dx \\
&= \langle f, -Hg \rangle, \qquad (4.1.10)
\end{aligned}
$$

namely, the dual/conjugate operator of H is $H' = -H$. Similarly, the adjoint operator H^* of H is uniquely defined via the identity

$$
(f, Hg) = \int_{\mathbb{R}} f \cdot \overline{Hg} dx = -\int_{\mathbb{R}} H f \bar{g} dx = (-Hf, g) =: (H^* f, g),
$$

that is, $H^* = -H$.

Note that for given $x \in \mathbb{R}$, $Hf(x)$ is defined for all $f \in L^1(\mathbb{R})$ satisfying the following Hölder condition near the point x:

$$
|f(x) - f(t)| \leqslant C_x |x - t|^{\varepsilon_x}
$$

for some $C_x > 0$ and $\varepsilon_x > 0$ whenever $|t - x| < \delta_x$. Indeed, suppose that this is the case, then

$$
\begin{aligned}
\lim_{y \to 0} Q_y * f(x) =& \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon < |x-t| < \delta_x} \frac{f(t)}{x - t} dt + \frac{1}{\pi} \int_{|x-t| \geqslant \delta_x} \frac{f(t)}{x - t} dt \\
=& \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon < |x-t| < \delta_x} \frac{f(t) - f(x)}{x - t} dt + \frac{1}{\pi} \int_{|x-t| \geqslant \delta_x} \frac{f(t)}{x - t} dt.
\end{aligned}
$$

Both integrals converge absolutely, and hence the limit of $Q_y * f(x)$ exists as $y \to 0$. Therefore, the Hilbert transform of a piece-wise smooth integrable function is well-defined at all points of Hölder-Lipschitz continuity of the function. On the other hand, observe that $Q_y * f$ is well-defined for all $f \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$, as it follows from the Hölder inequality, since $Q_y(x)$ is in $L^{p'}(\mathbb{R})$. Indeed,

$$
\begin{aligned}
\|\pi Q_y\|_{p'}^{p'} =& 2 \int_0^\infty \left(\frac{x}{x^2 + y^2} \right)^{p'} dx \\
=& 2 y^{1-p'} \int_0^\infty \left(\frac{x}{x^2 + 1} \right)^{p'} dx \\
=& 2 y^{1-p'} \int_0^{\pi/2} \sin^{p'} \theta \cos^{p'-2} \theta d\theta \quad (\text{let } x = \tan \theta) \\
=& y^{1-p'} B(p' + 1, p' - 1),
\end{aligned}
$$

where the Beta function converges if $p' - 1 > 0$. Thus, we obtain for $p' \in (1, \infty)$,

$$
\|Q_y\|_{p'} = \frac{1}{\pi} (B(p' + 1, p' - 1))^{1/p'} y^{-1/p}, \text{ and } \|Q_y\|_\infty = \frac{1}{2\pi y}.
$$

Definition 4.1.6.

The ***truncated Hilbert transform*** (at height ε) of a function $f \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$, is defined by

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

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 $|x| \geqslant \varepsilon$.

It is clear that the Hilbert transform of $f \in \mathscr{S}$ can be given by

$$Hf(x) = \lim_{\varepsilon\to 0} H^{(\varepsilon)} f(x). \tag{4.1.11}$$

Example 4.1.7. Consider the characteristic function $\chi_{[a,b]}$ of an interval $[a,b]$. It is a simple calculation to show that

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}. \tag{4.1.12}$$

Let us verify this identity. By the definition, we have

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \lim_{\varepsilon\to 0} \int_{|y|>\varepsilon} \frac{\chi_{[a,b]}(x-y)}{y} dy = \frac{1}{\pi} \lim_{\varepsilon\to 0} \int_{\substack{|y|>\varepsilon \\ x-b\leqslant y\leqslant x-a}} \frac{1}{y} dy.$$

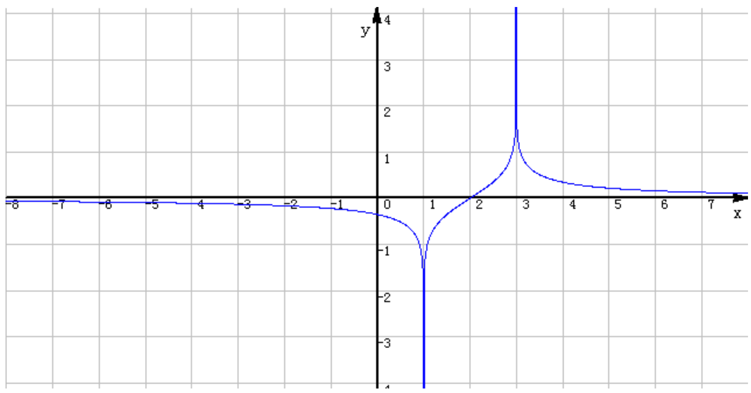
It is clear that it will be $-\infty$ and $+\infty$ at $x=a$ and $x=b$, respectively. Thus, we only need to consider three cases: $x-b>0$, $x-a<0$ and $x-b<0<x-a$. For the first two cases, we have

$$H(\chi_{[a,b]})(x) = \frac{1}{\pi} \int_{x-b}^{x-a} \frac{1}{y} dy = \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}.$$

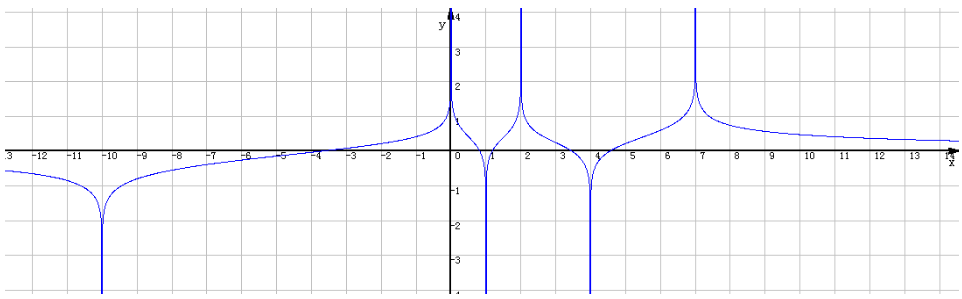
For the third case we get (without loss of generality, we can assume $\varepsilon < \min(|x-a|, |x-b|)$)

$$\begin{aligned} H(\chi_{[a,b]})(x) =& \frac{1}{\pi} \lim_{\varepsilon\to 0} \left(\int_{x-b}^{-\varepsilon} \frac{1}{y} dy + \int_{\varepsilon}^{x-a} \frac{1}{y} dy \right) \\ =& \frac{1}{\pi} \lim_{\varepsilon\to 0} \left(\ln \frac{|x-a|}{\varepsilon} + \ln \frac{\varepsilon}{|x-b|} \right) \\ =& \frac{1}{\pi} \ln \frac{|x-a|}{|x-b|}, \end{aligned}$$

where it is crucial to observe how the cancellation of the odd kernel $1/x$ is manifested. Note that $H(\chi_{[a,b]})(x)$ blows up logarithmically for x near the points a and b and decays like x^{-1} as $x \to \pm\infty$. See the following graph with $a=1$ and $b=3$:



The following is a graph of the function $H(\chi_{[-10,0]\cup[1,2]\cup[4,7]})$:



It is obvious, for the dilation operator δ^ε with $\varepsilon > 0$, by changes of variables ($\varepsilon y \to y$), that

$$\begin{aligned}(H\delta^\varepsilon)f(x) &= \lim_{\sigma\to 0}\frac{1}{\pi}\int_{|y|\geqslant\sigma}\frac{f(\varepsilon x-\varepsilon y)}{y}dy\\ &= \lim_{\sigma\to 0}\int_{|y|\geqslant\varepsilon\sigma}\frac{f(\varepsilon x-y)}{y}dy = (\delta^\varepsilon H)f(x),\end{aligned}$$

so $H\delta^\varepsilon = \delta^\varepsilon H$; and it follows obviously that $H\delta^\varepsilon = -\delta^\varepsilon H$, if $\varepsilon < 0$.

These simple considerations of dilation "invariance" and the obvious translation invariance in fact characterize the Hilbert transform.

Proposition 4.1.8: Characterization of Hilbert transform

Suppose T is a bounded linear operator on $L^2(\mathbb{R})$ which satisfies the following properties:
(a) T commutes with translations;
(b) T commutes with positive dilations;
(c) T anticommutes with the 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 2.5.6, there is a bounded function $m(\xi)$ such that $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$. The assumptions (b) and (c) may be written as $T\delta^\varepsilon f = \operatorname{sgn}(\varepsilon)\delta^\varepsilon Tf$ for all $f \in L^2(\mathbb{R})$. By part (iv) in Proposition 2.1.2, we have

$$\begin{aligned}\mathscr{F}(T\delta^\varepsilon f)(\xi) =& m(\xi)\mathscr{F}(\delta^\varepsilon f)(\xi) = m(\xi)|\varepsilon|^{-1}\hat{f}(\xi/\varepsilon),\\ \operatorname{sgn}(\varepsilon)\mathscr{F}(\delta^\varepsilon Tf)(\xi) =& \operatorname{sgn}(\varepsilon)|\varepsilon|^{-1}\widehat{Tf}(\xi/\varepsilon) = \operatorname{sgn}(\varepsilon)|\varepsilon|^{-1}m(\xi/\varepsilon)\hat{f}(\xi/\varepsilon),\end{aligned}$$

which means $m(\varepsilon\xi) = \operatorname{sgn}(\varepsilon)m(\xi)$, if $\varepsilon \neq 0$. This shows that $m(\xi) = c\operatorname{sgn}(\xi)$, and the proposition is proved. ■

§ 4.1.3 L^p boundedness of Hilbert transform

The next theorem shows that the Hilbert transform, now defined for functions in $\mathscr{S}$ or L^2, can be extended to functions in L^p, $1 \leqslant p < \infty$.

Theorem 4.1.9.

For $f \in \mathscr{S}(\mathbb{R})$, the following assertions hold:
(i) (Kolmogorov) H is of weak type $(1,1)$:

$$\mu(\{x\in\mathbb{R} : |Hf(x)| > \alpha\}) \leqslant \frac{C}{\alpha}\|f\|_1.$$

(ii) (M. Riesz) H is of type (p,p), $1 < p < \infty$:

$$\|Hf\|_p \leqslant C_p\|f\|_p.$$

Therefore, the Hilbert transform H admits an extension to a bounded operator on $L^p(\mathbb{R})$ when $1 < p < \infty$.

Proof. (i) Fix $\alpha > 0$. From the Calderón-Zygmund decomposition of f at height α (Theorem 3.3.4), there exist two functions g and b such that $f = g + b$ and

(1) $\|g\|_1 \leqslant \|f\|_1$ and $\|g\|_\infty \leqslant 2\alpha$.

(2) $b = \sum_j b_j$, where each b_j is supported in a dyadic interval I_j satisfying $\int_{I_j} b_j(x)dx = 0$ and $\|b_j\|_1 \leqslant 4\alpha\mu(I_j)$. Furthermore, the intervals I_j and I_k have disjoint interiors when $j \neq k$.

(3) $\sum_j \mu(I_j) \leqslant \alpha^{-1}\|f\|_1$.

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 $\mu(\Omega^*) \leqslant 2\mu(\Omega) \leqslant 2\alpha^{-1}\|f\|_1$.

Since $Hf = Hg + Hb$, from parts (iv) and (vi) of Proposition 1.1.3, (4.1.8) and (1), we have

$$\begin{aligned}
&(Hf)_*(\alpha) \leqslant (Hg)_*(\alpha/2) + (Hb)_*(\alpha/2)\\
\leqslant&(\alpha/2)^{-2}\int_{\mathbb{R}} |Hg(x)|^2 dx + \mu(\Omega^*) + \mu(\{x \notin \Omega^* : |Hb(x)| > \alpha/2\})\\
\leqslant&\frac{4}{\alpha^2}\int_{\mathbb{R}} |g(x)|^2 dx + 2\alpha^{-1}\|f\|_1 + 2\alpha^{-1}\int_{\mathbb{R}\setminus\Omega^*} |Hb(x)|dx\\
\leqslant&\frac{8}{\alpha}\int_{\mathbb{R}} |g(x)|dx + \frac{2}{\alpha}\|f\|_1 + \frac{2}{\alpha}\int_{\mathbb{R}\setminus\Omega^*}\sum_j |Hb_j(x)|dx\\
\leqslant&\frac{8}{\alpha}\|f\|_1 + \frac{2}{\alpha}\|f\|_1 + \frac{2}{\alpha}\sum_j\int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx.
\end{aligned}$$

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 $|x-y| \geqslant \mu(I_j)/2$ for $y \in I_j$. Denote the center of I_j by c_j, then, since b_j is mean value zero, we have

$$\begin{aligned}
\int_{\mathbb{R}\setminus 2I_j} |Hb_j(x)|dx =& \int_{\mathbb{R}\setminus 2I_j}\left|\frac{1}{\pi}\int_{I_j}\frac{b_j(y)}{x-y}dy\right| dx\\
=&\frac{1}{\pi}\int_{\mathbb{R}\setminus 2I_j}\left|\int_{I_j} b_j(y)\left(\frac{1}{x-y} - \frac{1}{x-c_j}\right)dy\right| dx\\
\leqslant&\frac{1}{\pi}\int_{I_j}|b_j(y)|\left(\int_{\mathbb{R}\setminus 2I_j}\frac{|y-c_j|}{|x-y||x-c_j|}dx\right)dy\\
\leqslant&\frac{1}{\pi}\int_{I_j}|b_j(y)|\left(\int_{\mathbb{R}\setminus 2I_j}\frac{\mu(I_j)}{|x-c_j|^2}dx\right)dy.
\end{aligned}$$

The last inequality follows from the fact that $|y-c_j| < \mu(I_j)/2$ and $|x-y| > |x-c_j|/2$. Since $|x-c_j| > \mu(I_j)$, the inner integral equals

$$2\mu(I_j)\int_{\mu(I_j)}^{\infty}\frac{1}{r^2}dr = 2\mu(I_j)\frac{1}{\mu(I_j)} = 2.$$

Thus, by (2) and (3),

$$\begin{aligned}
(Hf)_*(\alpha) \leqslant&\frac{10}{\alpha}\|f\|_1 + \frac{4}{\alpha\pi}\sum_j\int_{I_j}|b_j(y)|dy \leqslant \frac{10}{\alpha}\|f\|_1 + \frac{4}{\alpha\pi}\sum_j 4\alpha\mu(I_j)\\
\leqslant&\frac{10}{\alpha}\|f\|_1 + \frac{16}{\pi}\frac{1}{\alpha}\|f\|_1 = \frac{10+16/\pi}{\alpha}\|f\|_1.
\end{aligned}$$

(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 (4.1.10) 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. ■

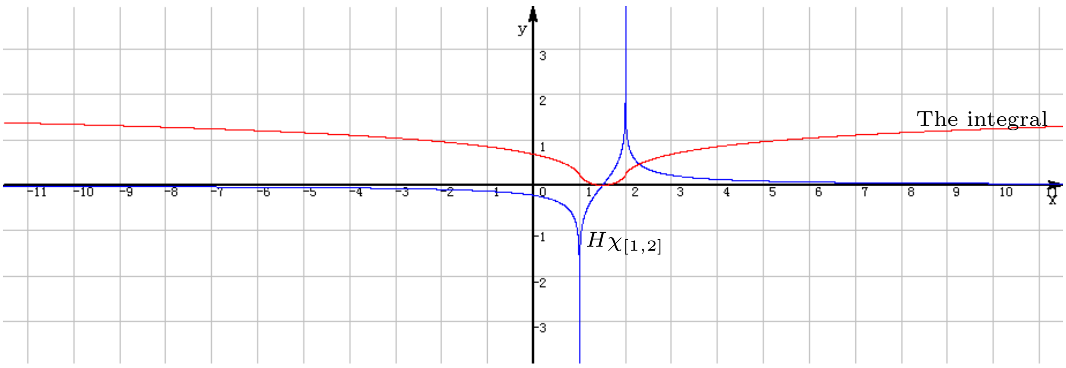
Remark 4.1.10. i) Recall from the proof of the Marcinkiewicz interpolation theorem that the coefficient

$$C_p = \begin{cases} \dfrac{(10+16/\pi)p}{p-1} + \dfrac{2\sqrt{2}}{2-p}, & 1 < p < 2, \\ (10+16/\pi)p + 2\sqrt{2}\dfrac{p-1}{p-2}, & p > 2. \end{cases}$$

So the constant C_p tends to infinity as p tends to 1 or ∞. More precisely,

$$C_p = O(p) \text{ as } p \to \infty, \text{ and } C_p = O((p-1)^{-1}) \text{ as } p \to 1.$$

ii) The strong (p,p) inequality is false if $p = 1$ or $p = \infty$, this can 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 4.1.9, 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 $\varepsilon > 0$,

$$\lim_{m,n\to\infty} \mu(\{x \in \mathbb{R} : |(Hf_n - Hf_m)(x)| > \varepsilon\}) = 0.$$

Therefore, it converges in measure to a measurable function which we define to be the Hilbert transform of f.

If $f \in L^p$, $1 < p < \infty$, and $\{f_n\}$ is a sequence of functions in $\mathscr{S}$ that converges to f in L^p, by the strong (p,p) inequality, $\{Hf_n\}$ is a Cauchy sequence in L^p, so it converges to a function in L^p which we call the Hilbert transform of f.

In either case, a subsequence of $\{Hf_n\}$, depending on f, converges pointwise almost everywhere to Hf as defined.

§ 4.1.4 The maximal Hilbert transform and its L^p boundedness

We now introduce the maximal Hilbert transform.

Definition 4.1.11.

The ***maximal Hilbert transform*** is the operator

$$H^{(*)}f(x) = \sup_{\varepsilon>0} |H^{(\varepsilon)}f(x)| \tag{4.1.13}$$

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 4.1.12. Using the result of Example 4.1.7, 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 disjoint 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 have the estimate $\|H^{(*)}f\|_p \leqslant C_p\|f\|_p$ for $f \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$, Theorem 3.2.13 yields 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 4.1.9 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 this ideas.

Theorem 4.1.13.

There exists a constant C such that for all $p \in (1,\infty)$, we have

$$\|H^{(*)}f\|_p \leqslant C\max(p,(p-1)^{-2})\|f\|_p. \tag{4.1.14}$$

Moreover, for all $f \in L^p(\mathbb{R})$, $H^{(\varepsilon)}f$ converges to Hf a.e. and in L^p.

Proof. Recall the kernels

$$P_\varepsilon = \frac{1}{\pi}\frac{\varepsilon}{x^2+\varepsilon^2}, \quad Q_\varepsilon = \frac{1}{\pi}\frac{x}{x^2+\varepsilon^2}.$$

From Corollary 2.1.23 and (4.1.7), we know

$$\widehat{P_\varepsilon}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{1/2} e^{-\varepsilon|\omega\xi|}, \quad \widehat{Q_\varepsilon}(\xi) = -i\,\mathrm{sgn}\,(\omega\xi)\left(\frac{|\omega|}{2\pi}\right)^{1/2} e^{-\varepsilon|\omega\xi|}.$$

Thus,

$$\widehat{f*Q_\varepsilon} = -i\,\mathrm{sgn}\,(\omega\xi)e^{-\varepsilon|\omega\xi|}\widehat{f} = e^{-\varepsilon|\omega\xi|}\widehat{Hf} = \left(\frac{|\omega|}{2\pi}\right)^{-1/2}\widehat{P_\varepsilon}\widehat{Hf} = \widehat{P_\varepsilon * Hf},$$

which implies for all $f \in L^p$

$$f*Q_\varepsilon = Hf*P_\varepsilon, \quad \varepsilon > 0. \tag{4.1.15}$$

Then, we have

$$H^{(\varepsilon)}f = H^{(\varepsilon)}f - f*Q_\varepsilon + Hf*P_\varepsilon. \tag{4.1.16}$$

Using the identity

$$H^{(\varepsilon)}f(x) - (f*Q_\varepsilon)(x) = -\frac{1}{\pi}\left[\int_{-\infty}^{\infty}\frac{tf(x-t)}{t^2+\varepsilon^2}dt - \int_{|t|\geqslant\varepsilon}\frac{f(x-t)}{t}dt\right]$$

$$= -\frac{1}{\pi}\int_{\mathbb{R}} f(x-t)\psi_\varepsilon(t)dt, \tag{4.1.17}$$

where $\psi_\varepsilon(x) = \varepsilon^{-1}\psi(\varepsilon^{-1}x)$ and

$$\psi(t) = \begin{cases} \dfrac{t}{t^2+1} - \dfrac{1}{t}, & \text{if } |t| \geqslant 1, \\ \dfrac{t}{t^2+1}, & \text{if } |t| < 1. \end{cases}$$

Note that ψ has integral zero since ψ is an odd function and is integrable over the line. Indeed,

$$\begin{aligned}
\int_{\mathbb{R}} |\psi(t)|dt =& \int_{|t|\geqslant 1} \left| \frac{t}{t^2+1} - \frac{1}{t} \right| dt + \int_{|t|<1} \frac{|t|}{t^2+1} dt \\
=& \int_{|t|\geqslant 1} \frac{1}{(t^2+1)|t|} dt + \int_{|t|<1} \frac{|t|}{t^2+1} dt \\
=& \int_1^\infty \frac{dt^2}{(t^2+1)t^2} + \int_0^1 \frac{dt^2}{t^2+1} \\
=& \int_1^\infty \frac{ds}{(s+1)s} + \int_0^1 \frac{ds}{s+1} \\
=& \int_1^\infty \left(\frac{1}{s} - \frac{1}{s+1} \right) ds + \int_0^1 \frac{ds}{s+1} \\
=& \left[\ln \left| \frac{s}{s+1} \right| \right]_1^\infty + [\ln |s+1|]_0^1 \\
=& 2\ln 2.
\end{aligned}$$

The least decreasing radial majorant of ψ is

$$\Psi(t) = \sup_{|s|\geqslant|t|} |\psi(s)| = \begin{cases} \dfrac{1}{(t^2+1)|t|}, & \text{if } |t| \geqslant 1, \\ \dfrac{1}{2}, & \text{if } |t| < 1, \end{cases}$$

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 3.2.12 that

$$\sup_{\varepsilon>0} |H^{(\varepsilon)}f(x) - (f * Q_\varepsilon)(x)| \leqslant \frac{\ln 2 + 1}{\pi} Mf(x). \tag{4.1.18}$$

In view of (4.1.16) and (4.1.18), from Theorem 4.1.3 we obtain for $f \in L^p(\mathbb{R})$ 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 Theorems 3.2.7 and 4.1.9 that $H^{(*)}$ is L^p bounded with norm at most $C\max(p,(p-1)^{-2})$.

Applying Corollary 2.1.16 to (4.1.17), we have $\lim\limits_{\varepsilon\to 0} \|H^{(\varepsilon)}f - (f * Q_\varepsilon)\|_p = 0$ since ψ has integral zero. By Theorem 2.1.15, we also have $\lim\limits_{\varepsilon\to 0} \|Hf * P_\varepsilon - Hf\|_p = 0$. Thus, from (4.1.16), it follows that $\lim\limits_{\varepsilon\to 0} \|H^{(\varepsilon)}f - Hf\|_p = 0$ and therefore we also have $H^{(\varepsilon)}f \to Hf$ a.e. as $\varepsilon \to 0$. ■

§4.2 Calderón-Zygmund singular integrals

From this section on, we are going to consider singular integrals whose kernels have the same essential properties as the kernel of the Hilbert transform. We can generalize Theorem 4.1.9 to get the following result.

Theorem 4.2.1: Calderón-Zygmund Theorem

Let K be a tempered distribution in $\mathbb{R}^n$ which coincides with a locally integrable function on $\mathbb{R}^n \setminus \{0\}$ and satisfies

$$|\widehat{K}(\xi)| \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} B, \tag{4.2.1}$$

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \quad y \in \mathbb{R}^n. \tag{4.2.2}$$

Then we have the strong type (p, p) estimate for $1 < p < \infty$

$$\|K * f\|_p \leqslant C_p \|f\|_p, \tag{4.2.3}$$

and the weak type $(1, 1)$ estimate

$$(K * f)_*(\alpha) \leqslant \frac{C}{\alpha}\|f\|_1. \tag{4.2.4}$$

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. Condition (4.2.2) is usually referred to as the ***Hörmander condition***; in practice it is often deduced from another stronger condition called the ***gradient condition*** (i.e., (4.2.5) as below).

Proposition 4.2.2.

The Hörmander condition (4.2.2) holds if for every $x \neq 0$

$$|\nabla K(x)| \leqslant \frac{C}{|x|^{n+1}}. \tag{4.2.5}$$

Proof. By the integral mean value theorem and (4.2.5), we have

$$\begin{aligned}
&\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant \int_{|x|\geqslant 2|y|} \int_0^1 |\nabla K(x-\theta y)||y|d\theta dx \\
\leqslant & \int_0^1 \int_{|x|\geqslant 2|y|} \frac{C|y|}{|x-\theta y|^{n+1}} dx d\theta \leqslant \int_0^1 \int_{|x|\geqslant 2|y|} \frac{C|y|}{(|x|/2)^{n+1}} dx d\theta \\
\leqslant & 2^{n+1} C|y|\omega_{n-1} \int_{2|y|}^\infty \frac{1}{r^2} dr = 2^{n+1}C|y|\omega_{n-1}\frac{1}{2|y|} = 2^n C\omega_{n-1}.
\end{aligned}$$

This completes the proof. ■

Proof of Theorem 4.2.1. Let $f \in \mathscr{S}$ and $Tf = K * f$. From (4.2.1), it follows that

$$\begin{aligned}
\|Tf\|_2 =& \|\widehat{Tf}\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|\widehat{K}\widehat{f}\|_2 \\
\leqslant & \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \|\widehat{K}\|_\infty \|\widehat{f}\|_2 \leqslant B\|\widehat{f}\|_2 \\
=& B\|f\|_2,
\end{aligned} \tag{4.2.6}$$

by the Plancherel theorem (Theorem 2.2.1) and part (vi) in Proposition 2.1.2.

It will suffice to prove that T is of weak type $(1,1)$ since the strong (p,p) inequality, $1<p<2$, follows from the interpolation, and for $p>2$ it follows from the duality since the conjugate operator T' has kernel $K'(x)=K(-x)$ which also satisfies (4.2.1) and (4.2.2). 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 4.1.9, we can write $f=g+b$, where

(i) $\|g\|_1\leqslant\|f\|_1$ and $\|g\|_\infty\leqslant 2^n\alpha$.

(ii) $b=\sum_j b_j$, where each b_j is supported in a dyadic cube Q_j satisfying

$$\int_{Q_j}b_j(x)dx=0 \text{ and } \|b_j\|_1\leqslant 2^{n+1}\alpha\mu(Q_j).$$

Furthermore, the cubes Q_j and Q_k have disjoint interiors when $j\neq k$.

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

The argument now proceeds as in Theorem 4.1.9, and the proof reduces to showing that

$$\int_{\mathbb{R}^n\setminus Q_j^*}|Tb_j(x)|dx\leqslant C\int_{Q_j}|b_j(x)|dx, \tag{4.2.7}$$

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 (4.2.7) follows from the Hörmander condition (4.2.2): since each b_j has zero average, if $x\notin Q_j^*$

$$Tb_j(x)=\int_{Q_j}K(x-y)b_j(y)dy=\int_{Q_j}[K(x-y)-K(x-c_j)]b_j(y)dy;$$

hence,

$$\int_{\mathbb{R}^n\setminus Q_j^*}|Tb_j(x)|dx\leqslant\int_{Q_j}\left(\int_{\mathbb{R}^n\setminus Q_j^*}|K(x-y)-K(x-c_j)|dx\right)|b_j(y)|dy.$$

By changing variables $x-c_j=x'$ and $y-c_j=y'$, and the fact that $|x-c_j|\geqslant 2|y-c_j|$ for all $x\notin Q_j^*$ and $y\in Q_j$ as an obvious geometric consideration shows, and (4.2.2), we get

$$\int_{\mathbb{R}^n\setminus Q_j^*}|K(x-y)-K(x-c_j)|dx\leqslant\int_{|x'|\geqslant 2|y'|}|K(x'-y')-K(x')|dx'\leqslant B.$$

Since the remainder proof is (essentially) a repetition of the proof of Theorem 4.1.9, we omit the details and complete the proof. ■

There is still an element which 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) An extraneous condition such as $K\in L^2$ subsists in the hypothesis; and for this reason our results do not directly treat the "principal-value" singular integrals, those which exist because of the cancelation of positive and negative values.

However, from what we have done, it is now a relatively simple matter to obtain a theorem which covers the cases of interest.

Definition 4.2.3.

Suppose that $K \in L^1_{\rm loc}(\mathbb{R}^n \setminus \{0\})$ and satisfies the following conditions:

$$\begin{aligned} &|K(x)| \leqslant B|x|^{-n}, \quad \forall x \neq 0, \\ &\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \quad \forall y \neq 0, \end{aligned} \tag{4.2.8}$$

and

$$\int_{R_1<|x|<R_2} K(x)dx = 0, \quad \forall 0 < R_1 < R_2 < \infty. \tag{4.2.9}$$

Then K is called the ***Calderón-Zygmund kernel***, where B is a constant independent of x and y.

Theorem 4.2.4.

Suppose that K is a Calderón-Zygmund kernel. For $\varepsilon > 0$ and $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, let

$$T_\varepsilon f(x) = \int_{|y|\geqslant\varepsilon} f(x-y)K(y)dy. \tag{4.2.10}$$

Then the following conclusions hold:
(i) We have

$$\|T_\varepsilon f\|_p \leqslant A_p\|f\|_p \tag{4.2.11}$$

where A_p is independent of f and ε.
(ii) For any $f \in L^p(\mathbb{R}^n)$, $\lim_{\varepsilon\to 0} T_\varepsilon(f)$ exists in the sense of L^p norm. That is, there exists an operator T such that

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

(iii) $\|Tf\|_p \leqslant A_p\|f\|_p$ for $f \in L^p(\mathbb{R}^n)$.

Remark 4.2.5. 1) The linear operator T defined by (ii) of Theorem 4.2.4 is called the ***Calderón-Zygmund singular integral operator***. T_ε is also called the ***truncated operator*** of T.

2) The cancelation property alluded to is contained in condition (4.2.9). This hypothesis, together with (4.2.8), allows us to prove the L^2 boundedness and the L^p convergence of the truncated integrals (4.2.11).

3) We should point out that the kernel $K(x) = \frac{1}{\pi x}$, $x \in \mathbb{R}$, clearly satisfies the hypotheses of Theorem 4.2.4. Therefore, we have the existence of the Hilbert transform in the sense that if $f \in L^p(\mathbb{R})$, $1 < p < \infty$, then

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

exists in the L^p norm and the resulting operator is bounded in L^p, as has shown in Theorem 4.1.9.

For L^2 boundedness, we have the following lemma.

Lemma 4.2.6.

Suppose that K satisfies the conditions (4.2.8) and (4.2.9) 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

$$\sup_\xi |\widehat{K_\varepsilon}(\xi)|\leqslant\left(\frac{|\omega|}{2\pi}\right)^{n/2}CB,\quad \varepsilon>0, \tag{4.2.12}$$

where C depends only on the dimension n.

Proof. First, we prove the inequality (4.2.12) for the special case $\varepsilon=1$. Since $\widehat{K_1}(0)=0$, thus we can assume $\xi\neq 0$ and have

$$\begin{aligned}\widehat{K_1}(\xi)=&\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}K_1(x)dx\\ =&\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{|x|<2\pi/(|\omega||\xi|)}e^{-\omega ix\cdot\xi}K_1(x)dx\\ &+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{2\pi/(|\omega||\xi|)\leqslant|x|}e^{-\omega ix\cdot\xi}K_1(x)dx\\ =:&I_1+I_2.\end{aligned}$$

By the condition (4.2.9), $\int_{1<|x|<2\pi/(|\omega||\xi|)}K(x)dx=0$ which implies

$$\int_{|x|<2\pi/(|\omega||\xi|)}K_1(x)dx=0.$$

Thus, $\int_{|x|<2\pi/(|\omega||\xi|)}e^{-\omega ix\cdot\xi}K_1(x)dx=\int_{|x|<2\pi/(|\omega||\xi|)}[e^{-\omega ix\cdot\xi}-1]K_1(x)dx$. Hence, from the fact $|e^{i\theta}-1|\leqslant|\theta|$ (see Section 2.1) and the first condition in (4.2.8), we get

$$\begin{aligned}\left(\frac{|\omega|}{2\pi}\right)^{-n/2}|I_1|\leqslant&\int_{|x|<2\pi/(|\omega||\xi|)}|\omega||x||\xi||K_1(x)|dx\leqslant|\omega|B|\xi|\int_{|x|<2\pi/(|\omega||\xi|)}|x|^{-n+1}dx\\ =&\omega_{n-1}B|\omega||\xi|\int_0^{2\pi/(|\omega||\xi|)}dr=2\pi\omega_{n-1}B.\end{aligned}$$

To estimate I_2, choose $z=z(\xi)$ such that $e^{-\omega i\xi\cdot z}=-1$. This choice can be realized if $z=\pi\xi/(\omega|\xi|^2)$, with $|z|=\pi/(|\omega||\xi|)$. Since, by changing variables $x+z=y$, we get

$$\begin{aligned}\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}K_1(x)dx=&-\int_{\mathbb{R}^n}e^{-\omega i(x+z)\cdot\xi}K_1(x)dx=-\int_{\mathbb{R}^n}e^{-\omega iy\cdot\xi}K_1(y-z)dy\\ =&-\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}K_1(x-z)dx,\end{aligned}$$

which implies $\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}K_1(x)dx=\frac{1}{2}\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}[K_1(x)-K_1(x-z)]dx$, then we have

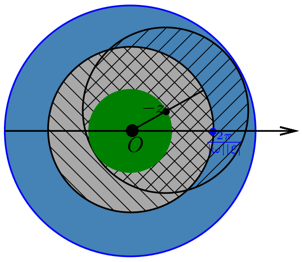
$$\begin{aligned}\left(\frac{|\omega|}{2\pi}\right)^{-n/2}I_2=&\left(\int_{\mathbb{R}^n}-\int_{|x|<2\pi/(|\omega||\xi|)}\right)e^{-\omega ix\cdot\xi}K_1(x)dx\\ =&\frac{1}{2}\int_{\mathbb{R}^n}e^{-\omega ix\cdot\xi}[K_1(x)-K_1(x-z)]dx\\ &-\int_{|x|<2\pi/(|\omega||\xi|)}e^{-\omega ix\cdot\xi}K_1(x)dx\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{2}\lim_{R\to\infty}\int_{2\pi/(|\omega||\xi|)\leqslant|x|\leqslant R} e^{-\omega ix\cdot\xi}[K_1(x)-K_1(x-z)]dx\\
&-\frac{1}{2}\int_{|x|<2\pi/(|\omega||\xi|)} e^{-\omega ix\cdot\xi}K_1(x)dx\\
&-\frac{1}{2}\int_{|x|<2\pi/(|\omega||\xi|)} e^{-\omega ix\cdot\xi}K_1(x-z)dx.
\end{aligned}
$$

The last two integrals are equal to, in view of the integration by parts,

$$
\begin{aligned}
&-\frac{1}{2}\int_{|x|<2\pi/(|\omega||\xi|)} e^{-\omega ix\cdot\xi}K_1(x)dx-\frac{1}{2}\int_{|y+z|<2\pi/(|\omega||\xi|)} e^{-\omega i(y+z)\cdot\xi}K_1(y)dy\\
=&-\frac{1}{2}\int_{|x|<2\pi/(|\omega||\xi|)} e^{-\omega ix\cdot\xi}K_1(x)dx+\frac{1}{2}\int_{|x+z|<2\pi/(|\omega||\xi|)} e^{-\omega ix\cdot\xi}K_1(x)dx\\
=&-\frac{1}{2}\int_{|x|<2\pi/(|\omega||\xi|)\leqslant|x+z|} e^{-\omega ix\cdot\xi}K_1(x)dx+\frac{1}{2}\int_{|x+z|<2\pi/(|\omega||\xi|)\leqslant|x|} e^{-\omega ix\cdot\xi}K_1(x)dx.
\end{aligned}
$$

For the first integral, we have $2\pi/(|\omega||\xi|)\geqslant|x|\geqslant|x+z|-|z|>2\pi/(|\omega||\xi|)-\pi/(|\omega||\xi|)=\pi/(|\omega||\xi|)$, and for the second one, $2\pi/(|\omega||\xi|)<|x|\leqslant|x+z|+|z|\leqslant 3\pi/(|\omega||\xi|)$. These two integrals are taken over a region contained in the spherical shell, $\pi/(|\omega||\xi|)<|x|\leqslant 3\pi/(|\omega||\xi|)$ (see the figure), and is bounded by $\frac{1}{2}B\omega_{n-1}\ln 3$ since $|K_1(x)|\leqslant B|x|^{-n}$. By $|z|=\pi/(|\omega||\xi|)$ and the condition (4.2.8), the first integral of I_2 is majorized by



$$
\begin{aligned}
&\frac{1}{2}\int_{|x|\geqslant 2\pi/(|\omega||\xi|)}|K_1(x-z)-K_1(x)|dx\\
=&\frac{1}{2}\int_{|x|\geqslant 2|z|}|K_1(x-z)-K_1(x)|dx\leqslant\frac{1}{2}B.
\end{aligned}
$$

Thus, we have obtained

$$
|\widehat{K_1}(\xi)|\leqslant\left(\frac{|\omega|}{2\pi}\right)^{n/2}\left(2\pi\omega_{n-1}B+\frac{1}{2}B+\frac{1}{2}B\omega_{n-1}\ln 3\right)\leqslant C_n\left(\frac{|\omega|}{2\pi}\right)^{n/2}B,
$$

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

$$
\begin{aligned}
\delta^{\varepsilon^{-1}}T\delta^\varepsilon f(x)=&\int_{\mathbb{R}^n}\varphi(\varepsilon^{-1}x-y)f(\varepsilon y)dy\\
=&\varepsilon^{-n}\int_{\mathbb{R}^n}\varphi(\varepsilon^{-1}(x-z))f(z)dz=\varphi_\varepsilon*f,
\end{aligned}
$$

where $\varphi_\varepsilon(x)=\varepsilon^{-n}\varphi(\varepsilon^{-1}x)$. In our case, if T corresponds to the kernel $K(x)$, then $\delta^{\varepsilon^{-1}}T\delta^\varepsilon$ corresponds to the kernel $\varepsilon^{-n}K(\varepsilon^{-1}x)$. ***Notice that if K satisfies the assumptions of our theorem, then $\varepsilon^{-n}K(\varepsilon^{-1}x)$ also satisfies these assumptions with the same bounds.*** (A similar remark holds for the assumptions of all the theorems in this chapter.) Now, with our K given, let $K'=\varepsilon^nK(\varepsilon x)$. Then K' satisfies the conditions of our lemma with the same bound B, and so if we denote

$$K_1'(x) = \begin{cases} K'(x), & |x| \geqslant 1, \\ 0, & |x| < 1, \end{cases}$$

then we know that $|\widehat{K_1'}(\xi)| \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} CB$. The Fourier transform of $\varepsilon^{-n} K_1'(\varepsilon^{-1}x)$ is $\widehat{K_1'}(\varepsilon\xi)$ which is again bounded by $\left(\frac{|\omega|}{2\pi}\right)^{n/2} CB$; however $\varepsilon^{-n} K_1'(\varepsilon^{-1}x) = K_\varepsilon(x)$, therefore the lemma is completely proved. ■

We can now prove Theorem 4.2.4.

Proof of Theorem 4.2.4. Since K satisfies the conditions (4.2.8) and (4.2.9), then $K_\varepsilon(x)$ satisfies the same conditions with bounds not greater than CB. By Lemma 4.2.6 and Theorem 4.2.1, 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 $f_1 \in \mathscr{C}_c^1(\mathbb{R}^n)$. In fact, we have

$$\begin{aligned} T_\varepsilon f_1(x) - T_\eta f_1(x) &= \int_{|y|\geqslant\varepsilon} K(y) f_1(x-y)dy - \int_{|y|\geqslant\eta} K(y) f_1(x-y)dy \\ &= \operatorname{sgn}(\eta-\varepsilon) \int_{\min(\varepsilon,\eta)\leqslant|y|\leqslant\max(\varepsilon,\eta)} K(y)[f_1(x-y) - f_1(x)]dy, \end{aligned}$$

because of the cancelation condition (4.2.9). For $p \in (1,\infty)$, we get, by the mean value theorem with some $\theta \in [0,1]$, Minkowski's inequality and (4.2.8), 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 the basic inequality (4.2.11) for f_2 to get $\|T_\varepsilon f_2\|_p \leqslant C\|f_2\|_p$, then we see that $\lim_{\varepsilon\to 0} T_\varepsilon f$ exists in L^p norm; that the limiting operator T also satisfies the inequality (4.2.11) is then obvious. Thus, we complete the proof of the theorem. ■

§4.3 L^2 boundedness of homogeneous singular integrals

Definition 4.3.1.

Let $\Omega \in L^1(S^{n-1})$ with mean value zero. For $0 < \varepsilon < N < \infty$ and $f \in \cup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$, we define the ***truncated singular integral***

$$T_\Omega^{(\varepsilon,N)} f(x) = \int_{\varepsilon\leqslant|y|\leqslant N} f(x-y)\frac{\Omega(y/|y|)}{|y|^n}dy. \tag{4.3.1}$$

Note that for $f \in L^p(\mathbb{R}^n)$, we have by Young's inequality

$$\begin{aligned}\|T_\Omega^{(\varepsilon,N)} f\|_p \leqslant &\|f\|_p \int_{\varepsilon\leqslant|y|\leqslant N} \frac{|\Omega(y/|y|)|}{|y|^n}dy\\ =&\|f\|_p \int_\varepsilon^N \int_{S^{n-1}} \frac{|\Omega(y')|}{r^n} r^{n-1} d\sigma(y')dr\\ =&\|f\|_p \|\Omega\|_{L^1(S^{n-1})} \ln\frac{N}{\varepsilon},\end{aligned}$$

which implies that (4.3.1) is finite a.e. and therefore well-defined.

We also note that the cancellation condition (4.2.9) is the same as the mean value zero condition

$$\int_{S^{n-1}} \Omega(x)d\sigma(x) = 0$$

where $K(x) = \frac{\Omega(x/|x|)}{|x|^n}$ and $d\sigma(x)$ is the induced Euclidean measure on S^{n-1}. In fact, this equation implies that

$$\begin{aligned}\int_{R_1<|x|<R_2} K(x)dx =& \int_{R_1}^{R_2}\int_{S^{n-1}} \frac{\Omega(x')}{r^n} d\sigma(x') r^{n-1}dr\\ =&\ln\left(\frac{R_2}{R_1}\right)\int_{S^{n-1}}\Omega(x')d\sigma(x').\end{aligned}$$

Definition 4.3.2.

We denote by T_Ω the singular integral operator whose kernel is $\text{p.v.}\,\frac{\Omega(x/|x|)}{|x|^n}$, i.e., for $f \in \mathscr{S}(\mathbb{R}^n)$

$$T_\Omega f(x) = \text{p.v.}\,\frac{\Omega(\cdot/|\cdot|)}{|\cdot|^n} * f(x) = \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} T_\Omega^{(\varepsilon,N)} f(x).$$

The associated ***maximal singular integral*** is defined by

$$T_\Omega^{(**)} f = \sup_{0<N<\infty}\sup_{0<\varepsilon<N} |T_\Omega^{(\varepsilon,N)} f|. \tag{4.3.2}$$

We note that if Ω is bounded, there is no need to use the upper truncations in the definition of $T_\Omega^{(\varepsilon,N)}$ given in (4.3.1). In this case, the maximal singular integrals could be defined as

$$T_\Omega^{(*)} f = \sup_{\varepsilon>0} |T_\Omega^{(\varepsilon)} f|, \tag{4.3.3}$$

where for $f \in \cup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$, $\varepsilon > 0$ and $x \in \mathbb{R}^n$, $T_\Omega^{(\varepsilon)} f(x)$ is defined in term 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 $\Omega \in L^\infty(S^{n-1})$, notice that

$$\left|T_\Omega^{(\varepsilon,N)} f(x)\right| = \left|\int_{\varepsilon \leqslant |y| \leqslant N} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy\right| \leqslant \sup_{0<N<\infty} \left|T_\Omega^{(\varepsilon,N)} f(x)\right|. \tag{4.3.4}$$

Then for $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, we let $N \to \infty$ on the l.h.s. in (4.3.4), and we note that the limit exists in view of the absolutely 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 $L^\infty(S^{n-1})$. This is the case with the Hilbert transform, that is, $H^{(**)}$ is comparable to $H^{(*)}$.

Next, we would like to compute the Fourier transforms of $\text{p.v.}\,\Omega(x/|x|)/|x|^n$. This provides information whether the operator T_Ω is L^2 bounded. We have the following result.

Theorem 4.3.3.

Let $\Omega \in L^1(S^{n-1})$ have mean value zero. Then the Fourier transform of $\left(\frac{|\omega|}{2\pi}\right)^{-n/2} \text{p.v.}\,\Omega(x/|x|)/|x|^n$ is a bounded homogeneous function of degree 0 given by

$$m(\xi) = \int_{S^{n-1}} \left[\ln(1/|\xi\cdot x|) - \frac{\pi i}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(\xi\cdot x)\right]\Omega(x)d\sigma(x), \quad |\xi| = 1. \tag{4.3.5}$$

Moreover, $m \in L^\infty(\mathbb{R}^n)$ and then T_Ω is L^2 bounded.

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 $\widehat{K_{\varepsilon,\eta} * f}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{K_{\varepsilon,\eta}}(\xi)\hat{f}(\xi)$.

We shall prove two facts about $\widehat{K_{\varepsilon,\eta}}(\xi)$.

(i) $\|\widehat{K_{\varepsilon,\eta}}\|_{L^\infty} \leqslant \left(\frac{|\omega|}{2\pi}\right)^{n/2} A$, with A independent of ε and η;

(ii) $\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}} \widehat{K_{\varepsilon,\eta}}(\xi) = m(\xi)$ a.e., see (4.3.5).

For this purpose, it is convenient to introduce polar coordinates. Let $x = rx'$, $r = |x|$, $x' = x/|x| \in S^{n-1}$, and $\xi = R\xi'$, $R = |\xi|$, $\xi' = \xi/|\xi| \in S^{n-1}$. Then we have

$$\begin{aligned}\left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{K_{\varepsilon,\eta}}(\xi) &= \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} K_{\varepsilon,\eta}(x)dx = \int_{\varepsilon\leqslant|x|\leqslant\eta} e^{-\omega i x\cdot\xi}\frac{\Omega(x/|x|)}{|x|^n}dx \\ &= \int_{S^{n-1}} \Omega(x')\left(\int_\varepsilon^\eta e^{-\omega i R r x'\cdot\xi'} r^{-n} r^{n-1} dr\right) d\sigma(x') \\ &= \int_{S^{n-1}} \Omega(x')\left(\int_\varepsilon^\eta e^{-\omega i R r x'\cdot\xi'} \frac{dr}{r}\right) d\sigma(x').\end{aligned}$$

Since

$$\int_{S^{n-1}} \Omega(x')d\sigma(x') = 0,$$

we can introduce the factor $\cos(|\omega|Rr)$ (which does not depend on x') in the integral defining $\widehat{K_{\varepsilon,\eta}}(\xi)$. We shall also need the auxiliary integral

$$I_{\varepsilon,\eta}(\xi,x')=\int_\varepsilon^\eta [e^{-\omega i Rr x'\cdot\xi'}-\cos(|\omega|Rr)]\frac{dr}{r},\quad R>0.$$

Thus, it follows

$$\left(\frac{|\omega|}{2\pi}\right)^{-n/2}\widehat{K_{\varepsilon,\eta}}(\xi)=\int_{S^{n-1}} I_{\varepsilon,\eta}(\xi,x')\Omega(x')d\sigma(x').$$

Now, we first consider $I_{\varepsilon,\eta}(\xi,x')$. For its imaginary part, we have, by changing variable $\omega Rr(x'\cdot\xi')=t$, that

$$\begin{aligned}\Im I_{\varepsilon,\eta}(\xi,x')=&-\int_\varepsilon^\eta \frac{\sin(\omega Rr(x'\cdot\xi'))}{r}dr\\ =&-\operatorname{sgn}(\omega)\operatorname{sgn}(x'\cdot\xi')\int_{|\omega|R\varepsilon|x'\cdot\xi'|}^{|\omega|R\eta|x'\cdot\xi'|}\frac{\sin t}{t}dt\end{aligned}$$

is uniformly bounded and converges to

$$-\operatorname{sgn}(\omega)\operatorname{sgn}(x'\cdot\xi')\int_0^\infty \frac{\sin t}{t}dt=-\frac{\pi}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(x'\cdot\xi'),$$

as $\varepsilon\to 0$ and $\eta\to\infty$.

For its real part, since $\cos r$ is an even function, we have

$$\Re I_{\varepsilon,\eta}(\xi,x')=\int_\varepsilon^\eta [\cos(|\omega|Rr|x'\cdot\xi'|)-\cos(|\omega|Rr)]\frac{dr}{r}.$$

If $x'\cdot\xi'=\pm1$, then $\Re I_{\varepsilon,\eta}(\xi,x')=0$. Now we assume $0<\varepsilon<1<\eta$. For the case $x'\cdot\xi'\neq\pm1$, we get the absolute value of its real part

$$\begin{aligned}\left|\Re I_{\varepsilon,\eta}(\xi,x')\right|\leqslant&\left|\int_\varepsilon^1 -2\sin\left(\frac{|\omega|}{2}Rr(|x'\cdot\xi'|+1)\right)\sin\left(\frac{|\omega|}{2}Rr(|x'\cdot\xi'|-1)\right)\frac{dr}{r}\right|\\ &+\left|\int_1^\eta \cos\left(|\omega|Rr|x'\cdot\xi'|\right)\frac{dr}{r}-\int_1^\eta\cos\left(|\omega|Rr\right)\frac{dr}{r}\right|\\ \leqslant&\frac{|\omega|^2}{2}R^2(1-|x'\cdot\xi'|^2)\int_\varepsilon^1 rdr\\ &+\left|\int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R\eta|\xi'\cdot x'|}\frac{\cos t}{t}dt-\int_{|\omega|R}^{|\omega|R\eta}\frac{\cos t}{t}dt\right|\\ \leqslant&\frac{|\omega|^2}{4}R^2+I_1.\end{aligned}$$

If $\eta|\xi'\cdot x'|>1$, then we have

$$\begin{aligned}I_1=&\left|\int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R}\frac{\cos t}{t}dt-\int_{|\omega|R\eta|\xi'\cdot x'|}^{|\omega|R\eta}\frac{\cos t}{t}dt\right|\\ \leqslant&\int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R}\frac{dt}{t}+\int_{|\omega|R\eta|\xi'\cdot x'|}^{|\omega|R\eta}\frac{dt}{t}\\ \leqslant&2\ln(1/|\xi'\cdot x'|).\end{aligned}$$

If $0<\eta|\xi'\cdot x'|\leqslant 1$, then

$$I_1\leqslant\int_{|\omega|R|\xi'\cdot x'|}^{|\omega|R/|\xi'\cdot x'|}\frac{dt}{t}\leqslant 2\ln(1/|\xi'\cdot x'|).$$

Thus,

$$\left|\Re I_{\varepsilon,\eta}(\xi,x')\right|\leqslant\frac{|\omega|^2}{4}R^2+2\ln(1/|\xi'\cdot x'|),$$

and so the real part converges as $\varepsilon \to 0$ and $\eta \to \infty$. By the fundamental theorem of calculus, we can write

$$\begin{aligned}
&\int_\varepsilon^\eta \frac{\cos(\lambda r)-\cos(\mu r)}{r}dr = -\int_\varepsilon^\eta\int_\mu^\lambda \sin(tr)dtdr = -\int_\mu^\lambda\int_\varepsilon^\eta \sin(tr)drdt\\
=&\int_\mu^\lambda\int_\varepsilon^\eta \frac{\partial_r\cos(tr)}{t}drdt = \int_\mu^\lambda \frac{\cos(t\eta)-\cos(t\varepsilon)}{t}dt\\
=&\int_{\mu\eta}^{\lambda\eta}\frac{\cos s}{s}ds - \int_\mu^\lambda \frac{\cos(t\varepsilon)}{t}dt = \frac{\sin s}{s}\Big|_{\mu\eta}^{\lambda\eta} + \int_{\mu\eta}^{\lambda\eta}\frac{\sin s}{s^2}ds - \int_\mu^\lambda\frac{\cos(t\varepsilon)}{t}dt\\
\to& 0 - \int_\mu^\lambda \frac{1}{t}dt = -\ln(\lambda/\mu) = \ln(\mu/\lambda), \text{ as } \eta\to\infty,\ \varepsilon\to 0,
\end{aligned}$$

by the Lebesgue 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 $\lambda = |\omega|R|x'\cdot\xi'|$, and $\mu = |\omega|R$. So

$$\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}} \Re(I_{\varepsilon,\eta}(\xi,x')) = \int_0^\infty [\cos|\omega|Rr(x'\cdot\xi') - \cos|\omega|Rr]\frac{dr}{r} = \ln(1/|x'\cdot\xi'|).$$

Next, we need to show (i) for all $\xi\in\mathbb{R}^n$. By the properties of $I_{\varepsilon,\eta}$ just proved, we have

$$\begin{aligned}
\left(\frac{|\omega|}{2\pi}\right)^{-n/2}|\widehat{K_{\varepsilon,\eta}}(\xi)| \leqslant& \int_{S^{n-1}}\left[4+\frac{|\omega|^2}{4}R^2 + 2\ln(1/|\xi'\cdot x'|)\right]|\Omega(x')|d\sigma(x') \qquad (4.3.6)\\
\leqslant&(4+\frac{|\omega|^2}{4}R^2)\|\Omega\|_{L^1(S^{n-1})} + 2\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x').
\end{aligned}$$

For $n=1$, we have $S^0=\{-1,1\}$ and then $\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x') = 2\ln 1 = 0$. For $n\geqslant 2$, if we can show

$$\int_{S^{n-1}}\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x')d\sigma(\xi') < \infty,$$

then, $\left(\frac{|\omega|}{2\pi}\right)^{-n/2}|\widehat{K_{\varepsilon,\eta}}(\xi)|$ is finite a.e. We can pick an orthogonal matrix A such that $Ae_1 = x'$, and so by changes of variables and using the notation $\bar{y} = (y_2, y_3, ..., y_n)$,

$$\begin{aligned}
&\int_{S^{n-1}}\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x')d\sigma(\xi')\\
=&\int_{S^{n-1}}\int_{S^{n-1}}\ln(1/|\xi'\cdot Ae_1|)d\sigma(\xi')|\Omega(x')|d\sigma(x')\\
=&\int_{S^{n-1}}\int_{S^{n-1}}\ln(1/|e_1\cdot A^{-1}\xi'|)d\sigma(\xi')|\Omega(x')|d\sigma(x')\\
\overset{A^{-1}\xi'=y}{=\!=\!=\!=}&\|\Omega\|_{L^1(\Omega)}\int_{S^{n-1}}\ln(1/|y_1|)d\sigma(y).
\end{aligned}$$

If for $\phi_j\in[0,\pi]$ $(j=1,\cdots,n-2)$ and $\phi_{n-1}\in[0,2\pi]$, let

$$\begin{aligned}
y_1 =&\cos\phi_1\\
y_2 =&\sin\phi_1\cos\phi_2\\
y_3 =&\sin\phi_1\sin\phi_2\cos\phi_3\\
&\vdots\\
y_{n-1} =&\sin\phi_1\cdots\sin\phi_{n-2}\cos\phi_{n-1}\\
y_n =&\sin\phi_1\cdots\sin\phi_{n-2}\sin\phi_{n-1},
\end{aligned}$$

then the volume element $d_{S^{n-1}}\sigma(y)$ of the $(n-1)$-sphere is given by

$$\begin{aligned}d_{S^{n-1}}\sigma(y) =&\sin^{n-2}(\phi_1)\sin^{n-3}(\phi_2)\cdots\sin(\phi_{n-2})\,d\phi_1\,d\phi_2\cdots d\phi_{n-1}\\ =&\sin^{n-3}(\phi_1)\sin^{n-3}(\phi_2)\cdots\sin(\phi_{n-2})\,dy_1\,d\phi_2\cdots d\phi_{n-1}\\ =&(1-y_1^2)^{(n-3)/2}dy_1 d_{S^{n-2}}\sigma(\bar{y}),\end{aligned}$$

due to $dy_1=\sin(\phi_1)d\phi_1$ and $\sin\phi_1=\sqrt{1-y_1^2}$. Thus, we get

$$\begin{aligned}&\int_{S^{n-1}}\ln(1/|y_1|)d\sigma(y)\\ =&\int_{-1}^1\ln(1/|y_1|)\int_{S^{n-2}}(1-y_1^2)^{(n-3)/2}d\sigma(\bar{y})dy_1\\ =&\omega_{n-2}\int_{-1}^1\ln(1/|y_1|)(1-y_1^2)^{(n-3)/2}dy_1\\ =&2\omega_{n-2}\int_0^1\ln(1/|y_1|)(1-y_1^2)^{(n-3)/2}dy_1\\ =&2\omega_{n-2}\int_0^{\pi/2}\ln(1/\cos\theta)(\sin\theta)^{n-2}d\theta \quad (\text{let } y_1=\cos\theta)\\ =&2\omega_{n-2}I_2.\end{aligned}$$

For $n\geqslant 3$, we have, by integration by parts,

$$I_2\leqslant\int_0^{\pi/2}\ln(1/\cos\theta)\sin\theta d\theta=\int_0^{\pi/2}\sin\theta d\theta=1.$$

For $n=2$, we have by 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 $I_2=\frac{\pi}{2}\ln 2$. Hence, $\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)|\Omega(x')|d\sigma(x')\leqslant C$ for any $\xi'\in S^{n-1}$.

Thus, we have proved the uniform boundedness of $\widehat{K_{\varepsilon,\eta}}(\xi)$, i.e., (i). In view of the limit of $I_{\varepsilon,\eta}(\xi,x')$ as $\varepsilon\to 0$, $\eta\to\infty$ just proved, and the dominated convergence theorem, we get

$$\left(\frac{|\omega|}{2\pi}\right)^{-n/2}\lim_{\substack{\varepsilon\to 0\\ \eta\to\infty}}\widehat{K_{\varepsilon,\eta}}(\xi)=m(\xi),\quad \text{a.e.}$$

By the Plancherel theorem, if $f\in L^2(\mathbb{R}^n)$, $K_{\varepsilon,\eta}*f$ converges in L^2 norm as $\varepsilon\to 0$ and $\eta\to\infty$, and the Fourier transform of this limit is $m(\xi)\widehat{f}(\xi)$. 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 4.3.4. 1) In the theorem, the condition that Ω is mean value zero on S^{n-1} is necessary and cannot be neglected. Since in the estimate

$$\int_{\mathbb{R}^n} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy = \left[\int_{|y|\leqslant 1} + \int_{|y|>1}\right] \frac{\Omega(y)}{|y|^n} f(x-y)dy,$$

the main difficulty lies in the first integral. For instance, if we assume $\Omega(x) \equiv 1 \in L^1(S^{n-1})$, $f(x) = \chi_{|x|\leqslant 1}(x) \in L^2(\mathbb{R}^n)$, 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., $\Im I_{\varepsilon,\eta}(\xi, x')$, we required only $\int_{S^{n-1}} |\Omega_o(x')|d\sigma(x') < \infty$for the odd part; and for the even part, the proof requires the uniform boundedness of

$$\int_{S^{n-1}} |\Omega_e(x')| \ln(1/|\xi' \cdot x'|)d\sigma(x').$$

This observation is suggestive of certain generalizations of Theorem 4.2.4, see [Ste70, §6.5, p.49–50]. In addition, $\ln(1/|\xi' \cdot x'|)$ is not bounded but any power (> 1) of it is integrable, we immediately get the following corollary.

Corollary 4.3.5.

Given a function Ω with mean value zero on S^{n-1}, suppose that $\Omega_o \in L^1(S^{n-1})$ and $\Omega_e \in L^q(S^{n-1})$ for some $q > 1$. Then, the Fourier transform of p.v. $\Omega(x')/|x|^n$ is bounded.

If $\Omega \in L^1(S^{n-1})$ is odd, i.e., $\Omega(-x) = -\Omega(x)$ for all $x \in S^{n-1}$, then

$$\int_{S^{n-1}} \Omega(x) \ln(1/|\xi \cdot x|)d\sigma(x) = 0$$

for all $\xi \in S^{n-1}$. Thus, $m \in L^\infty(\mathbb{R}^n)$ by Theorem 4.3.3. We have the following result by Theorem 2.5.6.

Corollary 4.3.6.

Given an odd function $\Omega \in L^1(S^{n-1})$, 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.

§ 4.4 Riesz transforms and spherical harmonics

§ 4.4.1 Riesz transforms

We look for the operators in $\mathbb{R}^n$ which have the analogous structural characterization as the Hilbert transform. We begin by making a few remarks about the interaction of rotations with the n-dimensional Fourier transform. We shall need the following elementary observation.

Let ρ denote any rotation about the origin in $\mathbb{R}^n$. Denote also by ρ its induced action on functions, $\rho(f)(x) = f(\rho x)$. Then

$$\begin{aligned}\mathscr{F}(\rho(f))(\xi) = &\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} f(\rho x)dx = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i \rho^{-1}y\cdot\xi} f(y)dy \\ = &\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i y\cdot\rho\xi} f(y)dy = \widehat{f}(\rho\xi) = \rho(\widehat{f})(\xi),\end{aligned}$$

that is,

$$\mathscr{F}\rho = \rho\mathscr{F}.$$

Let $\ell(x) = (\ell_1(x), \ell_2(x), ..., \ell_n(x))$ be an n-tuple of functions defined on $\mathbb{R}^n$. For any rotation ρ about the origin, write $\rho = (\rho_{jk})$ for its matrix realization. Suppose that ℓ transforms like a vector. Symbolically this can be written as

$$\ell(\rho x) = \rho(\ell(x)),$$

or more explicitly

$$\ell_j(\rho x) = \sum_k \rho_{jk}\ell_k(x), \quad \text{for every rotation } \rho. \tag{4.4.1}$$

Lemma 4.4.1.

Suppose ℓ is homogeneous of degree 0, i.e., $\ell(\varepsilon x) = \ell(x)$, for $\varepsilon > 0$. If ℓ transforms according to (4.4.1) then $\ell(x) = c\frac{x}{|x|}$ for some constant c; that is

$$\ell_j(x) = c\frac{x_j}{|x|}. \tag{4.4.2}$$

Proof. It suffices to consider $x \in S^{n-1}$ due to the homogeneousness of degree 0 for ℓ. Now, let e_1, e_2, ..., e_n denote the usual unit vectors along the axes. Set $c = \ell_1(e_1)$. We can see that $\ell_j(e_1) = 0$, if $j \neq 1$.

In fact, we take a rotation arbitrarily such that e_1 fixed under the acting of ρ, i.e., $\rho e_1 = e_1$. Thus, we also have $e_1 = \rho^{-1}\rho e_1 = \rho^{-1}e_1 = \rho^\top e_1$. From $\rho e_1 = \rho^\top e_1 = e_1$, we get $\rho_{11} = 1$ and $\rho_{1k} = \rho_{j1} = 0$ for $k \neq 1$ and $j \neq 1$. So $\rho = \begin{pmatrix} 1 & 0 \\ 0 & A\end{pmatrix}$. Because $\begin{pmatrix} 1 & 0 \\ 0 & A\end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & A^{-1}\end{pmatrix}$ and $\rho^{-1} = \rho^\top$, we obtain $A^{-1} = A^\top$ and $\det A = 1$, i.e., A is a rotation in $\mathbb{R}^{n-1}$. On the other hand, by (4.4.1), we get $\ell_j(e_1) = \sum_{k=2}^n \rho_{jk}\ell_k(e_1)$ for $j = 2, ..., n$. That is, the $n-1$ dimensional vector $(\ell_2(e_1), \ell_3(e_1), \cdots, \ell_n(e_1))$ is left fixed by all the rotations on this $n-1$ dimensional vector space. Thus, we have to take $\ell_2(e_1) = \ell_3(e_1) = \cdots = \ell_n(e_1) = 0$.

Inserting again in (4.4.1) gives $\ell_j(\rho e_1) = \rho_{j1}\ell_1(e_1) = c\rho_{j1}$. If we take a rotation such that $\rho e_1 = x$, then we have $\rho_{j1} = x_j$, so $\ell_j(x) = cx_j$, ($|x| = 1$), which proves the lemma. ■

We now define the n ***Riesz transforms***. For $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p < \infty$, we set

$$R_jf(x) = \lim_{\varepsilon\to 0} c_n \int_{|y|\geqslant\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)dy, \quad j = 1, ..., n, \tag{4.4.3}$$

with $c_n = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}$ where $1/c_n = \frac{\pi^{(n+1)/2}}{\Gamma((n+1)/2)}$ is half the surface area of the unit sphere S^n of $\mathbb{R}^{n+1}$. Thus, R_j is defined by the kernel $K_j(x) = \frac{\Omega_j(x)}{|x|^n}$, and $\Omega_j(x) = c_n \frac{x_j}{|x|}$.

Next, we derive the multipliers which correspond to the Riesz transforms, and which in fact justify their definition. Denote

$$\Omega(x) = (\Omega_1(x), \Omega_2(x), ..., \Omega_n(x)), \text{ and } m(\xi) = (m_1(\xi), m_2(\xi), ..., m_n(\xi)).$$

Let us recall the formula (4.3.5), i.e.,

$$m(\xi) = \int_{S^{n-1}} \Phi(\xi \cdot x)\Omega(x) d\sigma(x), \quad |\xi| = 1, \tag{4.4.4}$$

with $\Phi(t) = -\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(t) + \ln |1/t|$. For any rotation ρ, since Ω commutes with any rotations, i.e., $\Omega(\rho x) = \rho(\Omega(x))$, we have, by changes of variables,

$$\begin{aligned}
\rho(m(\xi)) =& \int_{S^{n-1}} \Phi(\xi \cdot x)\rho(\Omega(x)) d\sigma(x) = \int_{S^{n-1}} \Phi(\xi \cdot x)\Omega(\rho x) d\sigma(x) \\
=& \int_{S^{n-1}} \Phi(\xi \cdot \rho^{-1} y)\Omega(y) d\sigma(y) = \int_{S^{n-1}} \Phi(\rho\xi \cdot y)\Omega(y) d\sigma(y) \\
=& m(\rho\xi).
\end{aligned}$$

Thus, m commutes with rotations and so m satisfies (4.4.1). However, the m_j are each homogeneous of degree 0, so Lemma 4.4.1 shows that $m_j(\xi) = c\frac{\xi_j}{|\xi|}$, with

$$\begin{aligned}
c =& m_1(e_1) = \int_{S^{n-1}} \Phi(e_1 \cdot x)\Omega_1(x) d\sigma(x) \\
=& \int_{S^{n-1}} [-\frac{\pi i}{2} \operatorname{sgn}(\omega) \operatorname{sgn}(x_1) + \ln |1/x_1|] c_n x_1 d\sigma(x) \\
=& - \operatorname{sgn}(\omega) \frac{\pi i}{2} c_n \int_{S^{n-1}} |x_1| d\sigma(x) \text{ (the 2nd is 0 since it is odd w.r.t. } x_1) \\
=& - \operatorname{sgn}(\omega) \frac{\pi i}{2} \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \frac{2\pi^{(n-1)/2}}{\Gamma((n+1)/2)} = - \operatorname{sgn}(\omega) i.
\end{aligned}$$

Here we have used the fact $\int_{S^{n-1}} |x_1| d\sigma(x) = 2\pi^{(n-1)/2}/\Gamma((n+1)/2)$. Therefore, we obtain

$$\widehat{R_j f}(\xi) = - \operatorname{sgn}(\omega) i \frac{\xi_j}{|\xi|} \hat{f}(\xi), \quad j = 1, ..., n. \tag{4.4.5}$$

This identity and Plancherel's theorem also imply the following "unitary" character of the Riesz transforms

$$\sum_{j=1}^{n} \|R_j f\|_2^2 = \|f\|_2^2.$$

By $m(\rho\xi) = \rho(m(\xi))$ proved above, we have $m_j(\rho\xi) = \sum_k \rho_{jk} m_k(\xi)$ for any rotation ρ and then $m_j(\rho\xi)\hat{f}(\xi) = \sum_k \rho_{jk} m_k(\xi)\hat{f}(\xi)$. Taking the inverse Fourier transform, it follows

$$\begin{aligned}
\mathscr{F}^{-1}(m_j(\rho\xi)\hat{f}(\xi)) =& \mathscr{F}^{-1}(\sum_k \rho_{jk} m_k(\xi)\hat{f}(\xi)) \\
=& \sum_k \rho_{jk} \mathscr{F}^{-1}(m_k(\xi)\hat{f}(\xi)) = \sum_k \rho_{jk} R_k f.
\end{aligned}$$

But by changes of variables, we have

$$\begin{aligned}
&\mathscr{F}^{-1}(m_j(\rho\xi)\hat{f}(\xi)) \\
=& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{\omega i x \cdot \xi} m_j(\rho\xi)\hat{f}(\xi) d\xi
\end{aligned}$$

$$
\begin{aligned}
&=\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} e^{\omega i\rho x\cdot\eta}m_j(\eta)\widehat{f}(\rho^{-1}\eta)d\eta\\
&=(\mathscr{F}^{-1}(m_j(\xi)\widehat{f}(\rho^{-1}\xi)))(\rho x)=\rho\mathscr{F}^{-1}(m_j(\xi)\widehat{f}(\rho^{-1}\xi))(x)\\
&=\rho R_j\rho^{-1}f,
\end{aligned}
$$

since the Fourier transform commutes with rotations. Therefore, it reaches

$$
\rho R_j\rho^{-1}f=\sum_k \rho_{jk}R_kf, \tag{4.4.6}
$$

which is the statement that under rotations in $\mathbb{R}^n$, the Riesz operators transform in the same manner as the components of a vector.

We have the following characterization of Riesz transforms.

Proposition 4.4.2.

Let $T=(T_1,T_2,...,T_n)$ be an n-tuple of bounded linear transforms on $L^2(\mathbb{R}^n)$. Suppose
(a) Each T_j commutes with translations of $\mathbb{R}^n$;
(b) Each T_j commutes with dilations of $\mathbb{R}^n$;
(c) For every rotation $\rho=(\rho_{jk})$ of $\mathbb{R}^n$, $\rho T_j\rho^{-1}f=\sum_k\rho_{jk}T_kf$.
Then the T_j is a constant multiple of the Riesz transforms, i.e., there exists a constant c such that $T_j=cR_j$, $j=1,...,n$.

Proof. All the elements of the proof have already been discussed. We bring them together.

(i) Since the T_j is bounded linear on $L^2(\mathbb{R}^n)$ and commutes with translations, by Theorem 2.5.6 they can be each realized by bounded multipliers m_j, i.e., $\widehat{T_jf}=m_j\widehat{f}$.

(ii) Since the T_j commutes with dilations, i.e., $T_j\delta^\varepsilon f=\delta^\varepsilon T_jf$, in view of Proposition 2.1.2, we see that

$$
\widehat{T_j\delta^\varepsilon f}=m_j(\xi)\widehat{\delta^\varepsilon f}=m_j(\xi)\varepsilon^{-n}\delta^{\varepsilon^{-1}}\widehat{f}(\xi)=m_j(\xi)\varepsilon^{-n}\widehat{f}(\xi/\varepsilon)
$$

and

$$
\widehat{\delta^\varepsilon T_jf}=\varepsilon^{-n}\delta^{\varepsilon^{-1}}\widehat{T_jf}=\varepsilon^{-n}\delta^{\varepsilon^{-1}}(m_j\widehat{f})=\varepsilon^{-n}m_j(\xi/\varepsilon)\widehat{f}(\xi/\varepsilon),
$$

which imply $m_j(\xi)=m_j(\xi/\varepsilon)$ or equivalently $m_j(\varepsilon\xi)=m_j(\xi)$, $\varepsilon>0$; that is, each m_j is homogeneous of degree 0.

(iii) Finally, assumption (c) has a consequence by taking the Fourier transform, i.e., the relation (4.4.1), and so by Lemma 4.4.1, we can obtain the desired conclusion. ■

For the L^p boundedness, we have the following.

Theorem 4.4.3.

The Riesz transforms R_j, $j=1,\cdots,n$, are of weak-type $(1,1)$ and of strong-type (p,p) for $1<p<\infty$.

Proof. It suffices to show that $K_j(x)=c_n\,\text{p.v.}\,\frac{x_j}{|x|^{n+1}}$ satisfy the hypotheses of Theorem 4.2.1 for $j=1,\cdots,n$, respectively. Clearly K_j coincides with a locally integrable function on $\mathbb{R}^n\setminus\{0\}$. Moreover, by (4.4.5),

$$
\left(\frac{|\omega|}{2\pi}\right)^{-n/2}\widehat{K_j}(\xi)=-\operatorname{sgn}(\omega)i\frac{\xi_j}{|\xi|},\quad j=1,...,n,
$$

which is clearly bounded.

Finally, we have on $\mathbb{R}^n \setminus \{0\}$

$$
\begin{aligned}
|\nabla K_j(x)| \leqslant & c_n \frac{1}{|x|^{n+1}} + c_n(n+1)\frac{|x_j||x|}{|x|^{n+3}} \\
\leqslant & \frac{C}{|x|^{n+1}},
\end{aligned}
$$

that is, (4.2.5) is satisfied. Thus, by Theorem 4.2.1 we obtain the desired results. ■

One of the important applications of the Riesz transforms is that they can be used to mediate between various combinations of partial derivatives of a function.

Example 4.4.4. (Schauder estimate) Suppose $f \in \mathcal{C}_c^2(\mathbb{R}^n)$. Let $\Delta f = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j^2}$. Then we have the a priori bound

$$
\left\| \frac{\partial^2 f}{\partial x_j \partial x_k} \right\|_p \leqslant A_p \|\Delta f\|_p, \quad 1 < p < \infty. \tag{4.4.7}
$$

Proof. Since $\widehat{\partial_{x_j} f}(\xi) = \omega i \xi_j \widehat{f}(\xi)$, we have

$$
\begin{aligned}
\widehat{\frac{\partial^2 f}{\partial x_j \partial x_k}}(\xi) =& -\omega^2 \xi_j \xi_k \widehat{f}(\xi) \\
=& -\left(-\operatorname{sgn}(\omega)\frac{i\xi_j}{|\xi|}\right)\left(-\operatorname{sgn}(\omega)\frac{i\xi_k}{|\xi|}\right)(-\omega^2|\xi|^2)\widehat{f}(\xi) \\
=& -\widehat{R_j R_k \Delta f}(\xi).
\end{aligned}
$$

Thus, $\frac{\partial^2 f}{\partial x_j \partial x_k} = -R_j R_k \Delta f$. By the L^p boundedness of the Riesz transforms, we have the desired result. ■

Example 4.4.5. Suppose $f \in \mathcal{C}_c^1(\mathbb{R}^2)$. Then we have the a priori bound

$$
\left\| \frac{\partial f}{\partial x_1} \right\|_p + \left\| \frac{\partial f}{\partial x_2} \right\|_p \leqslant A_p \left\| \frac{\partial f}{\partial x_1} + i \frac{\partial f}{\partial x_2} \right\|_p, \quad 1 < p < \infty.
$$

Proof. The proof is similar to the previous one. Indeed, we have

$$
\begin{aligned}
\widehat{\partial_{x_j} f}(\xi) =& \omega i \xi_j \widehat{f}(\xi) = \omega \frac{i\xi_j}{|\xi|}|\xi|\widehat{f}(\xi) = \omega \frac{i\xi_j}{|\xi|}\frac{\xi_1^2+\xi_2^2}{|\xi|}\widehat{f}(\xi) \\
=& \omega \frac{i\xi_j}{|\xi|}\frac{(\xi_1 - i\xi_2)(\xi_1 + i\xi_2)}{|\xi|}\widehat{f}(\xi) \\
=& -\frac{-\operatorname{sgn}(\omega) i\xi_j}{|\xi|}\frac{-\operatorname{sgn}(\omega) i(\xi_1 - i\xi_2)}{|\xi|}\widehat{\partial_{x_1} f + i\partial_{x_2} f}(\xi) \\
=& -\widehat{R_j(R_1 - iR_2)(\partial_{x_1} f + i\partial_{x_2} f}(\xi).
\end{aligned}
$$

That is, $\partial_{x_j} f = -R_j(R_1 - iR_2)(\partial_{x_1} f + i\partial_{x_2} f)$. Also by the L^p boundedness of the Riesz transforms, we can obtain the result. ■

We shall now tie together the Riesz transforms and the theory of harmonic functions, more particularly Poisson integrals. Since we are interested here mainly in the formal aspects we shall restrict ourselves to the L^2 case. For L^p case, one can see the further results in [Ste70, §4.3 and §4.4, p.78].

Example 4.4.6. Let f and f_1, ..., f_n all belong to $L^2(\mathbb{R}^n)$, and let their respective Poisson integrals be $u_0(x,y) = P_y * f$, $u_1(x,y) = P_y * f_1$, ..., $u_n(x,y) = P_y * f_n$. Then a necessary and sufficient condition of

$$
f_j = R_j(f), \quad j = 1, ..., n, \tag{4.4.8}
$$

is that the following generalized Cauchy-Riemann equations hold:

$$\begin{cases} \displaystyle\sum_{j=0}^{n} \frac{\partial u_j}{\partial x_j} = 0, \\ \displaystyle\frac{\partial u_j}{\partial x_k} = \frac{\partial u_k}{\partial x_j}, \qquad j \neq k, \quad \text{with } x_0 = y. \end{cases} \tag{4.4.9}$$

Remark 4.4.7. At least locally, the system (4.4.9) is equivalent with the existence of a harmonic function g of the $n+1$ variables, such that $u_j = \frac{\partial g}{\partial x_j}$, $j = 0, 1, 2, \ldots, n$.

Proof. Suppose $f_j = R_j f$, then $\widehat{f_j}(\xi) = -\operatorname{sgn}(\omega)\frac{i\xi_j}{|\xi|}\widehat{f}(\xi)$, and so by (4.1.2)

$$u_j(x,y) = -\operatorname{sgn}(\omega)\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}\widehat{f}(\xi)\frac{i\xi_j}{|\xi|}e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi, \quad j = 1, \ldots, n,$$

and

$$u_0(x,y) = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}\widehat{f}(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi.$$

The equation (4.4.9) can then be immediately verified by differentiation under the integral sign, which is justified by the rapid convergence of the integrals in question.

Conversely, let $u_j(x,y) = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n}\widehat{f_j}(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi$, $j = 0, 1, \ldots, n$ with $f_0 = f$. Then the fact that $\frac{\partial u_0}{\partial x_j} = \frac{\partial u_j}{\partial x_0} = \frac{\partial u_j}{\partial y}$, $j = 1, \ldots, n$, and Fourier inversion theorem, show that

$$\omega i\xi_j\widehat{f_0}(\xi)e^{-|\omega\xi|y} = -|\omega\xi|\widehat{f_j}(\xi)e^{-|\omega\xi|y},$$

therefore $\widehat{f_j}(\xi) = -\operatorname{sgn}(\omega)\frac{i\xi_j}{|\xi|}\widehat{f_0}(\xi)$, and so $f_j = R_j f_0 = R_j f$ for $j = 1, \ldots, n$. ■

§ 4.4.2 Spherical harmonics and higher Riesz transforms

Consider now an open set $\Omega \subset \mathbb{R}^n$ and suppose u is a harmonic function (i.e., $\Delta u = 0$) within Ω. We next derive the important mean-value formulas, which declare that $u(x)$ equals both the average of u over the sphere $\partial B(x,r)$ and the average of u over the entire ball $B(x,r)$, provided $B(x,r) \subset \Omega$.

Theorem 4.4.8: Mean-value formula for harmonic functions

If $u \in \mathcal{C}^2(\Omega)$ is harmonic, then for each ball $B(x,r) \subset \Omega$,

$$u(x) = \frac{1}{\mu(\partial B(x,r))}\int_{\partial B(x,r)} u(y)d\sigma(y) = \frac{1}{\mu(B(x,r))}\int_{B(x,r)} u(y)dy.$$

Proof. Denote

$$f(r) = \frac{1}{\mu(\partial B(x,r))}\int_{\partial B(x,r)} u(y)d\sigma(y) = \frac{1}{\omega_{n-1}}\int_{S^{n-1}} u(x+rz)d\sigma(z).$$

Obviously,

$$f'(r) = \frac{1}{\omega_{n-1}}\int_{S^{n-1}}\sum_{j=1}^{n}\partial_{x_j}u(x+rz)z_j d\sigma(z) = \frac{1}{\omega_{n-1}}\int_{S^{n-1}}\frac{\partial u}{\partial\nu}(x+rz)d\sigma(z),$$

where $\frac{\partial}{\partial \nu}$ denotes the differentiation w.r.t. the outward normal. Thus, by changes of variable

$$f'(r) = \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B(x,r)} \frac{\partial u}{\partial \nu}(y) d\sigma(y).$$

By Stokes theorem, we get

$$f'(r) = \frac{1}{\omega_{n-1} r^{n-1}} \int_{B(x,r)} \Delta u(y) dy = 0.$$

Thus, $f(r) = \text{const}$. Since $\lim_{r\to 0} f(r) = u(x)$, hence, $f(r) = u(x)$.

Next, observe that our employing polar coordinates gives, by the first identity proved just now, that

$$\begin{aligned}\int_{B(x,r)} u(y) dy = &\int_0^r \left(\int_{\partial B(x,s)} u(y) d\sigma(y) \right) ds = \int_0^r \mu(\partial B(x,s)) u(x) ds \\ =& u(x) \int_0^r n V_n s^{n-1} ds = V_n r^n u(x).\end{aligned}$$

This completes the proof. ■

Theorem 4.4.9: Converse to mean-value property

If $u \in \mathcal{C}^2(\Omega)$ satisfies

$$u(x) = \frac{1}{\mu(\partial B(x,r))} \int_{\partial B(x,r)} u(y) d\sigma(y)$$

for each ball $B(x,r) \subset \Omega$, then u is harmonic.

Proof. If $\Delta u \not\equiv 0$, then there exists some ball $B(x,r) \subset \Omega$ such that, say, $\Delta u > 0$ within $B(x,r)$. But then for f as above,

$$0 = f'(r) = \frac{1}{r^{n-1} \omega_{n-1}} \int_{B(x,r)} \Delta u(y) dy > 0,$$

is a contradiction. ■

We return to the consideration of special transforms of the form

$$Tf(x) = \lim_{\varepsilon \to 0} \int_{|y| \geqslant \varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y) dy, \tag{4.4.10}$$

where Ω is homogeneous of degree 0 with mean value zero on S^{n-1}.

We have already considered the example, i.e., the case of Riesz transforms, $\Omega_j(y) = c\frac{y_j}{|y|}$, $j = 1, ..., n$. For $n = 1$, $\Omega(y) = c \operatorname{sgn} y$, this is the only possible case, i.e., the Hilbert transform. To study the matter further for $n > 1$, we recall the expression

$$m(\xi) = \int_{S^{n-1}} \Lambda(y \cdot \xi) \Omega(y) d\sigma(y), \quad |\xi| = 1$$

where m is the multiplier arising from the transform (4.4.10).

We have already remarked that the mapping $\Omega \to m$ commutes with rotations. We shall therefore consider the functions on the sphere S^{n-1} (more particularly the space $L^2(S^{n-1})$) from the point of view of its decomposition under the action of rotations. As is well known, this decomposition is in terms of the spherical harmonics, and it is with a brief review of their properties that we begin.

We fix our attention, as always, on $\mathbb{R}^n$, and we shall consider polynomials in $\mathbb{R}^n$ which are also harmonic.

Definition 4.4.10.

Let $\alpha = (\alpha_1, ..., \alpha_n)$ be a multi-index, $|\alpha| = \sum_{j=1}^n \alpha_j$ and $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Let $\mathscr{P}_k$ denote the linear space of all homogeneous polynomials of degree k, i.e.,

$$\mathscr{P}_k := \left\{ P(x) = \sum a_\alpha x^\alpha : |\alpha| = k \right\}.$$

Each such polynomial corresponds its dual object, the differential operator $P(\partial_x) = \sum a_\alpha \partial_x^\alpha$, where $\partial_x^\alpha = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}$. On $\mathscr{P}_k$, we define a positive inner product $\langle P, Q \rangle = P(\partial_x)\bar{Q}$. Note that two distinct monomials x^α and $x^{\alpha'}$ in $\mathscr{P}_k$ are orthogonal w.r.t. it, since there exists at least one i such that $\alpha_i \geqslant \alpha_i'$, then $\partial_{x_i}^{\alpha_i} x_i^{\alpha_i'} = 0$. $\langle P, P \rangle = \sum |a_\alpha|^2 \alpha!$ where $\alpha! = (\alpha_1!) \cdots (\alpha_n!)$.

Definition 4.4.11.

We define $\mathscr{H}_k$ to be the linear space of homogeneous polynomials of degree k which are harmonic: the ***solid spherical harmonics of degree*** k. That is,

$$\mathscr{H}_k := \{P(x) \in \mathscr{P}_k : \Delta P(x) = 0\}.$$

It will be convenient to restrict these polynomials to S^{n-1}, and then to define the standard inner product,

$$(P, Q) = \int_{S^{n-1}} P(x)\overline{Q(x)} d\sigma(x).$$

For a function f on S^{n-1}, we define the spherical Laplacean Δ_S by

$$\Delta_S f(x) = \Delta f(x/|x|),$$

where $f(x/|x|)$ is the degree zero homogeneous extension of the function f to $\mathbb{R}^n \setminus \{0\}$, and Δ is the Laplacian of the Euclidean space.

Proposition 4.4.12.

We have the following properties.
(1) The finite dimensional spaces $\{\mathscr{H}_k\}_{k=0}^\infty$ are mutually orthogonal.
(2) Every homogeneous polynomial $P \in \mathscr{P}_k$ can be written in the form $P = P_1 + |x|^2 P_2$, where $P_1 \in \mathscr{H}_k$ and $P_2 \in \mathscr{P}_{k-2}$.
(3) Let H_k denote the linear space of restrictions of $\mathscr{H}_k$ to the unit sphere.[a] The elements of H_k are the surface spherical harmonics of degree k, i.e.,

$$H_k = \{P(x) \in \mathscr{H}_k : |x| = 1\}.$$

Then $L^2(S^{n-1}) = \sum_{k=0}^\infty H_k$. Here the L^2 space is taken w.r.t. usual measure, and the infinite direct sum is taken in the sense of Hilbert space theory. That is, if $f \in L^2(S^{n-1})$, then f has the development

$$f(x) = \sum_{k=0}^\infty Y_k(x), \quad Y_k \in H_k, \tag{4.4.11}$$

where the convergence is in the $L^2(S^{n-1})$ norm, and

$$\int_{S^{n-1}} |f(x)|^2 d\sigma(x) = \sum_k \int_{S^{n-1}} |Y_k(x)|^2 d\sigma(x).$$

(4) If $Y_k \in H_k$, then $\Delta_S Y_k(x) = -k(k+n-2)Y_k(x)$.

(5) Suppose f has the development (4.4.11). Then f (after correction on a set of measure zero, if necessary) is indefinitely differentiable on S^{n-1} (i.e., $f \in \mathcal{C}^\infty(S^{n-1})$) if and only if

$$\int_{S^{n-1}} |Y_k(x)|^2 d\sigma(x) = O(k^{-N}), \quad \text{as } k \to \infty, \text{ for each fixed } N. \tag{4.4.12}$$

[a]Sometimes, in order to emphasize the distribution between $\mathscr{H}_k$ and H_k, the members of H_k are referred to as the ***surface spherical harmonics***.

Proof. (1) If $P \in \mathscr{P}_k$, i.e., $P(x) = \sum a_\alpha x^\alpha$ with $|\alpha| = k$, then

$$\sum_{j=1}^n x_j \partial_{x_j} P = \sum_{j=1}^n x_j \sum a_\alpha \alpha_j x_1^{\alpha_1} \cdots x_j^{\alpha_j - 1} \cdots x_n^{\alpha_n} = \sum_{j=1}^n \alpha_j \sum a_\alpha x^\alpha = kP.$$

On S^{n-1}, it follows $kP = \frac{\partial P}{\partial \nu}$ where $\frac{\partial}{\partial \nu}$ denotes differentiation w.r.t. the outward normal vector. Thus, for $P \in \mathscr{H}_k$, and $Q \in \mathscr{H}_j$, then by Green's formula

$$\begin{aligned}(k-j)\int_{S^{n-1}} P\bar{Q} d\sigma(x) =& \int_{S^{n-1}} \left(\bar{Q}\frac{\partial P}{\partial \nu} - P\frac{\partial \bar{Q}}{\partial \nu}\right) d\sigma(x) \\ =& \int_{|x|\leqslant 1} [\bar{Q}\Delta P - P\Delta\bar{Q}] dx = 0,\end{aligned}$$

where Δ is the Laplacean on $\mathbb{R}^n$.

(2) Let $|x|^2\mathscr{P}_{k-2}$ be the subspace of $\mathscr{P}_k$ of all polynomials of the form $|x|^2 P_2$ where $P_2 \in \mathscr{P}_{k-2}$. Then its orthogonal complement w.r.t. $\langle\cdot,\cdot\rangle$ is exactly $\mathscr{H}_k$. In fact, P_1 is in this orthogonal complement if and only if $\langle |x|^2 P_2, P_1\rangle = 0$ for all P_2. But $\langle |x|^2 P_2, P_1\rangle = (P_2(\partial_x)\Delta)\overline{P_1} = \langle P_2, \Delta P_1\rangle$, so $\Delta P_1 = 0$ and thus $\mathscr{P}_k = \mathscr{H}_k \oplus |x|^2\mathscr{P}_{k-2}$, which proves the conclusion. In addition, we have for $P \in \mathscr{P}_k$

$$P(x) = P_k(x) + |x|^2 P_{k-2}(x) + \cdots + \begin{cases} |x|^k P_0(x), & k \text{ even}, \\ |x|^{k-1} P_1(x), & k \text{ odd}, \end{cases}$$

where $P_j \in \mathscr{H}_j$ by noticing that $\mathscr{P}_j = \mathscr{H}_j$ for $j = 0, 1$.

(3) By the further result in (2), if $|x| = 1$, then we have

$$P(x) = P_k(x) + P_{k-2}(x) + \cdots . + \begin{cases} P_0(x), & k \text{ even}, \\ P_1(x), & k \text{ odd}, \end{cases}$$

with $P_j \in \mathscr{H}_j$. That is, the restriction of any polynomial on the unit sphere is a finite linear combination of spherical harmonics. Since the restriction of polynomials is dense in $L^2(S^{n-1})$ (see [SW71, Corollary 2.3, p.141]) by the Weierstrass approximation theorem, the conclusion (4.4.11) is established.

(4) For $|x| = 1$, we have

$$\begin{aligned}\Delta_S Y_k(x) =& \Delta(|x|^{-k} Y_k(x)) = |x|^{-k}\Delta Y_k + \Delta(|x|^{-k})Y_k + 2\nabla(|x|^{-k})\cdot\nabla Y_k \\ =& (k^2 + (2-n)k)|x|^{-k-2} Y_k - 2k^2 |x|^{-k-2} Y_k \\ =& -k(k+n-2)|x|^{k-2} Y_k = -k(k+n-2)Y_k,\end{aligned}$$

since $\sum_{j=1}^n x_j \partial_{x_j} Y_k = kY_k$ for $Y_k \in \mathscr{P}_k$.

(5) Write (4.4.11) as $f(x) = \sum_{k=0}^\infty a_k Y_k^0(x)$, where the Y_k^0 is normalized such that $\int_{S^{n-1}} |Y_k^0(x)|^2 d\sigma(x) = 1$. Our assertion is then equivalent with $a_k = O(k^{-N/2})$, as $k \to \infty$. If f is of class $\mathcal{C}^2$, then an application of Green's formula shows that

$$\int_{S^{n-1}} \Delta_S f \overline{Y_k^0} d\sigma = \int_{S^{n-1}} f \Delta_S \overline{Y_k^0} d\sigma.$$

Thus, if $f \in \mathcal{C}^\infty$, then by (4)

$$\int_{S^{n-1}} \Delta_S^r f \overline{Y_k^0} d\sigma = \int_{S^{n-1}} f \Delta_S^r \overline{Y_k^0} d\sigma = [-k(k+n-2)]^r \int_{S^{n-1}} \sum_{j=0}^{\infty} a_j Y_j^0 \overline{Y_k^0} d\sigma$$
$$=[-k(k+n-2)]^r a_k \int_{S^{n-1}} |Y_k^0|^2 d\sigma = a_k[-k(k+n-2)]^r.$$

So $a_k = O(k^{-2r})$ for every r and therefore (4.4.12) holds.

To prove the converse, from (4.4.12), we have for any $r \in \mathbb{N}$

$$\begin{aligned}\|\Delta_S^r f\|_2^2 =&(\Delta_S^r f, \Delta_S^r f) = \left(\sum_{j=0}^{\infty} \Delta_S^r Y_j(x), \sum_{k=0}^{\infty} \Delta_S^r Y_k(x)\right)\\
=&\left(\sum_{j=0}^{\infty}[-j(j+n-2)]^r Y_j(x), \sum_{k=0}^{\infty}[-k(k+n-2)]^r Y_k(x)\right)\\
=&\sum_{k=0}^{\infty}[-k(k+n-2)]^{2r}(Y_k(x), Y_k(x))\\
=&\sum_{k=0}^{\infty}[-k(k+n-2)]^{2r} O(k^{-N}) \leqslant C,\end{aligned}$$

if we take N large enough. Thus, $f \in \mathcal{C}^\infty(S^{n-1})$. ■

Theorem 4.4.13: Hecke's identity

It holds

$$\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}|x|^2}) = (-i\,\mathrm{sgn}\,(\omega))^k P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2}, \ \forall P_k \in \mathscr{H}_k(\mathbb{R}^n).$$

Proof. That is to prove

$$\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(x)e^{-\omega i x\cdot\xi - \frac{|\omega|}{2}|x|^2} dx = (-i\,\mathrm{sgn}\,(\omega))^k P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2}. \tag{4.4.13}$$

Applying the differential operator $P_k(\partial_\xi)$ to both sides of the identity (cf. Theorem 2.1.9)

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

we obtain

$$\left(\frac{|\omega|}{2\pi}\right)^{n/2} (-\omega i)^k \int_{\mathbb{R}^n} P_k(x)e^{-\omega i x\cdot\xi - \frac{|\omega|}{2}|x|^2} dx = Q(\xi)e^{-\frac{|\omega|}{2}|\xi|^2}.$$

Since $P_k(x)$ is polynomial, it has an obvious analytic continuation $P_k(z)$ to all of $\mathbb{C}^n$. Thus, by changes of variables, we get

$$\begin{aligned}Q(\xi) =&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(x)e^{-\omega i x\cdot\xi - \frac{|\omega|}{2}|x|^2 + \frac{|\omega|}{2}|\xi|^2} dx\\
=&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(x)e^{-\frac{|\omega|}{2}(x + i\,\mathrm{sgn}\,(\omega)\xi)^2} dx\\
=&(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(y - i\,\mathrm{sgn}\,(\omega)\xi)e^{-\frac{|\omega|}{2}|y|^2} dy.\end{aligned}$$

So,

$$Q(i\,\mathrm{sgn}\,(\omega)\xi) =(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} P_k(y+\xi)e^{-\frac{|\omega|}{2}|y|^2}dy$$
$$=(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_0^\infty r^{n-1}e^{-\frac{|\omega|}{2}r^2} \int_{S^{n-1}} P_k(\xi+ry')d\sigma(y')dr.$$

Since P_k is harmonic, it satisfies the mean value property, i.e., Theorem 4.4.8, thus

$$\int_{S^{n-1}} P_k(\xi+ry')d\sigma(y') = \omega_{n-1}P_k(\xi) = P_k(\xi)\int_{S^{n-1}} d\sigma(y').$$

Hence

$$Q(i\,\mathrm{sgn}\,(\omega)\xi) =(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} P_k(\xi) \int_0^\infty r^{n-1}e^{-\frac{|\omega|}{2}r^2} \int_{S^{n-1}} d\sigma(y')dr$$
$$=(-\omega i)^k \left(\frac{|\omega|}{2\pi}\right)^{n/2} P_k(\xi) \int_{\mathbb{R}^n} e^{-\frac{|\omega|}{2}|x|^2}dx = (-\omega i)^k P_k(\xi).$$

Thus, $Q(\xi) = (-\omega i)^k P_k(-i\,\mathrm{sgn}\,(\omega)\xi) = (-\omega i)^k(-i\,\mathrm{sgn}\,(\omega))^k P_k(\xi)$, which proves the theorem. ■

The theorem implies the following generalization, whose interest is that it links the various components of the decomposition of $L^2(\mathbb{R}^n)$, for different n.

If f is a radial function, we write $f = f(r)$, where $r = |x|$.

Corollary 4.4.14.

Let $P_k(x) \in \mathscr{H}_k(\mathbb{R}^n)$. Suppose that f is radial and $P_k(x)f(r) \in L^2(\mathbb{R}^n)$. Then the Fourier transform of $P_k(x)f(r)$ is also of the form $P_k(x)g(r)$, with g a radial function. Moreover, the induced transform $f \mapsto g$, $T_{n,k}f = g$, depends essentially only on $n+2k$. More precisely, we have Bochner's relation

$$T_{n,k} = (-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}. \tag{4.4.14}$$

Proof. Consider the Hilbert space of radial functions

$$\mathscr{R} = \left\{ f(r) : \|f\|^2 = \int_0^\infty |f(r)|^2 r^{2k+n-1}dr < \infty \right\},$$

with the indicated norm. Fix now $P_k(x)$, and assume that P_k is normalized, i.e.,

$$\int_{S^{n-1}} |P_k(x)|^2 d\sigma(x) = 1.$$

Our goal is to show that

$$(T_{n,k}f)(r) = (-i\,\mathrm{sgn}\,(\omega))^k (T_{n+2k,0}f)(r), \tag{4.4.15}$$

for each $f \in \mathscr{R}$.

We consider $e^{-\frac{|\omega|}{2}\varepsilon r^2}$ for a fixed $\varepsilon > 0$. By the homogeneity of P_k and the interplay of dilations with the Fourier transform (cf. Proposition 2.1.2), i.e., $\mathscr{F}\delta^\varepsilon = \varepsilon^{-n}\delta^{\varepsilon^{-1}}\mathscr{F}$, and Hecke's identity, we get

$$\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}) = \varepsilon^{-k/2}\mathscr{F}(P_k(\varepsilon^{1/2}x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2})$$
$$=\varepsilon^{-k/2-n/2}\delta^{\varepsilon^{-1/2}}\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}|x|^2})$$
$$=\varepsilon^{-k/2-n/2}(-i\,\mathrm{sgn}\,(\omega))^k\delta^{\varepsilon^{-1/2}}(P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2})$$
$$=(-i\,\mathrm{sgn}\,(\omega))^k\varepsilon^{-k/2-n/2}P_k(\varepsilon^{-1/2}\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}$$
$$=(-i\,\mathrm{sgn}\,(\omega))^k\varepsilon^{-k-n/2}P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}.$$

This shows that $T_{n,k}e^{-\frac{|\omega|}{2}\varepsilon r^2} = (-i\,\mathrm{sgn}\,(\omega))^k\varepsilon^{-k-n/2}e^{-\frac{|\omega|}{2}r^2/\varepsilon}$, and so

$$T_{n+2k,0}e^{-\frac{|\omega|}{2}\varepsilon r^2} =(-i\,\mathrm{sgn}\,(\omega))^0\varepsilon^{-0-(n+2k)/2}e^{-\frac{|\omega|}{2}r^2/\varepsilon} \\ =\varepsilon^{-k-n/2}e^{-\frac{|\omega|}{2}r^2/\varepsilon}.$$

Thus, $T_{n,k}e^{-\frac{|\omega|}{2}\varepsilon r^2} = (-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}e^{-\frac{|\omega|}{2}\varepsilon r^2}$ for $\varepsilon > 0$.

To finish the proof, it suffices to see that the linear combination of $\{e^{-\frac{|\omega|}{2}\varepsilon r^2}\}_{0<\varepsilon<\infty}$ is dense in $\mathscr{R}$. Suppose the contrary, then there exists a (almost everywhere) non-zero $g \in \mathscr{R}$, such that g is orthogonal to every $e^{-\frac{|\omega|}{2}\varepsilon r^2}$ in the sense of $\mathscr{R}$, i.e.,

$$\int_0^\infty e^{-\frac{|\omega|}{2}\varepsilon r^2} g(r) r^{2k+n-1} dr = 0, \tag{4.4.16}$$

for all $\varepsilon > 0$. Let $\psi(s) = \int_0^s e^{-r^2} g(r) r^{n+2k-1} dr$ for $s \geqslant 0$. Then, putting $\varepsilon = 2(m+1)/|\omega|$, where m is a positive integer, and by integration by parts, we have

$$0 = \int_0^\infty e^{-mr^2}\psi'(r)dr = 2m\int_0^\infty e^{-mr^2}\psi(r) r dr,$$

since $\psi(0) = 0$ and $0 \leqslant e^{-mr^2}\psi(r) \leqslant Ce^{-mr^2}r^{k+(n-1)/2} \to 0$ as $r \to \infty$ by the Hölder inequality. By the change of variable $z = e^{-r^2}$, this equality is equivalent to

$$0 = \int_0^1 z^{m-1}\psi(\sqrt{\ln 1/z})dz, \quad m = 1, 2, \ldots.$$

Since the polynomials are uniformly dense in the space of continuous functions on the closed interval $[0, 1]$, this can only be the case when $\psi(\sqrt{\ln 1/z}) = 0$ for all z in $[0, 1]$. Thus, $\psi'(r) = e^{-r^2} g(r) r^{n+2k-1} = 0$ for almost every $r \in (0, \infty)$, contradicting the hypothesis that $g(r)$ is not equal to 0 almost everywhere.

Since the operators $T_{n,k}$ and $(-i\,\mathrm{sgn}\,(\omega))^k T_{n+2k,0}$ are bounded and agree on the dense subspace, they must be equal. Thus, we have shown the desired result. ■

We come now to what has been our main goal in our discussion of spherical harmonics.

Theorem 4.4.15.

Let $P_k(x) \in \mathscr{H}_k$, $k \geqslant 1$. Then the multiplier corresponding to the transform (4.4.10) with the kernel $\frac{P_k(x)}{|x|^{k+n}}$ is

$$\gamma_k \frac{P_k(\xi)}{|\xi|^k}, \quad \text{with } \gamma_k = \left(\frac{|\omega|}{2}\right)^{n/2} (-i\,\mathrm{sgn}\,(\omega))^k \frac{\Gamma(k/2)}{\Gamma(k/2 + n/2)}.$$

Remark 4.4.16. 1) If $k \geqslant 1$, then $P_k(x)$ is orthogonal to the constants on the sphere, and so its mean value over any sphere centered at the origin is zero.

2) The statement of the theorem can be interpreted as

$$\mathscr{F}\left(\frac{P_k(x)}{|x|^{k+n}}\right) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \gamma_k \frac{P_k(\xi)}{|\xi|^k}. \tag{4.4.17}$$

3) As such it will be derived from the following closely related fact,

$$\mathscr{F}\left(\frac{P_k(x)}{|x|^{k+n-\alpha}}\right) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \gamma_{k,\alpha} \frac{P_k(\xi)}{|\xi|^{k+\alpha}}, \tag{4.4.18}$$

where $\gamma_{k,\alpha} = \left(\frac{|\omega|}{2}\right)^{\frac{n}{2}-\alpha} (-i\,\mathrm{sgn}\,(\omega))^k \frac{\Gamma(k/2+\alpha/2)}{\Gamma(k/2+n/2-\alpha/2)}$.

Lemma 4.4.17.

The identity (4.4.18) holds in the sense that
$$\int_{\mathbb{R}^n} \frac{P_k(x)}{|x|^{k+n-\alpha}} \widehat{\varphi}(x)dx = \gamma_{k,\alpha} \int_{\mathbb{R}^n} \frac{P_k(\xi)}{|\xi|^{k+\alpha}} \varphi(\xi)d\xi, \quad \forall \varphi \in \mathscr{S}. \tag{4.4.19}$$
It is valid for all non-negative integer k and for $0 < \alpha < n$.

Proof. From the proof of Corollary 4.4.14, we have already known that
$$\mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}) = (-i\,\text{sgn}\,(\omega))^k \varepsilon^{-k-n/2} P_k(\xi) e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon},$$
so we have by the multiplication formula,
$$\int_{\mathbb{R}^n} P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}\widehat{\varphi}(x)dx = \int_{\mathbb{R}^n} \mathscr{F}(P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2})(\xi)\varphi(\xi)d\xi$$
$$=(-i\,\text{sgn}\,(\omega))^k \varepsilon^{-k-n/2} \int_{\mathbb{R}^n} P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}\varphi(\xi)d\xi,$$
for $\varepsilon > 0$.

We now integrate both sides of the above w.r.t. ε, after having multiplied the equation by $\varepsilon^{\beta-1}$ (to be determined). That is
$$\begin{aligned}&\int_0^\infty \varepsilon^{\beta-1} \int_{\mathbb{R}^n} P_k(x)e^{-\frac{|\omega|}{2}\varepsilon|x|^2}\widehat{\varphi}(x)dxd\varepsilon \\ =&(-i\,\text{sgn}\,(\omega))^k \int_0^\infty \varepsilon^{\beta-1}\varepsilon^{-k-n/2} \int_{\mathbb{R}^n} P_k(\xi)e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}\varphi(\xi)d\xi d\varepsilon.\end{aligned} \tag{4.4.20}$$
By changing the order of the double integral and a change of variable, we get
$$\begin{aligned}\text{l.h.s. of (4.4.20)} =& \int_{\mathbb{R}^n} P_k(x)\widehat{\varphi}(x) \int_0^\infty \varepsilon^{\beta-1}e^{-\frac{|\omega|}{2}\varepsilon|x|^2}d\varepsilon dx \\ \overset{t=|\omega|\varepsilon|x|^2/2}{=\!=\!=\!=} & \int_{\mathbb{R}^n} P_k(x)\widehat{\varphi}(x)\left(\frac{|\omega|}{2}|x|^2\right)^{-\beta} \int_0^\infty t^{\beta-1}e^{-t}dtdx \\ =&\left(\frac{|\omega|}{2}\right)^{-\beta} \Gamma(\beta) \int_{\mathbb{R}^n} P_k(x)\widehat{\varphi}(x)|x|^{-2\beta}dx.\end{aligned}$$
Thus, we can take $\beta = (k+n-\alpha)/2$. Similarly,
$$\begin{aligned}\text{r.h.s. of (4.4.20)} =& (-i\,\text{sgn}\,(\omega))^k \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi) \\ & \int_0^\infty \varepsilon^{-(k/2+\alpha/2+1)}e^{-\frac{|\omega|}{2}|\xi|^2/\varepsilon}d\varepsilon d\xi \\ \overset{t=\frac{|\omega|}{2}|\xi|^2/\varepsilon}{=\!=\!=} & (-i\,\text{sgn}\,(\omega))^k \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi)\left(\frac{|\omega|}{2}|\xi|^2\right)^{-(k+\alpha)/2} \\ & \int_0^\infty t^{k/2+\alpha/2-1}e^{-t}dtd\xi \\ =& (-i\,\text{sgn}\,(\omega))^k \left(\frac{|\omega|}{2}\right)^{-(k+\alpha)/2} \Gamma(k/2+\alpha/2) \\ & \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi)|\xi|^{-(k+\alpha)}d\xi.\end{aligned}$$
Thus, we get
$$\begin{aligned}&\left(\frac{|\omega|}{2}\right)^{-(k+n-\alpha)/2} \Gamma((k+n-\alpha)/2) \int_{\mathbb{R}^n} P_k(x)\widehat{\varphi}(x)|x|^{-(k+n-\alpha)}dx \\ =&(-i\,\text{sgn}\,(\omega))^k \left(\frac{|\omega|}{2}\right)^{-(k+\alpha)/2} \Gamma(k/2+\alpha/2)\end{aligned}$$

$$\cdot \int_{\mathbb{R}^n} P_k(\xi)\varphi(\xi)|\xi|^{-(k+\alpha)}d\xi$$

which leads to (4.4.19).

Observe that when $0 < \alpha < n$ and $\varphi \in \mathscr{S}$, double integrals in the above converge absolutely. Thus, the formal argument just given establishes the lemma. ■

Remark 4.4.18. For the complex number α with $\Re\alpha \in (0, n)$, the lemma and (4.4.18) are also valid, see [SW71, Theorem 4.1, p.160-163].

Proof of Theorem 4.4.15. By the assumption that $k \geqslant 1$, we have that the integral of P_k over any sphere centered at the origin is zero. Thus for $\varphi \in \mathscr{S}$, we get

$$\int_{\mathbb{R}^n} \frac{P_k(x)}{|x|^{k+n-\alpha}}\widehat{\varphi}(x)dx = \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n-\alpha}}[\widehat{\varphi}(x) - \widehat{\varphi}(0)]dx + \int_{|x|>1} \frac{P_k(x)}{|x|^{k+n-\alpha}}\widehat{\varphi}(x)dx.$$

Obviously, the second term tends to $\int_{|x|>1} \frac{P_k(x)}{|x|^{k+n}}\widehat{\varphi}(x)dx$ as $\alpha \to 0$ by the dominated convergence theorem. It is clear that $\frac{P_k(x)}{|x|^{k+n}}[\widehat{\varphi}(x) - \widehat{\varphi}(0)]$ is locally integrable, thus we have, by the dominated convergence theorem, the limit of the first term in the r.h.s. of the above

$$\lim_{\alpha\to 0+} \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n-\alpha}}[\widehat{\varphi}(x) - \widehat{\varphi}(0)]dx = \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n}}[\widehat{\varphi}(x) - \widehat{\varphi}(0)]dx$$
$$= \int_{|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n}}\widehat{\varphi}(x)dx = \lim_{\varepsilon\to 0} \int_{\varepsilon\leqslant|x|\leqslant 1} \frac{P_k(x)}{|x|^{k+n}}\widehat{\varphi}(x)dx.$$

Thus, we obtain

$$\lim_{\alpha\to 0+} \int_{\mathbb{R}^n} \frac{P_k(x)}{|x|^{k+n-\alpha}}\widehat{\varphi}(x)dx = \lim_{\varepsilon\to 0} \int_{|x|\geqslant\varepsilon} \frac{P_k(x)}{|x|^{k+n}}\widehat{\varphi}(x)dx. \tag{4.4.21}$$

Similarly,

$$\lim_{\alpha\to 0+} \int_{\mathbb{R}^n} \frac{P_k(\xi)}{|\xi|^{k+\alpha}}\varphi(\xi)d\xi = \lim_{\varepsilon\to 0} \int_{|\xi|\geqslant\varepsilon} \frac{P_k(\xi)}{|\xi|^{k}}\varphi(\xi)d\xi.$$

Thus, by Lemma 4.4.14, we complete the proof with $\gamma_k = \lim_{\alpha\to 0}\gamma_{k,\alpha}$. ■

For fixed $k \geqslant 1$, the linear space of operators in (4.4.10), where $\Omega(y) = \frac{P_k(y)}{|y|^k}$ and $P_k \in \mathscr{H}_k$, form a natural generalization of the Riesz transforms; the latter arise in the special case $k = 1$. Those for $k > 1$, we call the ***higher Riesz transforms***, with k as the degree of the higher Riesz transforms, they can also be characterized by their invariant properties (see [Ste70, §4.8, p.79]).

Theorem 4.4.19.

The higher Riesz transforms are of weak-type $(1, 1)$ and of strong-type (p, p) for $1 < p < \infty$.

Proof. It suffices to show that $K(x) = \text{p.v.} \frac{P_k(x)}{|x|^{n+k}}$ with $P_k \in \mathscr{H}_k$ satisfy the hypotheses of Theorem 4.2.1. Clearly K coincides with a locally integrable function on $\mathbb{R}^n \setminus \{0\}$. Moreover, by Theorem 4.4.15 we get

$$\widehat{K}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \gamma_k \frac{P_k(\xi)}{|\xi|^k},$$

which is clearly bounded.

Finally, we have on $\mathbb{R}^n \setminus \{0\}$

$$|\nabla K(x)| \leqslant \frac{1}{|x|^{n+k}} \sum_{|\alpha|=k} \sum_{\substack{1\leqslant j\leqslant n\\ \alpha_j>1}} |a_\alpha \alpha_j x_1^{\alpha_1} \cdots x_j^{\alpha_j-1} \cdots x_n^{\alpha_n}| + (n+k)\frac{P_k(x)|x_j|}{|x|^{n+k+2}}$$
$$\leqslant \frac{C}{|x|^{n+1}},$$

that is, (4.2.5) is satisfied. Thus, by Theorem 4.2.1 we obtain the desired results. ■

§ 4.4.3 Equivalence between two classes of transforms

We now consider two classes of transforms, defined on $L^2(\mathbb{R}^n)$. The first class consists of all transforms of the form

$$Tf = c\cdot f + \lim_{\varepsilon\to 0}\int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy, \tag{4.4.22}$$

where c is a constant, $\Omega \in \mathcal{C}^\infty(S^{n-1})$ is a homogeneous function of degree 0, and the integral $\int_{S^{n-1}} \Omega(x)d\sigma(x) = 0$. The second class is given by those transform T for which

$$\widetilde{Tf}(\xi) = m(\xi)\widehat{f}(\xi) \tag{4.4.23}$$

where the multiplier $m \in \mathcal{C}^\infty(S^{n-1})$ is homogeneous of degree 0.

Theorem 4.4.20.

The two classes of transforms, defined by (4.4.22) and (4.4.23) respectively, are identical.

Proof. First, support that T is of the form (4.4.22). Then by Theorem 4.3.3, T is of the form (4.4.23) with m homogeneous of degree 0 and

$$m(\xi) = c + \int_{S^{n-1}} \left[\ln(1/|\xi\cdot x|) - \frac{\pi i}{2}\,\mathrm{sgn}\,(\omega)\,\mathrm{sgn}\,(\xi\cdot x)\right]\Omega(x)d\sigma(x),\ |\xi| = 1. \tag{4.4.24}$$

Now, we need to show $m \in \mathcal{C}^\infty(S^{n-1})$. Write the spherical harmonic developments

$$\Omega(x) = \sum_{k=1}^{\infty} Y_k(x),\ m(x) = \sum_{k=0}^{\infty} \widetilde{Y}_k(x),\ \Omega_N(x) = \sum_{k=1}^{N} Y_k(x),\ m_N(x) = \sum_{k=0}^{N} \widetilde{Y}_k(x), \tag{4.4.25}$$

where $Y_k, \widetilde{Y}_k \in H_k$ in view of part (3) in Proposition 4.4.12. k starts from 1 in the development of Ω, since $\int_{S^{n-1}} \Omega(x)dx = 0$ implies that $\Omega(x)$ is orthogonal to constants, and H_0 contains only constants.

Then, by Theorem 4.4.15 and the Plancherel theorem, we get that if $\Omega = \Omega_N$, then $m(x) = m_N(x)$, with

$$\widetilde{Y}_k(x) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \gamma_k Y_k(x), \quad k \geqslant 1.$$

But $m_M(x) - m_N(x) = \int_{S^{n-1}} \left[-\frac{\pi i}{2}\,\mathrm{sgn}\,(\omega)\,\mathrm{sgn}\,(y\cdot x) + \ln\frac{1}{|y\cdot x|}\right][\Omega_M(y) - \Omega_N(y)]d\sigma(y)$. Moreover, by Hölder's inequality, we have

$$\sup_{x\in S^{n-1}} |m_M(x) - m_N(x)|$$

$$\leqslant \left(\sup_x \int_{S^{n-1}} \left|-\frac{\pi i}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(y\cdot x)+\ln(1/|y\cdot x|)\right|^2 d\sigma(y)\right)^{1/2}$$

$$\times \left(\int_{S^{n-1}} |\Omega_M(y)-\Omega_N(y)|^2 d\sigma(y)\right)^{1/2} \to 0, \tag{4.4.26}$$

as $M, N\to\infty$, since for $n=1$, $S^0=\{-1,1\}$,

$$\int_{S^0}\left|-\frac{\pi i}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(y\cdot x)+\ln(1/|y\cdot x|)\right|^2 d\sigma(y)=\frac{\pi^2}{2},$$

and for $n\geqslant 2$, we can pick a orthogonal matrix A satisfying $Ae_1=x$ and $\det A=1$ for $|x|=1$, and then by a change of variable,

$$\begin{aligned}
&\sup_x \int_{S^{n-1}} \left|-\frac{\pi i}{2}\operatorname{sgn}(\omega)\operatorname{sgn}(y\cdot x)+\ln(1/|y\cdot x|)\right|^2 d\sigma(y)\\
=&\sup_x \int_{S^{n-1}} \left[\frac{\pi^2}{4}+(\ln(1/|y\cdot x|))^2\right] d\sigma(y)\\
=&\frac{\pi^2}{4}\omega_{n-1}+\sup_x\int_{S^{n-1}}(\ln|y\cdot Ae_1|)^2 d\sigma(y)\\
=&\frac{\pi^2}{4}\omega_{n-1}+\sup_x\int_{S^{n-1}}(\ln|A^{-1}y\cdot e_1|)^2 d\sigma(y)\\
\overset{z=A^{-1}y}{=\!=\!=\!=}&\frac{\pi^2}{4}\omega_{n-1}+\int_{S^{n-1}}(\ln|z_1|)^2 d\sigma(z)<\infty.
\end{aligned}$$

Here, we have used the boundedness of the integral in the r.h.s., i.e., (with the notation $\bar z=(z_2,\dots,z_n)$, as in the proof of Theorem 4.3.3,

$$\begin{aligned}
\int_{S^{n-1}}(\ln|z_1|)^2 d\sigma(z)=&\int_{-1}^1(\ln|z_1|)^2\int_{S^{n-2}}(1-z_1^2)^{(n-3)/2}d\sigma(\bar z)dz_1\\
=&\omega_{n-2}\int_{-1}^1(\ln|z_1|)^2(1-z_1^2)^{(n-3)/2}dz_1\\
\overset{z_1=\cos\theta}{=\!=\!=\!=}&\omega_{n-2}\int_0^\pi(\ln|\cos\theta|)^2(\sin\theta)^{n-2}d\theta=\omega_{n-2}I_1.
\end{aligned}$$

If $n\geqslant 3$, then, by integration by parts,

$$I_1\leqslant\int_0^\pi(\ln|\cos\theta|)^2\sin\theta d\theta=-2\int_0^\pi\ln|\cos\theta|\sin\theta d\theta=2\int_0^\pi\sin\theta d\theta=4.$$

If $n=2$, then, by the formula[1] $\int_0^{\pi/2}(\ln(\cos\theta))^2d\theta=\frac{\pi}{2}[(\ln 2)^2+\pi^2/12]$, cf. [GR07, 4.225.8, p.531], we get

$$I_1=\int_0^\pi(\ln|\cos\theta|)^2d\theta=2\int_0^{\pi/2}(\ln(\cos\theta))^2d\theta=\pi[(\ln 2)^2+\pi^2/12].$$

Thus, (4.4.26) shows that

$$m(x)=c+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\sum_{k=1}^\infty\gamma_kY_k(x).$$

Since $\Omega\in\mathcal{C}^\infty$, we have, in view of part (5) of Proposition 4.4.12, that

$$\int_{S^{n-1}}|Y_k(x)|^2d\sigma(x)=O(k^{-N})$$

[1] One can see https://math.stackexchange.com/questions/58654 or http://www.doc88.com/p-9798925245778.html for some detailed solutions.

as $k \to \infty$ for every fixed N. However, by the explicit form of γ_k, we see that $\gamma_k \sim k^{-n/2}$, so $m(x)$ is also indefinitely differentiable on the unit sphere, i.e., $m \in \mathscr{C}^\infty(S^{n-1})$.

Conversely, suppose $m(x) \in \mathscr{C}^\infty(S^{n-1})$ and let its spherical harmonic development be as in (4.4.25). Set $c = \widetilde{Y}_0$, and $Y_k(x) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \frac{1}{\gamma_k} \widetilde{Y}_k(x)$. Then $\Omega(x)$, given by (4.4.25), has mean value zero in the sphere, and is again indefinitely differentiable there. But as we have just seen the multiplier corresponding to this transform is m; so the theorem is proved. ■

As an application of this theorem, we shall give the generalization of the estimates for partial derivatives given in 4.4.1.

Let $P(x) \in \mathscr{P}_k(\mathbb{R}^n)$. We shall say that P is ***elliptic*** if $P(x)$ vanishes only at the origin. For any polynomial P, we consider also its corresponding differential polynomial. Thus, if $P(x) = \sum a_\alpha x^\alpha$, we write $P(\partial_x) = \sum a_\alpha \partial_x^\alpha$ as in the previous definition.

Corollary 4.4.21.

Suppose P is a homogeneous elliptic polynomial of degree k. Let ∂_x^α be any differential monomial of degree k. Assume $f \in \mathscr{C}_c^k$, then we have the a priori estimate

$$\|\partial_x^\alpha f\|_p \leqslant A_p \|P(\partial_x) f\|_p, \quad 1 < p < \infty. \tag{4.4.27}$$

Proof. From the Fourier transform of $\partial_x^\alpha f$ and $P(\partial_x) f$,

$$\widehat{P(\partial_x) f}(\xi) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\omega i x\cdot\xi} P(\partial_x) f(x) dx = (\omega i)^k P(\xi) \widehat{f}(\xi),$$

and

$$\widehat{\partial_x^\alpha f}(\xi) = (\omega i)^k \xi^\alpha \widehat{f}(\xi),$$

we have the following relation

$$P(\xi) \widehat{\partial_x^\alpha f}(\xi) = \xi^\alpha \widehat{P(\partial_x) f}(\xi).$$

Since $P(\xi)$ is non-vanishing except at the origin, $\frac{\xi^\alpha}{P(\xi)}$ is homogeneous of degree 0 and is indefinitely differentiable on the unit sphere. Thus

$$\partial_x^\alpha f = T(P(\partial_x) f),$$

where T is one of the transforms of the type given by (4.4.23). By Theorem 4.4.20, T is also given by (4.4.22) and hence by the result of Theorem 4.2.1 and Proposition 4.2.2, we get the estimate (4.4.27). ■

§ 4.5 The method of rotations and singualr integral with odd kernels

A simple procedure called the method of rotations plays a crucial role in the study of operators T_Ω when Ω is an odd function. This method is based on the use of the directional Hilbert transforms.

Fix a unit vector $\theta \in \mathbb{R}^n$. For $f \in \mathscr{S}(\mathbb{R}^n)$, let

$$H_\theta f(x) = \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} f(x - t\theta) \frac{dt}{t}. \tag{4.5.1}$$

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 (4.5.1) is well-defined, since it converges rapidly at infinity and by subtracting the constant $f(x)$, it also converges near zero.

Now, we define the ***directional maximal Hilbert transforms***. For a function $f \in \cup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$ 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 almost everywhere. 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 for almost all $x\in\mathbb{R}^n$. Thus, $H_\theta^{(**)} f$ is well-defined for $f \in \cup_{1\leqslant p<\infty} L^p(\mathbb{R}^n)$.

Theorem 4.5.1.

If Ω is odd and integrable over 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 S^{n-1}. The operator H_{e_1} is the directional Hilbert transform in the 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 4.1.9, we have

$$\begin{aligned} \|H_{e_1} f\|_{L^p(\mathbb{R}^n)}^p =& \left\| \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|t|\geqslant\varepsilon} f(x-te_1)\frac{dt}{t}\right\|_{L^p(\mathbb{R}^n)}^p \\ =& \left\| \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|t|\geqslant\varepsilon} f(x_1-t,x_2,\cdots,x_n)\frac{dt}{t}\right\|_{L^p(\mathbb{R}^n)}^p \\ \leqslant& \left\| \|H\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})}^p \|f(x_1,x')\|_{L^p_{x_1}(\mathbb{R})}^p \right\|_{L^p_{x'}(\mathbb{R}^{n-1})}^p \\ =& \|H\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})}^p \|f\|_{L^p(\mathbb{R}^n)}^p. \end{aligned}$$

Next, observe that the following identity is valid for all matrices $A\in O(n)$ (the set of all $n\times n$ orthogonal matrices):

$$\begin{aligned} H_{Ae_1} f(x) =& \frac{1}{\pi}\,\text{p.v.}\int_{-\infty}^{\infty} f(x-tAe_1)\frac{dt}{t} \\ =& \frac{1}{\pi}\,\text{p.v.}\int_{-\infty}^{\infty} f(A(A^{-1}x-te_1))\frac{dt}{t} \\ =& H_{e_1}(f\circ A)(A^{-1}x). \end{aligned} \tag{4.5.2}$$

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 $\theta\in S^{n-1}$.

Identity (4.5.2) 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 of 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:

$$\begin{aligned}\int_{\varepsilon\leqslant|y|\leqslant N} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy =& \int_{S^{n-1}} \Omega(\theta) \int_{\varepsilon}^{N} f(x-r\theta)\frac{dr}{r} d\sigma(\theta)\\ =& -\int_{S^{n-1}} \Omega(\theta) \int_{\varepsilon}^{N} f(x+r\theta)\frac{dr}{r} d\sigma(\theta)\\ =& \int_{S^{n-1}} \Omega(\theta) \int_{-N}^{-\varepsilon} f(x-r\theta)\frac{dr}{r} d\sigma(\theta),\end{aligned}$$

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

$$\begin{aligned}&\int_{\varepsilon\leqslant|y|\leqslant N} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)dy\\ =&\frac{1}{2}\int_{S^{n-1}} \Omega(\theta) \int_{\varepsilon\leqslant|r|\leqslant N} \frac{f(x-r\theta)}{r} dr d\sigma(\theta) \qquad (4.5.3)\\ =&\frac{\pi}{2}\int_{S^{n-1}} \Omega(\theta) H_\theta^{(\varepsilon,N)} f(x) d\sigma(\theta). \qquad (4.5.4)\end{aligned}$$

Since Ω is odd and so it has mean value zero, we can get

$$\begin{aligned}(4.5.3) =&\frac{1}{2}\int_{S^{n-1}} \Omega(\theta) \int_{\varepsilon\leqslant|r|\leqslant 1} \frac{f(x-r\theta)-f(x)}{r} dr d\sigma(\theta)\\ &+\frac{1}{2}\int_{S^{n-1}} \Omega(\theta) \int_{1<|r|\leqslant N} \frac{f(x-r\theta)}{r} dr d\sigma(\theta).\end{aligned}$$

Because $f \in \mathscr{S}$, the inner integrals is uniformly bounded, so we can apply the Lebesgue dominated convergence theorem to get

$$T_\Omega f(x) = \frac{\pi}{2}\int_{S^{n-1}} \Omega(\theta) H_\theta f(x) d\sigma(\theta). \qquad (4.5.5)$$

From (4.5.4), we conclude that

$$T_\Omega^{(**)} f(x) \leqslant \frac{\pi}{2}\int_{S^{n-1}} |\Omega(\theta)| H_\theta^{(**)} f(x) d\sigma(\theta). \qquad (4.5.6)$$

The L^p boundedness of T_Ω and $T_\Omega^{(**)}$ for Ω odd are then trivial consequences of (4.5.6) and (4.5.5) via Minkowski's integral inequality. ■

Remark 4.5.2. It follows from the proof of Theorem 4.5.1 and from Theorems 4.1.9 and 4.1.13 that whenever Ω is an odd function on S^{n-1}, we have

$$\begin{aligned}\|T_\Omega\|_{L^p\to L^p} \leqslant& C\|\Omega\|_1 \begin{cases} p, & \text{if } p\geqslant 2,\\ (p-1)^{-1}, & \text{if } 1<p<2,\end{cases}\\ \|T_\Omega^{(**)}\|_{L^p\to L^p} \leqslant& C\|\Omega\|_1 \begin{cases} p, & \text{if } p\geqslant 2,\\ (p-1)^{-1}, & \text{if } 1<p<2,\end{cases}\end{aligned}$$

for some $C>0$ independent of p and the dimension.

Corollary 4.5.3.

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. By the way, the boundedness for R_j, we have obtained in Theorem 4.4.3. The assertion follows from the fact that the Riesz transforms have odd kernels. Since the kernel of R_j decays like $|x|^{-n}$ near infinity, it follows that $R_j^{(*)} f$ is well-defined for $f \in L^p(\mathbb{R}^n)$. Since $R_j^{(*)}$ is point-wise bounded by $2R_j^{(**)}$, the conclusion follows from Theorem 4.5.1. ∎

§4.6 Singular integral operators with Dini-type condition

§4.6.1 L^p boundedness of homogeneous singular integrals

In this section, we shall consider those operators which not only commute with translations but also with dilations. Among these we shall study the class of singular integral operators, falling under the scope of Theorem 4.2.4.

If T corresponds to the kernel $K(x)$, then as we have already pointed out, $\delta^{\varepsilon^{-1}} T \delta^{\varepsilon}$ corresponds to the kernel $\varepsilon^{-n} K(\varepsilon^{-1} x)$. So if $\delta^{\varepsilon^{-1}} T \delta^{\varepsilon} = T$ we are back to the requirement $K(x) = \varepsilon^{-n} K(\varepsilon^{-1} x)$, i.e., $K(\varepsilon x) = \varepsilon^{-n} K(x)$, $\varepsilon > 0$; that is K is homogeneous of degree $-n$. Put another way

$$K(x) = \frac{\Omega(x)}{|x|^n}, \tag{4.6.1}$$

with Ω homogeneous of degree 0, i.e., $\Omega(\varepsilon x) = \Omega(x)$, $\varepsilon > 0$. This condition on Ω is equivalent with the fact that it is constant on rays emanating from the origin; in particular, Ω is completely determined by its restriction to the unit sphere S^{n-1}.

Let us try to reinterpret the conditions of Theorem 4.2.4 in terms of Ω.

1) By (4.2.8), $\Omega(x)$ must be bounded and consequently integrable on S^{n-1}; and another condition $\int_{|x|\geqslant 2|y|} \left| \frac{\Omega(x-y)}{|x-y|^n} - \frac{\Omega(x)}{|x|^n} \right| dx \leqslant C$ which is not easily restated precisely in terms of Ω. However, what is evident is that it requires a certain continuity of Ω. Here we shall content ourselves in treating the case where Ω satisfies the following "***Dini-type***" condition suggested by (4.2.8):

$$\text{if } w(\eta) := \sup_{\substack{|x-x'|\leqslant \eta \\ |x|=|x'|=1}} |\Omega(x) - \Omega(x')|, \quad \text{then } \int_0^1 \frac{w(\eta)}{\eta} d\eta < \infty. \tag{4.6.2}$$

Of course, any Ω which is of class $\mathscr{C}^1$, or even merely Lipschitz continuous, satisfies the condition (4.6.2).

2) The cancelation condition (4.2.9) is then the same as the mean value zero of Ω on S^{n-1}.

Theorem 4.6.1.

Let $\Omega \in L^\infty(S^{n-1})$ be homogeneous of degree 0 with mean value zero on S^{n-1}, and suppose that Ω satisfies the smoothness property (4.6.2). For $1 <$

$p < \infty$, and $f \in L^p(\mathbb{R}^n)$, let

$$T_\varepsilon f(x) = \int_{|y|\geqslant\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy. \tag{4.6.3}$$

(a) Then there exists a bound A_p (independent of f and ε) such that

$$\|T_\varepsilon f\|_p \leqslant A_p \|f\|_p.$$

(b) $\lim_{\varepsilon\to 0} T_\varepsilon f = Tf$ exists in L^p norm, and

$$\|Tf\|_p \leqslant A_p \|f\|_p.$$

(c) If $f \in L^2(\mathbb{R}^n)$, then the Fourier multiplier m corresponding to T is a homogeneous function of degree 0 expressed in (4.3.5).

Proof. The conclusions (a) and (b) are immediately consequences of Theorem 4.2.4, once we have shown that any $K(x)$ of the form $\frac{\Omega(x)}{|x|^n}$ satisfies

$$\int_{|x|\geqslant 2|y|} |K(x-y) - K(x)|dx \leqslant B, \tag{4.6.4}$$

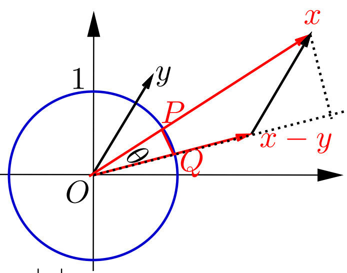
if Ω is as in condition (4.6.2). Indeed,

$$K(x-y) - K(x) = \frac{\Omega(x-y) - \Omega(x)}{|x-y|^n} + \Omega(x)\left[\frac{1}{|x-y|^n} - \frac{1}{|x|^n}\right].$$

The second group of terms is bounded since Ω is bounded and

$$\begin{aligned}
&\int_{|x|\geqslant 2|y|} \left|\frac{1}{|x-y|^n} - \frac{1}{|x|^n}\right| dx = \int_{|x|\geqslant 2|y|} \left|\frac{|x|^n - |x-y|^n}{|x-y|^n|x|^n}\right| dx \\
=&\int_{|x|\geqslant 2|y|} \frac{||x| - |x-y|| \sum_{j=0}^{n-1} |x|^{n-1-j}|x-y|^j}{|x-y|^n|x|^n} dx \\
\leqslant&\int_{|x|\geqslant 2|y|} |y| \sum_{j=0}^{n-1} |x|^{-j-1}|x-y|^{j-n} dx \\
\leqslant&\int_{|x|\geqslant 2|y|} |y| \sum_{j=0}^{n-1} |x|^{-j-1}(|x|/2)^{j-n} dx \quad (\because |x-y| \geqslant |x| - |y| \geqslant |x|/2) \\
=&\int_{|x|\geqslant 2|y|} |y| \sum_{j=0}^{n-1} 2^{n-j}|x|^{-n-1} dx = 2(2^n-1)|y| \int_{|x|\geqslant 2|y|} |x|^{-n-1}dx \\
=&2(2^n-1)|y|\omega_{n-1}\frac{1}{2|y|} = (2^n-1)\omega_{n-1}.
\end{aligned}$$

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

$$\begin{aligned}2^n\int_{|x|\geqslant 2|y|} w\left(2\frac{|y|}{|x|}\right)\frac{dx}{|x|^n}=&2^n\int_{|z|\geqslant 2} w(2/|z|)\frac{dz}{|z|^n}\\ =2^n\omega_{n-1}\int_2^\infty w(2/r)\frac{dr}{r}=&2^n\omega_{n-1}\int_0^1\frac{w(\eta)d\eta}{\eta}<\infty\end{aligned}$$

in view of changes of variables $x=|y|z$ and the Dini-type condition (4.6.2).

For part (c), it is the same as the proof of Theorem 4.3.3 with minor modification. Indeed, we only need to simplify the proof of (4.3.6) due to $\Omega\in L^\infty(S^{n-1})$ here. We can control (4.3.6) by

$$\omega_{n-1}(4+\frac{|\omega|^2}{4}R^2)\|\Omega\|_{L^\infty(S^{n-1})}+2\|\Omega\|_{L^\infty(S^{n-1})}\int_{S^{n-1}}\ln(1/|\xi'\cdot x'|)d\sigma(x'),$$

where the integral in the last term is equal to

$$\int_{S^{n-1}}\ln(1/|y_1|)d\sigma(y)$$

which have been estimated in Theorem 4.3.3. Thus, we have completed the proof. ■

§ 4.6.2 The maximal singular integral operator

Theorem 4.6.1 guaranteed the existence of the singular integral

$$\lim_{\varepsilon\to 0}\int_{|y|\geqslant\varepsilon}\frac{\Omega(y)}{|y|^n}f(x-y)dy \tag{4.6.5}$$

in the sense of convergence in the L^p norm. The natural counterpart of this result is that of convergence almost everywhere. For the questions involving almost everywhere convergence, it is best to consider also the corresponding maximal function.

Theorem 4.6.2.

Suppose that Ω satisfies the conditions of Theorem 4.6.1. For $f\in L^p(\mathbb{R}^n)$, $1\leqslant p<\infty$, consider

$$T_\varepsilon f(x)=\int_{|y|\geqslant\varepsilon}\frac{\Omega(y)}{|y|^n}f(x-y)dy,\quad \varepsilon>0.$$

(The integral converges absolutely for every x.)

(a) $\lim_{\varepsilon\to 0} T_\varepsilon f(x)$ exists for almost every x.

(b) Let $T^*f(x)=\sup_{\varepsilon>0}|T^{(\varepsilon)}f(x)|$. If $f\in L^1(\mathbb{R}^n)$, then the mapping $f\to T^*f$ is of weak type $(1,1)$.

(c) If $1<p<\infty$, then $\|T^*f\|_p\leqslant A_p\|f\|_p$.

Proof. The argument for the theorem presents itself in three stages.

The first one is the proof of inequality (c) which can be obtained as a relatively easy consequence of the L^p norm existence of $\lim_{\varepsilon\to 0}T^{(\varepsilon)}$, already proved, and certain general properties of "approximations to the identity".

Let $Tf(x)=\lim_{\varepsilon\to 0}T_\varepsilon f(x)$, where the limit is taken in the L^p norm. Its existence is guaranteed by Theorem 4.6.1. We shall prove this part by showing the following ***Cotlar inequality***

$$T^*f(x)\leqslant M(Tf)(x)+CMf(x).$$

Let φ be a smooth non-negative function on $\mathbb{R}^n$, which is supported in the unit ball, has integral equal to one, and which is also radial and decreasing in $|x|$. Consider

$$K_\varepsilon(x) = \begin{cases} \frac{\Omega(x)}{|x|^n}, & |x| \geqslant \varepsilon, \\ 0, & |x| < \varepsilon. \end{cases}$$

This leads us to another function Φ defined by

$$\Phi = \varphi * K - K_1, \tag{4.6.6}$$

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 3.2.12).

In fact, if $|x| < 1$, then

$$\begin{aligned} |\Phi| =&|\varphi * K| = \left|\int_{\mathbb{R}^n} K(y)\varphi(x-y)dy\right| = \left|\int_{\mathbb{R}^n} K(y)(\varphi(x-y)-\varphi(x))dy\right| \\ \leqslant& \int_{\mathbb{R}^n} |K(y)||\varphi(x-y)-\varphi(x)|dy \leqslant C\int_{\mathbb{R}^n} \frac{|\varphi(x-y)-\varphi(x)|}{|y|^n}dy \leqslant C, \end{aligned}$$

since the mean value zero of Ω on 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 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 (4.6.4), we can get the bound for $|y| \leqslant 1$ and so $|x| \geqslant 2|y|$,

$$\begin{aligned} |K(x-y)-K(x)| \leqslant& 2^n w\left(\frac{2|y|}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_\infty |y||x|^{-(n+1)} \\ \leqslant& 2^n w\left(\frac{2}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_\infty |x|^{-(n+1)}, \end{aligned}$$

as in the proof of Theorem 4.6.1, since w is increasing. Thus, due to $\|\varphi\|_1 = 1$, we obtain for $|x| \geqslant 2$

$$|\Phi(x)| \leqslant 2^n w\left(\frac{2}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_\infty |x|^{-(n+1)}.$$

Therefore, we get $|\Psi| \leqslant C$ for $|x| < 2$, and

$$|\Psi(x)| \leqslant 2^n w\left(\frac{2}{|x|}\right)|x|^{-n} + 2(2^n-1)\|\Omega\|_\infty |x|^{-(n+1)},$$

for $|x| \geqslant 2$, and then we can proved that $\Psi \in L^1(\mathbb{R}^n)$ with the help of the Dini-type condition.

From (4.6.6), it follows, because the singular integral operator $\varphi \to \varphi * K$ commutes with dilations, that

$$\varphi_\varepsilon * K - K_\varepsilon = \Phi_\varepsilon, \quad \text{with } \Phi_\varepsilon(x) = \varepsilon^{-n}\Phi(x/\varepsilon). \tag{4.6.7}$$

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

$$(\varphi_\varepsilon * K) * f(x) = Tf * \varphi_\varepsilon(x), \tag{4.6.8}$$

where the identity holds for every x. In fact, we notice first that

$$(\varphi_\varepsilon * K_\delta) * f(x) = T_\delta f * \varphi_\varepsilon(x), \quad \text{for every } \delta > 0 \tag{4.6.9}$$

because both sides of (4.6.9) are equal for each x to the absolutely convergent double integral $\int_{z\in\mathbb{R}^n}\int_{|y|\geqslant\delta} K(y)f(z-y)\varphi_\varepsilon(x-z)dydz$. Moreover, $\varphi_\varepsilon \in L^{p'}(\mathbb{R}^n)$, with

$1/p + 1/p' = 1$, so $\varphi_\varepsilon * K_\delta \to \varphi_\varepsilon * K$ in $L^{p'}$ norm, and $T_\delta f \to Tf$ in L^p norm, as $\delta \to 0$, by Theorem 4.6.1. This proves (4.6.8), and so by (4.6.7)

$$T_\varepsilon f = K_\varepsilon * f = \varphi_\varepsilon * K * f - \Phi_\varepsilon * f = Tf * \varphi_\varepsilon - f * \Phi_\varepsilon.$$

Passing to the supremum over ε, we obtain the Cotlar inequality, and applying Theorem 3.2.12, Theorem 3.2.7 for maximal funtions and Theorem 4.6.1, we get

$$\begin{aligned}\|T^*f\|_p \leqslant & \|\sup_{\varepsilon>0} |Tf * \varphi_\varepsilon|\|_p + \|\sup_{\varepsilon>0} |f * \Phi_\varepsilon|\|_p \\ \leqslant & C\|M(Tf)\|_p + C\|Mf\|_p \leqslant C\|Tf\|_p + C\|f\|_p \leqslant C\|f\|_p.\end{aligned}$$

Thus, we have proved (c).

The second and most difficult stage of the proof is the conclusion (b). Here the argument proceeds in the main as in the proof of the weak type $(1,1)$ result for singular integrals in Theorem 4.2.1. We review it with deliberate brevity so as to avoid a repetition of details already examined.

For a given $\alpha > 0$, we split $f = g + b$ as in the proof of Theorem 4.2.1. We also consider for each cube Q_j its mate Q_j^*, which has the same center c_j but whose side length is expanded $2\sqrt{n}$ times. The following geometric remarks concerning these cubes are nearly obvious (The first one has given in the proof of Theorem 4.2.1).

(i) If $x \notin Q_j^*$, then $|x - c_j| \geqslant 2|y - c_j|$ for all $y \in Q_j$, as an obvious geometric consideration shows.

(ii) Suppose $x \in \mathbb{R}^n \setminus Q_j^*$ and assume that for some $y \in Q_j$, $|x - y| = \varepsilon$. Then the closed ball centered at x, of radius $\gamma_n\varepsilon$, contains Q_j, i.e., $B(x, r) \supset Q_j$, if $r = \gamma_n\varepsilon$.

(iii) Under the same hypotheses as (ii), we have that $|x - y| \geqslant \gamma_n'\varepsilon$, for every $y \in Q_j$.

Here γ_n and γ_n' depend only on the dimension n, and not the particular cube Q_j.

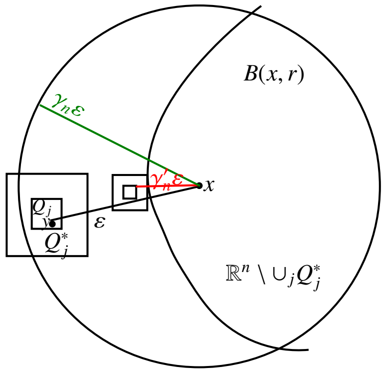


Figure 4.1: Observation for (ii) and (iii)

With these observations, and following the development in the proof of Theorem 4.2.1, we shall prove that if $x \in \mathbb{R}^n \setminus \cup_j Q_j^*$,

$$\begin{aligned}\sup_{\varepsilon>0} |T_\varepsilon b(x)| \leqslant & \sum_j \int_{Q_j} |K(x-y) - K(x-c_j)||b_j(y)|dy \\ & + C \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |b(y)|dy,\end{aligned} \tag{4.6.10}$$

with $K(x) = \frac{\Omega(x)}{|x|^n}$.

The addition of the maximal function to the r.h.s of (4.6.10) is the main new element of the proof.

To prove (4.6.10), 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

$$T_\varepsilon b(x) = \sum_j \int_{Q_j} K_\varepsilon(x-y) b_j(y) dy. \tag{4.6.11}$$

Case 1). $K_\varepsilon(x-y)=0$ if $|x-y|<\varepsilon$, and so the integral over the cube Q_j in (4.6.11) is zero.

Case 2). $K_\varepsilon(x-y)=K(x-y)$, if $|x-y|>\varepsilon$, and therefore this integral over Q_j equals

$$\int_{Q_j} K(x-y)b_j(y)dy = \int_{Q_j}[K(x-y)-K(x-c_j)]b_j(y)dy.$$

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 (4.6.10).

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 the fact that $\Omega(x)$ 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,

$$\left|\int_{Q_j} K_\varepsilon(x-y)b_j(y)dy\right| \leqslant \frac{C}{m(B(x,r))}\int_{Q_j\cap B(x,r)}|b_j(y)|dy.$$

If we add over all cubes Q_j, we finally obtain, for $r=\gamma_n\varepsilon$,

$$|T_\varepsilon b(x)| \leqslant \sum_j \int_{Q_j}|K(x-y)-K(x-c_j)||b_j(y)|dy$$
$$+\frac{C}{m(B(x,r))}\int_{B(x,r)}|b(y)|dy.$$

Taking the supremum over ε gives (4.6.10).

This inequality can be written in the form

$$|T^*b(x)| \leqslant \Sigma(x)+CMb(x), \quad x\in\mathbb{R}^n\setminus\cup_jQ_j^*,$$

and so

$$|\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:|T^*b(x)|>\alpha/2\}|$$
$$\leqslant|\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:\Sigma(x)>\alpha/4\}|+|\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:CMb(x)>\alpha/4\}|.$$

The first term in the r.h.s. is similar to (4.2.7), and we can get

$$\int_{\mathbb{R}^n\setminus\cup_jQ_j^*}\Sigma(x)dx \leqslant C\|b\|_1$$

which implies $|\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:\Sigma(x)>\alpha/4\}| \leqslant \frac{4C}{\alpha}\|b\|_1$.

For the second one, by Theorem 3.2.7, i.e., the weak type estimate for the maximal function M, we get $|\{x\in\mathbb{R}^n\setminus\cup_jQ_j^*:CMb(x)>\alpha/4\}| \leqslant \frac{C}{\alpha}\|b\|_1$.

The weak type $(1,1)$ property of T^* then follows as in the proof of the same property for T, in Theorem 4.2.1 for more details.

The final stage of the proof, the passage from the inequalities of T^* to the existence of the limits almost everywhere, follows the familiar pattern described in the proof of the Lebesgue differential theorem (i.e., Theorem 3.2.14).

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 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 4.2.4 that $T_\varepsilon f_1$ converges uniformly as $\varepsilon \to 0$, so $\Lambda f_1(x) \equiv 0$. By (4.2.11), we have $\|\Lambda f_2\|_p \leqslant 2A_p \|f_2\|_p \leqslant 2A_p \delta$ if $1 < p < \infty$. This shows $\Lambda f_2 = 0$, almost everywhere, thus by $\Lambda f(x) \leqslant \Lambda f_1(x) + \Lambda f_2(x)$, we have $\Lambda f = 0$ almost everywhere. So $\lim_{\varepsilon \to 0} T_\varepsilon f$ exists almost everywhere if $1 < p < \infty$.

In the case $p = 1$, we get similarly

$$|\{x : \Lambda f(x) > \alpha\}| \leqslant \frac{A}{\alpha} \|f_2\|_1 \leqslant \frac{A\delta}{\alpha},$$

and so again $\Lambda f(x) = 0$ almost everywhere, which implies that $\lim_{\varepsilon \to 0} T_\varepsilon f(x)$ exists almost everywhere. ■

§4.7 Vector-valued analogues

It is interesting to point out that the results of this chapter, where our functions were assumes to take real or complex values, can be extended to the case of functions taking their values in a Hilbert space. We present this generalization because it can be put to good use in several problems. An indication of this usefulness is given in the Littlewood-Paley theory.

We begin by reviewing quickly certain aspects of integration theory in this context.

Let $\mathcal{H}$ 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 non-negative 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 $L(\mathcal{H}_1, \mathcal{H}_2)$ 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 $L(\mathcal{H}_1, \mathcal{H}_2)$ 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 $L^p(\mathbb{R}^n, L(\mathcal{H}_1, \mathcal{H}_2))$, as before; here again $|\cdot|$ denotes the norm, this time in $L(\mathcal{H}_1, \mathcal{H}_2)$.

The usual facts about convolution hold in this setting. For example, let $f \in L^p(\mathbb{R}^n, \mathcal{H}_1)$ and $K \in L^q(\mathbb{R}^n, L(\mathcal{H}_1, \mathcal{H}_2))$, 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

$$|g(x)| \leqslant \int_{\mathbb{R}^n} |K(x-y) f(y)| dy \leqslant \int_{\mathbb{R}^n} |K(x-y)| |f(y)| dy.$$

Also $\|g\|_r \leqslant \|K\|_q \|f\|_p$, if $1/r = 1/p + 1/q - 1$, with $1 \leqslant r \leqslant \infty$.

Suppose that $f \in L^1(\mathbb{R}^n, \mathcal{H})$. Then we can define its Fourier transform

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

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 $L(\mathcal{H}_1, \mathcal{H}_2)$. Then

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

whenever defined, takes values in $\mathcal{H}_2$.

Theorem 4.7.1.

The results in this chapter, in particular Theorems 4.2.1, 4.2.4, 4.6.1 and 4.6.2, and Proposition 4.2.2 are valid in the more general context where f takes its value in $\mathcal{H}_1$, K takes its values in $L(\mathcal{H}_1, \mathcal{H}_2)$ and Tf and $T_\varepsilon f$ take their value in $\mathcal{H}_2$, and where throughout the absolute value $|\cdot|$ is replaced by the appropriate norm in $\mathcal{H}_1$, $L(\mathcal{H}_1, \mathcal{H}_2)$ 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 case, if we take into account the remarks made in the above paragraphs. So, we leave the proof to the interested reader.

Remark 4.7.2. 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 case.

2) Most of the argument goes through in the even greater generality of Banach space-valued functions, appropriately defined, one can refer to [Gra14, pp.385-414]. The Hilbert space structure is used only in the L^2 theory when applying the variant of Plancherel's formula.

The Hilbert space structure also enters in the following corollary.

Corollary 4.7.3.

With the same assumptions as in Theorem 4.7.1, 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 the $L^2(\mathbb{R}^n, \mathcal{H}_j)$ 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 so by the general theory of inner products 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

$$\langle Tf_1, f_2\rangle_2 = \langle f_1, T^* f_2\rangle_1, \quad \text{with } f_j \in L^2(\mathbb{R}^n, \mathcal{H}_j).$$

But our assumption is equivalent with the identity (see the theory of Hilbert spaces, e.g., [Din07, Chapter 6])

$$\langle Tf, Tg\rangle_2 = c^2\langle f, g\rangle_1, \quad \text{for all } f, g \in L^2(\mathbb{R}^n, \mathcal{H}_1).$$

Thus using the definition of the adjoint, $\langle T^*Tf, g\rangle_1 = c^2\langle f, g\rangle_1$, and so the assumption can be restated as

$$T^*Tf = c^2 f, \quad f \in L^2(\mathbb{R}^n, \mathcal{H}_1). \tag{4.7.1}$$

T^* is again an operator of the same kind as T but it takes 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 $L(\mathcal{H}_1, \mathcal{H}_2)$.

This is obvious on the formal level since

$$\begin{aligned}\langle Tf_1, f_2\rangle_2 &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (K(x-y)f_1(y), f_2(x))_2 dydx \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (f_1(y), K^*(-(y-x))f_2(x))_1 dxdy = \langle f_1, T^*f_2\rangle_1.\end{aligned}$$

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 (4.7.1),

$$c^2\|f\|_p = \|T^*Tf\|_p \leqslant A_p\|Tf\|_p.$$

This proves the corollary with $A_p' = A_p/c^2$. ■

Remark 4.7.4. This corollary applies in particular to the singular integrals commuted with dilations, then the condition required is that the multiplier $m(\xi)$ have constant absolute value. This is the case, for example, when T is the Hilbert transform, $K(x) = \frac{1}{\pi x}$, and $m(\xi) = -i\,\mathrm{sgn}\,(\omega)\,\mathrm{sgn}\,(\xi)$.

References

[Din07] Guanggui Ding. *New talk of functional analysis*. Science press, Beijing, 2007.

[Duo01] Javier Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.

[GR07] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam, seventh edition, 2007. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger, With one CD-ROM (Windows, Macintosh and UNIX).

[Gra14] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[Mia04] Changxing Miao. *Harmonic analysis and its application in PDEs*, volume 89 of *Fundamental Series in Mordon Mathematics*. Science Press, Beijing, 2nd edition, 2004.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[SW71] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

V
Littlewood-Paley Theory and Multiplier Theorems

In harmonic analysis, Littlewood-Paley theory is a term used to describe a theoretical framework used to extend certain results about L^2 functions to L^p functions for $1 < p < \infty$. It is typically used as a substitute for orthogonality arguments which only apply to L^p functions when $p = 2$. One implementation involves studying a function by decomposing it in terms of functions with localized frequencies, and using the Littlewood-Paley g-function to compare it with its Poisson integral. The one-variable case was originated by Littlewood and Paley (1931, 1937, 1938) and developed further by Zygmund and Marcinkiewicz in the 1930s using complex function theory (Zygmund 2002 [1935], chapters XIV, XV). Stein later extended the theory to higher dimensions using real variable techniques.

§ 5.1 Three approach functions and L^p boundedness

The g-function is a nonlinear operator which allows one to give a useful characterization of the L^p norm of a function on $\mathbb{R}^n$ in terms of the behavior of its Poisson integral. This characterization will be used not only in this chapter, but also in the succeeding chapter dealing with function spaces.

Let $f \in L^p(\mathbb{R}^n)$ and write $u(x, y)$ for its Poisson integral

$$u(x,y) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{\omega i\xi\cdot x} e^{-|\omega\xi| y} \widehat{f}(\xi) d\xi = \int_{\mathbb{R}^n} P_y(t) f(x-t) dt = P_y * f(x) \tag{5.1.1}$$

as defined in (4.1.2) and (4.1.4). Let Δ denote the Laplace operator in $\mathbb{R}^{n+1}_+$, i.e., $\Delta = \frac{\partial^2}{\partial y^2} + \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$; ∇ is the corresponding gradient, $|\nabla u(x,y)|^2 = |\frac{\partial u}{\partial y}|^2 + |\nabla_x u(x,y)|^2$, where $|\nabla_x u(x,y)|^2 = \sum_{j=1}^n |\frac{\partial u}{\partial x_j}|^2$.

Definition 5.1.1.

With the above notations, we define the ***Littlewood-Paley g-function*** $g(f)(x)$, by

$$g(f)(x) = \left(\int_0^\infty |\nabla u(x,y)|^2 y dy\right)^{1/2}. \tag{5.1.2}$$

We can also define two ***partial g-functions***, one dealing with the y differentiation and the other with the x differentiation, i.e.,

$$g_1(f)(x) = \left(\int_0^\infty \left|\frac{\partial u}{\partial y}(x,y)\right|^2 ydy\right)^{1/2}, \; g_x(f)(x) = \left(\int_0^\infty |\nabla_x u(x,y)|^2 ydy\right)^{1/2}. \tag{5.1.3}$$

Obviously, $g^2 = g_1^2 + g_x^2$.

The basic result for g is as follows.

Theorem 5.1.2.

Suppose $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Then we have $g(f) \in L^p(\mathbb{R}^n)$, and

$$A'_p\|f\|_p \leqslant \|g(f)\|_p \leqslant A_p\|f\|_p. \tag{5.1.4}$$

Proof. Step 1: We first consider the simple case $p = 2$. For $f \in L^2(\mathbb{R}^n)$, we have

$$\|g(f)\|_2^2 = \int_{\mathbb{R}^n}\int_0^\infty |\nabla u(x,y)|^2 ydydx = \int_0^\infty y\int_{\mathbb{R}^n} |\nabla u(x,y)|^2 dxdy.$$

In view of the identity (5.1.1), we have

$$\frac{\partial u}{\partial y} = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} -|\omega\xi|\widehat{f}(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi,$$

and

$$\frac{\partial u}{\partial x_j} = \left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^n} \omega i\xi_j\widehat{f}(\xi)e^{\omega i\xi\cdot x}e^{-|\omega\xi|y}d\xi.$$

It follows from Plancherel's formula that

$$\begin{aligned}
\int_{\mathbb{R}^n}|\nabla u(x,y)|^2dx =& \int_{\mathbb{R}^n}\left[\left|\frac{\partial u}{\partial y}\right|^2 + \sum_{j=1}^n\left|\frac{\partial u}{\partial x_j}\right|^2\right]dx \\
=&\left\|\frac{\partial u}{\partial y}\right\|_{L_x^2}^2 + \sum_{j=1}^n\left\|\frac{\partial u}{\partial x_j}\right\|_{L_x^2}^2 \\
=&\|\mathscr{F}^{-1}(-|\omega\xi|\widehat{f}(\xi)e^{-|\omega\xi|y})\|_2^2 + \sum_{j=1}^n\|\mathscr{F}^{-1}(\omega i\xi_j\widehat{f}(\xi)e^{-|\omega\xi|y})\|_2^2 \\
=&\|-|\omega\xi|\widehat{f}(\xi)e^{-|\omega\xi|y}\|_2^2 + \sum_{j=1}^n\|\omega i\xi_j\widehat{f}(\xi)e^{-|\omega\xi|y}\|_2^2 \\
=&2\omega^2\||\xi|\widehat{f}(\xi)e^{-|\omega\xi|y}\|_2^2 \\
=&\int_{\mathbb{R}^n}2\omega^2|\xi|^2|\widehat{f}(\xi)|^2e^{-2|\omega\xi|y}d\xi,
\end{aligned}$$

and by integration by parts,

$$\begin{aligned}
\|g(f)\|_2^2 =& \int_0^\infty y\int_{\mathbb{R}^n}2\omega^2|\xi|^2|\widehat{f}(\xi)|^2e^{-2|\omega\xi|y}d\xi dy \\
=&\int_{\mathbb{R}^n}2\omega^2|\xi|^2|\widehat{f}(\xi)|^2\int_0^\infty ye^{-2|\omega\xi|y}dyd\xi \\
=&\int_{\mathbb{R}^n}2\omega^2|\xi|^2|\widehat{f}(\xi)|^2\frac{1}{4\omega^2|\xi|^2}d\xi
\end{aligned}$$

$$=\frac{1}{2}\|\widehat{f}\|_2^2=\frac{1}{2}\|f\|_2^2.$$

Hence, we get

$$\|g(f)\|_2=2^{-1/2}\|f\|_2. \tag{5.1.5}$$

We have also obtained $\|g_1(f)\|_2=\|g_x(f)\|_2=\frac{1}{2}\|f\|_2$.

Step 2: We consider the case $p\neq 2$ and prove $\|g(f)\|_p\leqslant A_p\|f\|_p$. We define the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ which are to be considered now. Let $\mathcal{H}_1$ be the one-dimensional Hilbert space of complex numbers. To define $\mathcal{H}_2$, we define first $\mathcal{H}_2^0$ as the L^2 space on $(0,\infty)$ with measure ydy, i.e.,

$$\mathcal{H}_2^0=\left\{f:|f|^2=\int_0^\infty |f(y)|^2ydy<\infty\right\}.$$

Let $\mathcal{H}_2$ be the direct sum of $n+1$ copies of $\mathcal{H}_2^0$; so the elements of $\mathcal{H}_2$ can be represented as $(n+1)$ component vectors whose entries belong to $\mathcal{H}_2^0$. Since $\mathcal{H}_1$ is the same as the complex numbers, $L(\mathcal{H}_1,\mathcal{H}_2)$ is of course identifiable with $\mathcal{H}_2$.

Now let $\varepsilon>0$, and keep it temporarily fixed. Define

$$K_\varepsilon(x)=\left(\frac{\partial P_{y+\varepsilon}(x)}{\partial y},\frac{\partial P_{y+\varepsilon}(x)}{\partial x_1},\cdots,\frac{\partial P_{y+\varepsilon}(x)}{\partial x_n}\right).$$

Notice that for each fixed x, $K_\varepsilon(x)\in\mathcal{H}_2$. This is the same as saying that

$$\int_0^\infty\left|\frac{\partial P_{y+\varepsilon}(x)}{\partial y}\right|^2 ydy<\infty \text{ and } \int_0^\infty\left|\frac{\partial P_{y+\varepsilon}(x)}{\partial x_j}\right|^2 ydy<\infty, \text{ for } j=1,...,n.$$

In fact, $P_y(x)=\frac{c_ny}{(|x|^2+y^2)^{(n+1)/2}}$ implies that both $\frac{\partial P_y}{\partial y}$ and $\frac{\partial P_y}{\partial x_j}$ are bounded by $\frac{A}{(|x|^2+y^2)^{(n+1)/2}}$. For the norm of $K_\varepsilon(x)$ in $\mathcal{H}_2$, we have

$$\begin{aligned}|K_\varepsilon(x)|^2\leqslant &A^2(n+1)\int_0^\infty\frac{ydy}{(|x|^2+(y+\varepsilon)^2)^{n+1}}\\ \leqslant &A^2(n+1)\int_0^\infty\frac{dy}{(y+\varepsilon)^{2n+1}}\leqslant C_\varepsilon,\end{aligned}$$

and then

$$|K_\varepsilon(x)|\in L^1_{\text{loc}}(\mathbb{R}^n\setminus\{0\}). \tag{5.1.6}$$

Similarly,

$$\left|\frac{\partial K_\varepsilon(x)}{\partial x_j}\right|^2\leqslant C\int_\varepsilon^\infty\frac{ydy}{(|x|^2+y^2)^{n+2}}\leqslant C|x|^{-2n-2},$$

thus, K_ε satisfies the gradient condition, i.e.,

$$\left|\frac{\partial K_\varepsilon(x)}{\partial x_j}\right|\leqslant C|x|^{-(n+1)}, \tag{5.1.7}$$

with C independent of ε.

Now we consider the operator T_ε defined by

$$T_\varepsilon f(x)=K_\varepsilon*f(x)=\nabla P_{y+\varepsilon}*f(x)=\nabla(P_{y+\varepsilon}*f)(x)=\nabla u(x,y+\varepsilon).$$

The function f is complex-valued (take its value in $\mathcal{H}_1$), but $T_\varepsilon f(x)$ takes its value in $\mathcal{H}_2$. Observe that

$$|T_\varepsilon f(x)|=\left(\int_0^\infty|\nabla u(x,y+\varepsilon)|^2ydy\right)^{\frac{1}{2}}\leqslant\left(\int_\varepsilon^\infty|\nabla u(x,y)|^2ydy\right)^{\frac{1}{2}}\leqslant g(f)(x). \tag{5.1.8}$$

Hence, $\|T_\varepsilon f\|_2 \leqslant 2^{-1/2}\|f\|_2$, if $f \in L^2(\mathbb{R}^n)$, by (5.1.5). Therefore, by Theorem 2.5.6, we get

$$\left(\frac{|\omega|}{2\pi}\right)^{-n/2}|\widehat{K}_\varepsilon(x)| \leqslant \left\|\left(\frac{|\omega|}{2\pi}\right)^{-n/2}|\widehat{K}_\varepsilon(x)|\right\|_{L^\infty_x(\mathbb{R}^n)} = \|T\| \leqslant 2^{-1/2}. \tag{5.1.9}$$

Because of (5.1.6), (5.1.7) and (5.1.9), by Theorem 4.7.1 (cf. Theorem 4.2.1 and Proposition 4.2.2), we get $\|T_\varepsilon f\|_p \leqslant A_p\|f\|_p$, $1 < p < \infty$ with A_p independent of ε. By (5.1.8), for each x, $|T_\varepsilon f(x)|$ increases to $g(f)(x)$, as $\varepsilon \to 0$, thus we obtain finally

$$\|g(f)\|_p \leqslant A_p\|f\|_p, \quad 1 < p < \infty. \tag{5.1.10}$$

Step 3: To derive the converse inequalities:

$$A'_p\|f\|_p \leqslant \|g(f)\|_p, \quad 1 < p < \infty. \tag{5.1.11}$$

In the first step, we have shown that $\|g_1(f)\|_2 = \frac{1}{2}\|f\|_2$ for $f \in L^2(\mathbb{R}^n)$. Let u_1 and u_2 be the Poisson integrals of $f_1, f_2 \in L^2$, respectively. Then we have $\|g_1(f_1 + f_2)\|_2^2 = \frac{1}{4}\|f_1 + f_2\|_2^2$, i.e.,

$$\int_{\mathbb{R}^n}\int_0^\infty \left|\frac{\partial(u_1+u_2)}{\partial y}\right|^2 ydydx = \frac{1}{4}\int_{\mathbb{R}^n}|f_1+f_2|^2dx.$$

It leads to the identity

$$4\int_{\mathbb{R}^n}\int_0^\infty \frac{\partial u_1}{\partial y}(x,y)\overline{\frac{\partial u_2}{\partial y}(x,y)}ydydx = \int_{\mathbb{R}^n} f_1(x)\overline{f_2(x)}dx,$$

which, in turn, leads to the inequality, by Hölder's inequality and the definition of g_1,

$$\frac{1}{4}\left|\int_{\mathbb{R}^n} f_1(x)\overline{f_2(x)}dx\right| \leqslant \int_{\mathbb{R}^n} g_1(f_1)(x)g_1(f_2)(x)dx.$$

Suppose now in addition that $f_1 \in L^p(\mathbb{R}^n)$ and $f_2 \in L^{p'}(\mathbb{R}^n)$ with $\|f_2\|_{p'} \leqslant 1$ and $1/p + 1/p' = 1$. Then by Hölder inequality and the result (5.1.10), we get

$$\left|\int_{\mathbb{R}^n} f_1(x)\overline{f_2(x)}dx\right| \leqslant 4\|g_1(f_1)\|_p\|g_1(f_2)\|_{p'} \leqslant 4A_{p'}\|g_1(f_1)\|_p. \tag{5.1.12}$$

Now we take the supremum in (5.1.12) as f_2 ranges over all function in $L^2 \cap L^{p'}$, with $\|f_2\|_{p'} \leqslant 1$. Then, we obtain the desired result (5.1.11), with $A'_p = 1/4A_{p'}$, but where f is restricted to be in $L^2 \cap L^p$. The passage to the general case is provided by an easy limiting argument. Let f_m be a sequence of functions in $L^2 \cap L^p$, which converges in L^p norm to f. Notice that

$$\begin{aligned}|g(f_m)(x) - g(f_n)(x)| =& \left|\|\nabla u_m\|_{L^2(0,\infty;ydy)} - \|\nabla u_n\|_{L^2(0,\infty;ydy)}\right| \\ \leqslant& \|\nabla u_m - \nabla u_n\|_{L^2(0,\infty;ydy)} \\ =& g(f_m - f_n)(x)\end{aligned}$$

by the triangle inequality. Thus, $\{g(f_m)\}$ is a Cauchy sequence in L^p and so converges to $g(f)$ in L^p, and we obtain the inequality (5.1.11) for f as a result of the corresponding inequalities for f_m. ∎

We have incidentally also proved the following, which we state as a corollary.

Corollary 5.1.3.

Suppose $f \in L^2(\mathbb{R}^n)$, and $g_1(f) \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Then $f \in L^p(\mathbb{R}^n)$, and $A'_p\|f\|_p \leqslant \|g_1(f)\|_p$.

Remark 5.1.4. There are some very simple variants of the above that should be pointed out:

(i) The results hold also with $g_x(f)$ instead of $g(f)$. The direct inequality $\|g_x(f)\|_p \leqslant A_p\|f\|_p$ is of course a consequence of the one for g. The converse inequality is then proved in the same way as that for g_1.

(ii) For any integer $k > 1$, define

$$g_k(f)(x) = \left(\int_0^\infty \left|\frac{\partial^k u}{\partial y^k}(x,y)\right|^2 y^{2k-1}dy\right)^{1/2}.$$

Then the L^p inequalities hold for g_k as well. Both (i) and (ii) are stated more systematically in [Ste70, Chapter IV, §7.2, p.112-113].

(iii) For later purpose, it will be useful to note that for each x, $g_k(f)(x) \geqslant A_k g_1(f)(x)$ where the bound A_k depends only on k.

It is easily verified from the Poisson integral formula that if $f \in L^p(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$, then

$$\frac{\partial^k u(x,y)}{\partial y^k} \to 0 \text{ for each } x, \quad \text{as } y \to \infty,$$

which yields

$$\frac{\partial^k u(x,y)}{\partial y^k} = -\int_y^\infty \frac{\partial^{k+1}u(x,s)}{\partial s^{k+1}} s^k \frac{ds}{s^k}.$$

By Schwarz's inequality, we get

$$\begin{aligned}\left|\frac{\partial^k u(x,y)}{\partial y^k}\right|^2 \leqslant& \left(\int_y^\infty \left|\frac{\partial^{k+1}u(x,s)}{\partial s^{k+1}}\right|^2 s^{2k}ds\right)\left(\int_y^\infty s^{-2k}ds\right)\\ =&\frac{1}{2k-1}y^{-2k+1}\left(\int_y^\infty \left|\frac{\partial^{k+1}u(x,s)}{\partial s^{k+1}}\right|^2 s^{2k}ds\right).\end{aligned}$$

Hence, by Fubini's theorem, we have

$$\begin{aligned}(g_k(f)(x))^2 =& \int_0^\infty \left|\frac{\partial^k u}{\partial y^k}(x,y)\right|^2 y^{2k-1}dy\\ \leqslant&\frac{1}{2k-1}\int_0^\infty\left(\int_y^\infty \left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2k}ds\right)dy\\ =&\frac{1}{2k-1}\int_0^\infty\left(\int_0^s dy\right)\left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2k}ds\\ =&\frac{1}{2k-1}\int_0^\infty \left|\frac{\partial^{k+1}u}{\partial s^{k+1}}(x,s)\right|^2 s^{2k+1}ds\\ =&\frac{1}{2k-1}(g_{k+1}(f)(x))^2.\end{aligned}$$

Thus, the assertion is proved by the induction on k.

The proof given for the L^p inequalities of the g-function did not, in any essential way, depend on the theory of harmonic functions, despite the fact that this function was defined in terms of the Poisson integral. In effect, all that was really used are the fact that the Poisson kernels are suitable approximations to the identity.

There is, however, another approach, which can be carried out without recourse to the theory of singular integrals, but which leans heavily on characteristic properties of harmonic functions. We present it here (more precisely, we present that part which deals with $1 < p \leqslant 2$, for the inequality (5.1.10)), because its ideas can be adapted to other situations where the methods of Chapter 4 are not applicable. Everything will be based on the following three observations.

Lemma 5.1.5.

Suppose u is harmonic and strictly positive. Then

$$\Delta u^p = p(p-1)u^{p-2}|\nabla u|^2. \tag{5.1.13}$$

Proof. The proof is straightforward. Indeed,

$$\partial_{x_j} u^p = pu^{p-1}\partial_{x_j} u, \quad \partial^2_{x_j} u^p = p(p-1)u^{p-2}(\partial_{x_j} u)^2 + pu^{p-1}\partial^2_{x_j} u,$$

which implies by summation

$$\Delta u^p = p(p-1)u^{p-2}|\nabla u|^2 + pu^{p-1}\Delta u = p(p-1)u^{p-2}|\nabla u|^2,$$

since $\Delta u = 0$. ■

Lemma 5.1.6.

Suppose that $F(x,y) \in \mathcal{C}(\overline{\mathbb{R}^{n+1}_+}) \cap \mathcal{C}^2(\mathbb{R}^{n+1}_+)$ satisfies $\Delta F \geqslant 0$, and for some $\varepsilon > 0$, $|F| = O(r^{-n-\varepsilon})$ and $|\nabla F| = O(r^{-n-1-\varepsilon})$ as $r = |(x,y)| \to \infty$. Then

$$\int_{\mathbb{R}^{n+1}_+} y\Delta F(x,y)dxdy = \int_{\mathbb{R}^n} F(x,0)dx. \tag{5.1.14}$$

Proof. We use Green's theorem

$$\int_D (u\Delta v - v\Delta u)dxdy = \int_{\partial D}\left(u\frac{\partial v}{\partial \mathcal{N}} - v\frac{\partial u}{\partial \mathcal{N}}\right)d\sigma$$

where $D = B_r \cap \mathbb{R}^{n+1}_+$, with B_r the ball of radius r in $\mathbb{R}^{n+1}$ centered at the origin, and $\mathcal{N}$ is the outward normal vector. We take $v = F$, and $u = y$ to obtain

$$\int_D (y\Delta F - F\Delta y)dxdy = \int_{\partial D}\left(y\frac{\partial F}{\partial \mathcal{N}} - F\frac{\partial y}{\partial \mathcal{N}}\right)d\sigma,$$

i.e.,

$$\int_D y\Delta F dxdy = \int_{\partial D_0}\left(y\frac{\partial F}{\partial \mathcal{N}} - F\frac{\partial y}{\partial \mathcal{N}}\right)d\sigma + \int_{\mathbb{R}^n} F(x,0)dx, \tag{5.1.15}$$

due to $\Delta y = 0$ in D and $\frac{\partial y}{\partial \mathcal{N}} = -1$ on the boundary $\left\{(x,y) \in \mathbb{R}^{n+1}_+ : y = 0\right\} = \mathbb{R}^n$, where ∂D_0 is the spherical part of the boundary of D. Since $\Delta F \geqslant 0$, by Levi's monotone convergence theorem, we get

$$\int_D y\Delta F(x,y)dxdy = \int y\Delta F(x,y)\chi_D(x,y)dxdy \to \int_{\mathbb{R}^{n+1}_+} y\Delta F(x,y)dxdy, \tag{5.1.16}$$

as $r \to \infty$. Let $y = r\sin\theta$ on ∂D_0 with $\theta \in [0,\pi]$, we have

$$\begin{aligned}\int_{\partial D_0}\left(y\frac{\partial F}{\partial \mathcal{N}} - F\frac{\partial y}{\partial \mathcal{N}}\right)d\sigma =& \int_{\partial D_0}\left(r\sin\theta\frac{\partial F}{\partial \mathcal{N}} - F\sin\theta\right)d\sigma\\ =& r^n\int_{S^n_+}\left(r\sin\theta\frac{\partial F}{\partial \mathcal{N}} - F\sin\theta\right)d\sigma\\ =& O(r^{-\varepsilon})\int_{S^n_+}\sin\theta d\sigma \leqslant \frac{\omega_n}{2}O(r^{-\varepsilon}) \to 0,\end{aligned} \tag{5.1.17}$$

as $r \to \infty$. Thus, combining (5.1.15), (5.1.16) and (5.1.17), we obtain the desired result (5.1.14). ■

Lemma 5.1.7.

If $u(x,y)$ is the Poisson integral of f, then

$$\sup_{y>0} |u(x,y)| \leqslant Mf(x). \tag{5.1.18}$$

Proof. This is the same as the part (a) of Theorem 4.1.3. It can be proved with a similar argument as in the proof of Theorem 3.2.12. ■

Now we use these lemmas to give another proof for the inequality

$$\|g(f)\|_p \leqslant A_p \|f\|_p, \quad 1 < p \leqslant 2.$$

Another proof of $\|g(f)\|_p \leqslant A_p\|f\|_p$, $1 < p \leqslant 2$. Suppose first $0 \leqslant f \in \mathscr{D}(\mathbb{R}^n)$ (and at least $f \neq 0$ on a nonzero measurable set). Then the Poisson integral u of f, $u(x,y) = \int_{\mathbb{R}^n} P_y(t) f(x-t)dt > 0$, since $P_y > 0$ for any $x \in \mathbb{R}^n$ and $y > 0$; and the majorizations $u^p(x,y) = O(r^{-np})$ and $|\nabla u^p| = O(r^{-np-1})$, as $r = |(x,y)| \to \infty$ are valid. We have, by Lemmas 5.1.5 and 5.1.7, and the hypothesis $1 < p \leqslant 2$,

$$\begin{aligned}(g(f)(x))^2 &= \int_0^\infty y|\nabla u(x,y)|^2 dy = \frac{1}{p(p-1)} \int_0^\infty y u^{2-p} \Delta u^p dy \\ &\leqslant \frac{[Mf(x)]^{2-p}}{p(p-1)} \int_0^\infty y \Delta u^p dy.\end{aligned}$$

We can write this as

$$g(f)(x) \leqslant C_p (Mf(x))^{(2-p)/2} (I(x))^{1/2}, \tag{5.1.19}$$

where $I(x) = \int_0^\infty y \Delta u^p dy$. However, by Lemma 5.1.6,

$$\int_{\mathbb{R}^n} I(x)dx = \int_{\mathbb{R}^{n+1}_+} y\Delta u^p dydx = \int_{\mathbb{R}^n} u^p(x,0)dx = \|f\|_p^p. \tag{5.1.20}$$

This immediately gives the desired result for $p = 2$.

Next, suppose $1 < p < 2$. By (5.1.19), Hölder's inequality, Theorem 3.2.7 and (5.1.20), we have, for $0 \leqslant f \in \mathscr{D}(\mathbb{R}^n)$,

$$\begin{aligned}&\int_{\mathbb{R}^n} (g(f)(x))^p dx \leqslant C_p^p \int_{\mathbb{R}^n} (Mf(x))^{p(2-p)/2} (I(x))^{p/2} dx \\ \leqslant & C_p^p \left(\int_{\mathbb{R}^n} (Mf(x))^p dx\right)^{1/r'} \left(\int_{\mathbb{R}^n} I(x)dx\right)^{1/r} \leqslant C_p' \|f\|_p^{p/r'} \|f\|_p^{p/r} = C_p' \|f\|_p^p,\end{aligned}$$

where $r = 2/p \in (1,2)$ and $1/r + 1/r' = 1$, then $r' = 2/(2-p)$.

Thus, $\|g(f)\|_p \leqslant A_p\|f\|_p$, $1 < p \leqslant 2$, whenever $0 \leqslant f \in \mathscr{D}(\mathbb{R}^n)$.

For general $f \in L^p(\mathbb{R}^n)$ (which we assume for simplicity to be real-valued), write $f = f^+ - f^-$ as its decomposition into positive and negative part; then we need only approximate in norm f^+ and f^-, each by a sequences of positive functions in $\mathscr{D}(\mathbb{R}^n)$. We omit the routine details that are needed to complete the proof. ■

Unfortunately, the elegant argument just given is not valid for $p > 2$. There is, however, a more intricate variant of the same idea which does work for the case $p > 2$, but we do not intend to reproduce it here.

We shall, however, use the ideas above to obtain a significant generalization of the inequality for the g-functions.

Definition 5.1.8.

Define the positive function

$$(g_\lambda^*(f)(x))^2 = \int_0^\infty \int_{\mathbb{R}^n} \left(\frac{y}{|t|+y}\right)^{\lambda n} |\nabla u(x-t,y)|^2 y^{1-n} dt dy. \qquad (5.1.21)$$

Before going any further, we shall make a few comments that will help to clarify the meaning of the complicated expression (5.1.21).

First, $g_\lambda^*(f)(x)$ will turn out to be a pointwise majorant of $g(f)(x)$. To understand this situation better we have to introduce still another quantity, which is roughly midway between g and g_λ^*. It is defined as follows.

Definition 5.1.9.

Let Γ be a fixed proper cone in $\mathbb{R}_+^{n+1}$ with vertex at the origin and which contains $(0,1)$ in its interior. The exact form of Γ will not really matter, but for the sake of definiteness let us choose for Γ the up circular cone:

$$\Gamma = \left\{(t,y) \in \mathbb{R}_+^{n+1} : |t| < y, y > 0\right\}.$$

For any $x \in \mathbb{R}^n$, let $\Gamma(x)$ be the cone Γ translated such that its vertex is at x. Now define the positive ***Lusin's S-function*** $S(f)(x)$ by

$$[S(f)(x)]^2 = \int_{\Gamma(x)} |\nabla u(t,y)|^2 y^{1-n} dy dt = \int_\Gamma |\nabla u(x-t,y)|^2 y^{1-n} dy dt. \qquad (5.1.22)$$

We assert, as we shall momentarily prove, that

Proposition 5.1.10.

$$g(f)(x) \leqslant CS(f)(x) \leqslant C_\lambda g_\lambda^*(f)(x). \qquad (5.1.23)$$

What interpretation can we put on the inequalities relating these three quantities? A hint is afforded by considering three corresponding approaches to the boundary for harmonic functions.

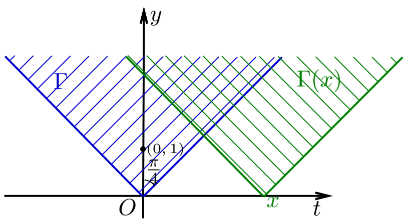


Figure 5.1: Γ and $\Gamma(x)$ for $n=1$

(a) With $u(x,y)$ the Poisson integral of $f(x)$, the simplest approach to the boundary point $x \in \mathbb{R}^n$ is obtained by letting $y \to 0$, (with x fixed). This is the ***perpendicular approach***, and for it the appropriate limit exists almost everywhere, as we already know.

(b) Wider scope is obtained by allowing the variable point (t,y) to approach $(x,0)$ through any cone $\Gamma(x)$ whose vertex is x. This is the ***nontangential approach*** which will be so important for us later. As the reader may have already realized, the relation of the S-function to the g-function is in some sense analogous to the relation between the nontangential and the perpendicular approaches; we should add that the S-function is of decisive significance in its own right, but we shall not pursue that matter now.

(c) Finally, the widest scope is obtained by allowing the variable point (t, y) to approach $(x, 0)$ in an arbitrary manner, i.e., the ***unrestricted approach***. The function g_λ^* has the analogous role: it takes into account the unrestricted approach for Poisson integrals.

Notice that $g_\lambda^*(x)$ depends on λ. For each x, the smaller λ the greater $g_\lambda^*(x)$, and this behavior is such that that L^p boundedness of g_λ^* depends critically on the correct relation between p and λ. This last point is probably the main interest in g_λ^*, and is what makes its study more difficult than g or S.

After these various heuristic and imprecise indications, let us return to firm ground. The only thing for us to prove here is the assertion (5.1.23).

Proof of Proposition 5.1.10. The inequality $S(f)(x) \leqslant C_\lambda g_\lambda^*(f)(x)$ is obvious, since the integral (5.1.21) majorizes that part of the integral taken only over Γ, and

$$\left(\frac{y}{|t|+y}\right)^{\lambda n} \geqslant \frac{1}{2^{\lambda n}}$$

since $|t| < y$ there. The non-trivial part of the assertion is:

$$g(f)(x) \leqslant CS(f)(x).$$

It suffices to prove this inequality for $x = 0$. Let us denote by B_y the ball in $\mathbb{R}_+^{n+1}$ centered at $(0, y)$ and tangent to the boundary of the cone Γ; the radius of B_y is then proportional to y. Now the partial derivatives $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial x_k}$ are, like u, harmonic functions. Thus, by the mean value theorem of harmonic functions (i.e., Theorem 4.4.8 by noticing $(0, y)$ is the center of B_y), we get

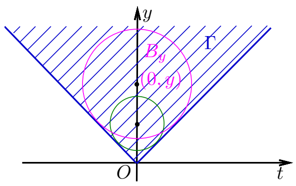


Figure 5.2: Γ and B_y

$$\frac{\partial u(0,y)}{\partial y} = \frac{1}{|B_y|}\int_{B_y} \frac{\partial u(x,s)}{\partial s} dxds$$

where $|B_y|$ is the $n+1$ dimensional measure of B_y, i.e., $|B_y| = cy^{n+1}$ for an appropriate constant c. By Schwarz's inequality, we have

$$\begin{aligned}\left|\frac{\partial u(0,y)}{\partial y}\right|^2 \leqslant & \frac{1}{|B_y|^2}\int_{B_y}\left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds \int_{B_y} dxds \\ = & \frac{1}{|B_y|}\int_{B_y}\left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds.\end{aligned}$$

If we integrate this inequality, we obtain

$$\int_0^\infty y\left|\frac{\partial u(0,y)}{\partial y}\right|^2 dy \leqslant \int_0^\infty c^{-1}y^{-n}\left(\int_{B_y}\left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds\right) dy.$$

However, $(x, s) \in B_y$ clearly implies that $c_1 s \leqslant y \leqslant c_2 s$, for two positive constants c_1 and c_2. Thus, apart from a multiplicative factor by changing the order of the double integrals, the last integral is majorized by

$$\int_\Gamma\left(\int_{c_1 s}^{c_2 s} y^{-n}dy\right)\left|\frac{\partial u(x,s)}{\partial s}\right|^2 dxds \leqslant c'\int_\Gamma\left|\frac{\partial u(x,s)}{\partial s}\right|^2 s^{1-n}dxds.$$

This is another way of saying that,

$$\int_0^\infty y\left|\frac{\partial u(0,y)}{\partial y}\right|^2 dy \leqslant c''\int_\Gamma\left|\frac{\partial u(x,y)}{\partial y}\right|^2 y^{1-n}dxdy.$$

The same is true for the derivatives $\frac{\partial u}{\partial x_j}$, $j = 1, ..., n$, and adding the corresponding estimates proves our assertion. ■

We are now in a position to state the L^p estimates concerning g_λ^*.

Theorem 5.1.11.

Let $\lambda > 1$ be a parameter. Suppose $f \in L^p(\mathbb{R}^n)$. Then,
(a) For every $x \in \mathbb{R}^n$, $g(f)(x) \leqslant C_\lambda g_\lambda^*(f)(x)$.
(b) If $1 < p < \infty$, and $p > 2/\lambda$, then

$$\|g_\lambda^*(f)\|_p \leqslant A_{p,\lambda}\|f\|_p. \tag{5.1.24}$$

Proof. The part (a) has already been proved in Proposition 5.1.10. Now, we prove (b).

For the case $p \geqslant 2$, only the assumption $\lambda > 1$ is relevant since $2/\lambda < 2 \leqslant p$.

Let ψ denote a positive function on $\mathbb{R}^n$, we claim that

$$\int_{\mathbb{R}^n} (g_\lambda^*(f)(x))^2\psi(x)dx \leqslant A_\lambda \int_{\mathbb{R}^n} (g(f)(x))^2(M\psi)(x)dx. \tag{5.1.25}$$

The l.h.s. of (5.1.25) equals

$$\int_0^\infty \int_{t\in\mathbb{R}^n} y|\nabla u(t,y)|^2 \left[\int_{x\in\mathbb{R}^n} \frac{\psi(x)}{(|t-x|+y)^{\lambda n}} y^{\lambda n} y^{-n} dx\right] dtdy,$$

so to prove (5.1.25), we must show that

$$\sup_{y>0} \int_{x\in\mathbb{R}^n} \frac{\psi(x)}{(|t-x|+y)^{\lambda n}} y^{\lambda n} y^{-n} dx \leqslant A_\lambda M\psi(t). \tag{5.1.26}$$

However, we know by Theorem 3.2.12, that

$$\sup_{\varepsilon>0} (\psi * \varphi_\varepsilon)(t) \leqslant AM\psi(t)$$

for appropriate φ, with $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$. Here, we have in fact $\varphi(x) = (1 + |x|)^{-\lambda n}$, $\varepsilon = y$, and so with $\lambda > 1$ the hypotheses of that theorem are satisfied. This proves (5.1.26) and thus also (5.1.25).

The case $p = 2$ follows immediately from (5.1.25) by inserting in this inequality the function $\psi = 1$ (or by the definitions of $g_\lambda^*(f)$ and $g(f)$ directly), and using the L^2 result for g.

Suppose now $p > 2$; let us set $1/q + 2/p = 1$, and take the supremum of the l.h.s. of (5.1.25) over all $\psi \geqslant 0$, such that $\psi \in L^q(\mathbb{R}^n)$ and $\|\psi\|_q \leqslant 1$. Then, it gives $\|g_\lambda^*(f)\|_p^2$; Hölder's inequality yields an estimate for the right side:

$$A_\lambda \|g(f)\|_p^2 \|M\psi\|_q.$$

However, by the inequalities for the g-function, $\|g(f)\|_p \leqslant A_p'\|f\|_p$; and by the theorem of the maximal function $\|M\psi\|_q \leqslant A_q\|\psi\|_q \leqslant A_q$, since $q > 1$, if $p < \infty$. If we substitute these in the above, we get the result:

$$\|g_\lambda^*(f)\|_p \leqslant A_{p,\lambda}\|f\|_p, \quad 2 \leqslant p < \infty, \quad \lambda > 1.$$

The inequalities for $p < 2$ will be proved by an adaptation of the reasoning used for g. Lemmas 5.1.5 and 5.1.6 will be equally applicable in the present situation, but we need more general version of Lemma 5.1.7, in order to majorize the unrestricted approach to the boundary of a Poisson integral.

It is at this stage where results which depend critically on the L^p class first make their appearance. Matters will depend on a variant of the maximal function which we define as follows. Let $\mu \geqslant 1$, and write $M_\mu f(x)$ for

$$M_\mu f(x) = \left(\sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|^\mu dy \right)^{1/\mu}. \tag{5.1.27}$$

Then $M_1 f(x) = Mf(x)$, and $M_\mu f(x) = ((M|f|^\mu)(x))^{1/\mu}$. From the theorem of the maximal function, it immediately follows that, for $p > \mu$,

$$\begin{aligned} \|M_\mu f\|_p =& \|((M|f|^\mu)(x))^{1/\mu}\|_p = \|((M|f|^\mu)(x))\|_{p/\mu}^{1/\mu} \\ \leqslant& C\||f|^\mu\|_{p/\mu}^{1/\mu} = C\|f\|_p. \end{aligned} \tag{5.1.28}$$

This inequality fails for $p \leqslant \mu$, as in the special case $\mu = 1$.

The substitute for Lemma 5.1.7 is as follows.

Lemma 5.1.12.

Let $f \in L^p(\mathbb{R}^n)$, $p \geqslant \mu \geqslant 1$ and $u(x,y)$ be the Poisson integral of f, then

$$|u(x-t,y)| \leqslant A \left(1 + \frac{|t|}{y}\right)^n Mf(x), \tag{5.1.29}$$

and more generally

$$|u(x-t,y)| \leqslant A_\mu \left(1 + \frac{|t|}{y}\right)^{n/\mu} M_\mu f(x). \tag{5.1.30}$$

We shall now complete the proof of the inequality (5.1.24) for the case $1 < p < 2$, with the restriction $p > 2/\lambda$.

Let us observe that we can always find a $\mu \in [1,p)$ such that if we set $\lambda' = \lambda - \frac{2-p}{\mu}$, then one still has $\lambda' > 1$. In fact, if $\mu = p$, then $\lambda - \frac{2-p}{\mu} > 1$ since $\lambda > 2/p$; this inequality can then be maintained by a small variation of μ. With this choice of μ, we have by Lemma 5.1.12

$$|u(x-t,y)| \left(\frac{y}{y+|t|}\right)^{n/\mu} \leqslant A_\mu M_\mu f(x). \tag{5.1.31}$$

We now proceed the argument with which we treated the function g.

$$\begin{aligned} &(g_\lambda^*(f)(x))^2 \\ =& \frac{1}{p(p-1)} \int_{\mathbb{R}_+^{n+1}} y^{1-n} \left(\frac{y}{y+|t|}\right)^{\lambda n} u^{2-p}(x-t,y)\Delta u^p(x-t,y)dtdy \\ \leqslant& \frac{1}{p(p-1)} A_\mu^{2-p}(M_\mu f(x))^{2-p} I^*(x), \end{aligned} \tag{5.1.32}$$

where

$$I^*(x) = \int_{\mathbb{R}_+^{n+1}} y^{1-n} \left(\frac{y}{y+|t|}\right)^{\lambda' n} \Delta u^p(x-t,y)dtdy.$$

It is clear that

$$\begin{aligned} \int_{\mathbb{R}^n} I^*(x)dx =& \int_{\mathbb{R}_+^{n+1}} \int_{\mathbb{R}_x^n} y^{1-n} \left(\frac{y}{y+|t-x|}\right)^{\lambda' n} \Delta u^p(t,y)dxdtdy \\ =& C_{\lambda'} \int_{\mathbb{R}_+^{n+1}} y\Delta u^p(t,y)dtdy, \end{aligned}$$

where the last step follows from the fact that if $\lambda' > 1$

$$y^{-n} \int_{\mathbb{R}^n} \left(\frac{y}{y+|t-x|}\right)^{\lambda' n} dx = y^{-n} \int_{\mathbb{R}^n} \left(\frac{y}{y+|x|}\right)^{\lambda' n} dx$$

$$\overset{x=yz}{=\!=}\int_{\mathbb{R}^n}\left(\frac{1}{1+|z|}\right)^{\lambda' n}dz$$
$$=C_{\lambda'}<\infty.$$

So, by Lemma 5.1.6, we have

$$\int_{\mathbb{R}^n} I^*(x)dx = C_{\lambda'}\int_{\mathbb{R}^n} u^p(t,0)dt = C_{\lambda'}\|f\|_p^p. \tag{5.1.33}$$

Therefore, by (5.1.32), Hölder's inequality, (5.1.28) and (5.1.33), we obtain

$$\|g_\lambda^*(f)\|_p \leqslant C\|M_\mu f(x)^{1-p/2}(I^*(x))^{1/2}\|_p \leqslant C\|M_\mu f\|_p^{1-p/2}\|I^*\|_1^{1/2} \leqslant C\|f\|_p,$$

which is the desired result. ■

Finally, we prove Lemma 5.1.12.

Proof of Lemma 5.1.12. One notices that (5.1.29) is unchanged by the dilation $(x,t,y)\to(\delta x,\delta t,\delta y)$, it is then clear that it suffices to prove (5.1.29) with $y=1$.

Setting $y=1$ in the Poisson kernel, we have $P_1(x)=c_n(1+|x|^2)^{-(n+1)/2}$, and $u(x-t,1)=(f*P_1)(x-t)$, for each t. Theorem 3.2.12 shows that $|u(x-t,1)|\leqslant A_tMf(x)$, where $A_t=\int Q_t(x)dx$, and $Q_t(x)$ is the smallest decreasing radial majorant of $P_1(x-t)$, i.e.,

$$Q_t(x)=c_n\sup_{|x'|\geqslant|x|}\frac{1}{(1+|x'-t|^2)^{(n+1)/2}}.$$

For $Q_t(x)$, we have the easy estimates, $Q_t(x)\leqslant c_n$ for $|x|\leqslant 2t$ and $Q_t(x)\leqslant A'(1+|x|^2)^{-(n+1)/2}$, for $|x|\geqslant 2|t|$, from which it is obvious that $A_t\leqslant A(1+|t|)^n$ and hence (5.1.29) is proved.

Since $u(x-t,y)=\int_{\mathbb{R}^n}P_y(s)f(x-t-s)ds$, and $\int_{\mathbb{R}^n}P_y(s)ds=1$, by Hölder inequality, we have

$$\begin{aligned}u(x-t,y)\leqslant&\|P_y^{1/\mu}f(x-t-\cdot)\|_\mu\|P_y^{1/\mu'}\|_{\mu'}\\ \leqslant&\left(\int_{\mathbb{R}^n}P_y(s)|f(x-t-s)|^\mu ds\right)^{1/\mu}=U^{1/\mu}(x-t,y),\end{aligned}$$

where U is the Poisson integral of $|f|^\mu$. Apply (5.1.29) to U, it gives

$$\begin{aligned}|u(x-t,y)|\leqslant&A^{1/\mu}(1+|t|/y)^{n/\mu}(M(|f|^\mu)(x))^{1/\mu}\\ =&A_\mu(1+|t|/y)^{n/\mu}M_\mu f(x),\end{aligned}$$

and the Lemma is established. ■

§ 5.2 Mikhlin and Hörmander multiplier theorem

The first application of the theory of the functions g and g_λ^* will be in the study of multipliers. Our main tool when proving theorems for the Sobolev spaces, defined in the following chapter, is the following theorem.

Theorem 5.2.1: Mikhlin multiplier theorem

Suppose that $m(\xi)\in\mathcal{C}^k(\mathbb{R}^n\setminus\{0\})$ where $k>n/2$ is an integer. Assume also that for every differential monomial $\left(\frac{\partial}{\partial\xi}\right)^\alpha$, $\alpha=(\alpha_1,\alpha_2,...,\alpha_n)$, with $|\alpha|=\alpha_1+\alpha_2+...+\alpha_n$, we have Mikhlin's condition

$$\left|\left(\frac{\partial}{\partial\xi}\right)^\alpha m(\xi)\right|\leqslant A|\xi|^{-|\alpha|},\quad\text{whenever } |\alpha|\leqslant k. \tag{5.2.1}$$

Then $m \in \mathcal{M}_p$, $1 < p < \infty$, and

$$\|m\|_{\mathcal{M}_p} \leqslant C_{p,n} A.$$

The proof of the theorem leads to a generalization of its statement which we formulate as a corollary.

Corollary 5.2.2: Hörmander multiplier theorem

The assumption (5.2.1) can be replaced by the weaker assumptions, i.e., Hörmander's condition

$$\begin{aligned} |m(\xi)| \leqslant & A, \\ \sup_{0<R<\infty} R^{2|\alpha|-n} \int_{R\leqslant|\xi|\leqslant 2R} \left| \left(\frac{\partial}{\partial\xi}\right)^{\alpha} m(\xi) \right|^2 d\xi \leqslant & A, \quad |\alpha| \leqslant k. \end{aligned} \tag{5.2.2}$$

The theorem and its corollary will be consequences of the following lemma. Its statement illuminates at the same time the nature of the multiplier transforms considered here, and the role played by the g-functions and their variants. We also remark here that while the Mikhlin and Hörmander multiplier theorem provides useful sufficient conditions for L^p boundedness of Fourier multipliers, it is not discerning enough for a satisfactory theory; it does not distinguish between p so long as $1 < p < \infty$. Note that $1 < p < \infty$ here in contrast to the case in Theorem 2.6.5.

Lemma 5.2.3.

Under the assumptions of Theorem 5.2.1 or Corollary 5.2.2, let us set for $f \in L^2(\mathbb{R}^n)$

$$F(x) = T_m f(x) = \left(\frac{|\omega|}{2\pi}\right)^{n/2} (\widecheck{m} * f)(x).$$

Then

$$g_1(F)(x) \leqslant A_\lambda g_\lambda^*(f)(x), \quad \text{where } \lambda = 2k/n. \tag{5.2.3}$$

Thus in view of the lemma, the g-functions and their variants are the characterizing expressions which deal at once with all the multipliers considered. On the other hand, the fact that the relation (5.2.3) is pointwise shows that to a large extent the mapping T_m is "semi-local".

Proof of Theorem 5.2.1 and Corollary 5.2.2. The conclusion is deduced from the lemma as follows. Our assumption on k is such that $\lambda = 2k/n > 1$. Thus, Theorem 5.1.11 shows us that

$$\|g_\lambda^*(f)\|_p \leqslant A_{\lambda,p} \|f\|_p, \quad 2 \leqslant p < \infty, \text{ if } f \in L^2 \cap L^p.$$

However, by Corollary 5.1.3, $A'_p \|F\|_p \leqslant \|g_1(F)\|_p$, therefore by Lemma 5.2.3,

$$\|T_m f\|_p = \|F\|_p \leqslant A_\lambda \|g_\lambda^*(f)\|_p \leqslant A_p \|f\|_p, \quad \text{if } 2 \leqslant p < \infty \text{ and } f \in L^2 \cap L^p.$$

That is, $m \in \mathcal{M}_p$, $2 \leqslant p < \infty$. By duality, i.e., (2.6.2), we have also $m \in \mathcal{M}_p$, $1 < p \leqslant 2$, which gives the assertion of the theorem. ■

Now we shall prove Lemma 5.2.3.

Proof of Lemma 5.2.3. Let $u(x, y)$ denote the Poisson integral of f, and $U(x, y)$ the Poisson integral of F. Then with $\hat{}$ denoting the Fourier transform w.r.t. the x variable, we have

$$\widehat{u}(\xi, y) = e^{-|\omega\xi|y}\widehat{f}(\xi), \text{ and } \widehat{U}(\xi, y) = e^{-|\omega\xi|y}\widehat{F}(\xi) = e^{-|\omega\xi|y}m(\xi)\widehat{f}(\xi).$$

Define $M(x, y) = \left(\frac{|\omega|}{2\pi}\right)^n \int_{\mathbb{R}^n} e^{\omega ix\cdot\xi}e^{-|\omega\xi|y}m(\xi)d\xi$. It is clear that

$$\left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{M}(\xi, y) = e^{-|\omega\xi|y}m(\xi),$$

and then

$$\widehat{U}(\xi, y_1 + y_2) = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \widehat{M}(\xi, y_1)\widehat{u}(\xi, y_2), \quad y = y_1 + y_2, \quad y_1, y_2 > 0.$$

This can be written as

$$U(x, y_1 + y_2) = \int_{\mathbb{R}^n} M(t, y_1)u(x - t, y_2)dt.$$

We differentiate this relation k times w.r.t. y_1 and once w.r.t. y_2, and set $y_1 = y_2 = y/2$. This gives us the identity

$$U^{(k+1)}(x, y) = \int_{\mathbb{R}^n} M^{(k)}(t, y/2)u^{(1)}(x - t, y/2)dt. \tag{5.2.4}$$

Here the superscripts denote the differentiation w.r.t. y.

Next, we translate the assumptions (5.2.1) (or (5.2.2)) on m in terms of $M(x, y)$. The result is

$$|M^{(k)}(t, y)| \leqslant A'y^{-n-k}, \tag{5.2.5}$$

$$\int_{\mathbb{R}^n} |t|^{2k}|M^{(k)}(t, y)|^2 dt \leqslant A'y^{-n}, \tag{5.2.6}$$

where A' depends only on n and k.

In fact, from the definition of M and the condition $|m(\xi)| \leqslant A$, it follows that

$$\begin{aligned}
|M^{(k)}(x, y)| \leqslant & \left(\frac{|\omega|}{2\pi}\right)^n |\omega|^k \int_{\mathbb{R}^n} |\xi|^k e^{-|\omega\xi|y}|m(\xi)|d\xi \\
\leqslant & A\omega_{n-1}\left(\frac{|\omega|}{2\pi}\right)^n |\omega|^k \int_0^\infty r^k e^{-|\omega|ry}r^{n-1}dr \\
= & A\omega_{n-1}\left(\frac{1}{2\pi}\right)^n y^{-n-k}\int_0^\infty e^{-R}R^{k+n-1}dR \\
= & A\omega_{n-1}\left(\frac{1}{2\pi}\right)^n \Gamma(k+n)y^{-n-k},
\end{aligned}$$

which is (5.2.5).

To prove (5.2.6), let us show more particularly that

$$\int_{\mathbb{R}^n} |x^\alpha M^{(k)}(x, y)|^2 dx \leqslant A'y^{-n},$$

where $|\alpha| = k$.

By Plancherel's theorem and Proposition 2.1.2, we have

$$\|x^\alpha M^{(k)}(x, y)\|_2 = \left(\frac{|\omega|}{2\pi}\right)^{n/2} \left\|\left(\frac{\partial}{\partial\xi}\right)^\alpha (|\xi|^k m(\xi)e^{-|\omega\xi|y})\right\|_2. \tag{5.2.7}$$

Then we need to evaluate, by using Leibniz' rule, that

$$\left(\frac{\partial}{\partial\xi}\right)^\alpha (|\xi|^k m(\xi)e^{-|\omega\xi|y}) = \sum_{\beta+\gamma=\alpha} C_{\beta,\gamma}\left(\frac{\partial}{\partial\xi}\right)^\beta (|\xi|^k m(\xi))\left(\frac{\partial}{\partial\xi}\right)^\gamma e^{-|\omega\xi|y}. \tag{5.2.8}$$

Case I: (5.2.1) $\Longrightarrow$ (5.2.6)**.** By the hypothesis (5.2.1) and Leibniz' rule again, we have

$$\left|\left(\frac{\partial}{\partial\xi}\right)^\beta (|\xi|^k m(\xi))\right| \leqslant A'|\xi|^{k-|\beta|}, \text{ with } |\beta|\leqslant k.$$

Thus,

$$\begin{aligned}&\left|\left(\frac{\partial}{\partial\xi}\right)^\alpha (|\xi|^k m(\xi)e^{-|\omega\xi|y})\right|\\ \leqslant & C\sum_{|\beta|+|\gamma|=k}|\xi|^{k-|\beta|}(|\omega|y)^{|\gamma|}e^{-|\omega\xi|y}\leqslant C\sum_{0\leqslant r\leqslant k}|\xi|^r(|\omega|y)^r e^{-|\omega\xi|y}.\end{aligned}$$

Since for $r\geqslant 0$

$$\begin{aligned}(|\omega|y)^{2r}\int_{\mathbb{R}^n}|\xi|^{2r}e^{-2|\omega\xi|y}d\xi =&\omega_{n-1}(|\omega|y)^{2r}\int_0^\infty R^{2r}e^{-2|\omega|Ry}R^{n-1}dR\\ =&\omega_{n-1}2^{-(2r+n)}(|\omega|y)^{-n}\int_0^\infty z^{2r+n-1}e^{-z}dz\\ =&\omega_{n-1}(|\omega|y)^{-n}2^{-(2r+n)}\Gamma(2r+n),\end{aligned}$$

we get for $|\alpha|=k$

$$\begin{aligned}\|x^\alpha M^{(k)}(x,y)\|_2^2\leqslant&\left(\frac{|\omega|}{2\pi}\right)^n\omega_{n-1}(|\omega|y)^{-n}\left[\sum_{0\leqslant r\leqslant k}(2^{-(2r+n)}\Gamma(2r+n))^{1/2}\right]^2\\ \leqslant& C_{k,n}y^{-n},\end{aligned}$$

which proves the assertion (5.2.6) in view of (2.3.6).

Case II: (5.2.2) $\Longrightarrow$ (5.2.6)**.** From (5.2.7) and (5.2.8), we have, by Leibniz' rule again and (5.2.2),

$$\begin{aligned}&\left(\frac{|\omega|}{2\pi}\right)^{-n/2}\|x^\alpha M^{(k)}(x,y)\|_2\\ \leqslant&\sum_{|\beta'|+|\beta''|+|\gamma|=k}C_{\beta'\beta''\gamma}\\ &\left(\int_{\mathbb{R}^n}\left|\left(\frac{\partial}{\partial\xi}\right)^{\beta'}|\xi|^k\right|^2\left|\left(\frac{\partial}{\partial\xi}\right)^{\beta''}m(\xi)\right|^2 e^{-2|\omega\xi|y}(|\omega|y)^{2|\gamma|}d\xi\right)^{1/2}\\ \leqslant& C\sum_{|\beta'|+|\beta''|+|\gamma|=k}(|\omega|y)^{|\gamma|}\left(\sum_{j\in\mathbb{Z}}\int_{2^j<|\xi|\leqslant 2^{j+1}}|\xi|^{2(k-|\beta'|)}\left|\left(\frac{\partial}{\partial\xi}\right)^{\beta''}m(\xi)\right|^2 e^{-2|\omega\xi|y}d\xi\right)^{1/2}\\ \leqslant& C\sum_{|\beta'|+|\beta''|+|\gamma|=k}(|\omega|y)^{|\gamma|}\left[\sum_{j\in\mathbb{Z}}(2^{j+1})^{2(k-|\beta'|)}e^{-|\omega|2^{j+1}y}\right.\\ &\left.\cdot(2^j)^{-2|\beta''|+n}\left((2^j)^{2|\beta''|-n}\int_{2^j<|\xi|\leqslant 2^{j+1}}\left|\left(\frac{\partial}{\partial\xi}\right)^{\beta''}m(\xi)\right|^2 d\xi\right)\right]^{1/2}\\ \leqslant& CA^{1/2}2^k\sum_{0\leqslant r\leqslant k}(|\omega|y)^r\left(\sum_{j\in\mathbb{Z}}2^j(2^j)^{2r+n-1}e^{-|\omega|2^{j+1}y}\right)^{1/2}\\ \leqslant& CA^{1/2}2^k\sum_{0\leqslant r\leqslant k}(|\omega|y)^r\left(\int_0^\infty R^{2r+n-1}e^{-|\omega|Ry}dR\right)^{1/2}\end{aligned}$$

$$
\begin{aligned}
=&CA^{1/2}2^k \sum_{0\leqslant r\leqslant k} (|\omega|y)^{-n/2} \left(\int_0^\infty z^{2r+n-1}e^{-z}dz\right)^{1/2}\\
\leqslant&CA^{1/2}(|\omega|y)^{-n/2}2^k \sum_{0\leqslant r\leqslant k} \Gamma^{1/2}(2r+n),
\end{aligned}
$$

which yields (5.2.6) in view of (2.3.6) again.

Now, we return to the identity (5.2.4), and for each y divide the range of integration into two parts, $|t| \leqslant y/2$ and $|t| > y/2$. In the first range, we use the estimate (5.2.5) on $M^{(k)}$ and in the second range, we use the estimate (5.2.6). This together with Schwarz' inequality gives immediately

$$
\begin{aligned}
|U^{(k+1)}(x,y)|^2 \leqslant&Cy^{-n-2k}\int_{|t|\leqslant y/2} |u^{(1)}(x-t,y/2)|^2dt\\
&+Cy^{-n}\int_{|t|>y/2} \frac{|u^{(1)}(x-t,y/2)|^2dt}{|t|^{2k}}\\
=:&I_1(y)+I_2(y).
\end{aligned}
$$

Now

$$
(g_{k+1}(F)(x))^2 = \int_0^\infty |U^{(k+1)}(x,y)|^2 y^{2k+1}dy \leqslant \sum_{j=1}^2 \int_0^\infty I_j(y)y^{2k+1}dy.
$$

However, by a change of variable $y/2 \to y$,

$$
\begin{aligned}
\int_0^\infty I_1(y)y^{2k+1}dy =&C\int_0^\infty \int_{|t|\leqslant y/2} |u^{(1)}(x-t,y/2)|^2 y^{-n+1}dtdy\\
\leqslant&C\int_\Gamma |\nabla u(x-t,y)|^2 y^{-n+1}dtdy = C(S(f)(x))^2\\
\leqslant&C_\lambda(g_\lambda^*(f)(x))^2.
\end{aligned}
$$

Similarly, with $n\lambda = 2k$,

$$
\begin{aligned}
\int_0^\infty I_2(y)y^{2k+1}dy \leqslant&C\int_0^\infty \int_{|t|>y} y^{-n+2k+1}|t|^{-2k}|\nabla u(x-t,y)|^2dtdy\\
\leqslant&C(g_\lambda^*(f)(x))^2.
\end{aligned}
$$

This shows that $g_{k+1}(F)(x) \leqslant C_\lambda g_\lambda^*(f)(x)$. However, by Remark 5.1.4 (iii) of g-functions, we know that $g_1(F)(x) \leqslant C_k g_{k+1}(F)(x)$. Thus, the proof of the lemma is concluded. ■

§ 5.3 The partial sums operators

We shall now develop the second main tool in the Littlewood-Paley theory, (the first being the usage of the functions g and g^*).

Let ρ denote an arbitrary rectangle in $\mathbb{R}^n$. By rectangle we shall mean, in the rest of this chapter, a possibly infinite rectangle with sides parallel to the axes, i.e., the Cartesian product of n intervals.

Definition 5.3.1.

For each rectangle ρ denote by S_ρ the ***partial sum operator***, that is the multiplier operator with $m = \chi_\rho$, i.e., characteristic function of the rectangle ρ.

So

$$\widehat{S_\rho f} = \chi_\rho \widehat{f}, \quad f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n). \tag{5.3.1}$$

For this operator, we immediately have the following theorem in view of the Mikhlin multiplier theorem.

Theorem 5.3.2.

Let $1 < p < \infty$, we have

$$\|S_\rho f\|_p \leqslant A_p \|f\|_p, \quad f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n),$$

where the constant A_p does not depend on the rectangle ρ.

However, we shall need a more extended version of the theorem which arises when we replace complex-valued functions by functions taking their value in a Hilbert space.

Let $\mathcal{H}$ be the sequence Hilbert space,

$$\mathcal{H} = \{(c_j)_{j=1}^\infty : (\sum_j |c_j|^2)^{1/2} = |c| < \infty\}.$$

Then we can represent a function $f \in L^p(\mathbb{R}^n, \mathcal{H})$ as sequences

$$f(x) = (f_1(x), \cdots, f_j(x), \cdots),$$

where each f_j is complex-valued and $|f(x)| = (\sum_{j=1}^\infty |f_j(x)|^2)^{1/2}$. Let $\mathfrak{R}$ be a sequence of rectangle, $\mathfrak{R} = \{\rho_j\}_{j=1}^\infty$. Then we can define the operator $S_\mathfrak{R}$, mapping $L^2(\mathbb{R}^n, \mathcal{H})$ to itself, by the rule

$$S_\mathfrak{R} f = (S_{\rho_1} f_1, \cdots, S_{\rho_j} f_j, \cdots), \text{ where } f = (f_1, \cdots, f_j, \cdots). \tag{5.3.2}$$

We first give a lemma, which will be used in the proof of the theorem or its generalization. Recall the Hilbert transform $f \to H(f)$, which corresponds to the multiplier $-i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi)$ in one dimension.

Lemma 5.3.3.

Let $f(x) = (f_1(x), \cdots, f_j(x), \cdots) \in L^2(\mathbb{R}^n, \mathcal{H}) \cap L^p(\mathbb{R}^n, \mathcal{H})$. Denote $\widetilde{H} f(x) = (Hf_1(x), \cdots, Hf_j(x), \cdots)$. Then

$$\|\widetilde{H} f\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty,$$

where A_p is the same constant as in the scalar case, i.e., when $\mathcal{H}$ is one-dimensional.

Proof. We use the vector-valued version of the Hilbert transform, as is described more generally in Sec. 4.7. Let the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ be both identical with $\mathcal{H}$. Take in $\mathbb{R}$, $K(x) = I \cdot 1/\pi x$, where I is the identity mapping on $\mathcal{H}$. Then the kernel $K(x)$ satisfies all the assumptions of Theorem 4.7.1 and Theorem 4.6.1. Moreover,

$$\lim_{\varepsilon \to 0} \int_{|y| \geqslant \varepsilon} K(y) f(x-y) dy = \widetilde{H} f(x),$$

and thus the lemma is proved. ■

The generalization of Theorem 5.3.2 is then as follows.

Theorem 5.3.4.

Let $f \in L^2(\mathbb{R}^n, \mathcal{H}) \cap L^p(\mathbb{R}^n, \mathcal{H})$. Then

$$\|S_{\Re} f\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{5.3.3}$$

where A_p does not depend on the family $\Re$ of rectangles.

Proof. The theorem will be proved in four steps, the first two of which already contain the essence of the matter.

Step 1: $n = 1$, and the rectangles $\rho_1, \rho_2, \cdots, \rho_j, \cdots$ are the semi-infinite intervals $(-\infty, 0)$.

It is clear that $\widehat{S_{(-\infty,0)} f} = \chi_{(-\infty,0)} \widehat{f} = \frac{1-\operatorname{sgn}(\xi)}{2} \widehat{f}$, so

$$S_{(-\infty,0)} = \frac{I - i \operatorname{sgn}(\omega) H}{2}, \tag{5.3.4}$$

where I is the identity, and $S_{(-\infty,0)}$ is the partial sum operator corresponding to the interval $(-\infty, 0)$.

Now if all the rectangles are the intervals $(-\infty, 0)$, then by (5.3.4),

$$S_{\Re} = \frac{I - i \operatorname{sgn}(\omega) \widetilde{H}}{2}$$

and so by Lemma 5.3.3, we have the desired result.

Step 2: $n = 1$, and the rectangles are the intervals $(-\infty, a_1)$, $(-\infty, a_2)$, $\cdots$, $(-\infty, a_j)$, $\cdots$.

Notice that $\widehat{f(x) e^{-\omega i x \cdot a}} = \widehat{f}(\xi + a)$, therefore

$$\widehat{H(e^{-\omega i x \cdot a} f(x))} = -i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi) \widehat{f}(\xi + a),$$

and hence $\widehat{e^{\omega i x \cdot a} H(e^{-\omega i x \cdot a} f(x))} = -i \operatorname{sgn}(\omega) \operatorname{sgn}(\xi - a) \widehat{f}(\xi)$. From this, we see that

$$(S_{(-\infty, a_j)} f_j)(x) = \frac{f_j - i \operatorname{sgn}(\omega) e^{\omega i x \cdot a_j} H(e^{-\omega i x \cdot a_j} f_j)}{2}. \tag{5.3.5}$$

If we now write symbolically $e^{-\omega i x \cdot a} f$ for

$$(e^{-\omega i x \cdot a_1} f_1, \cdots, e^{-\omega i x \cdot a_j} f_j, \cdots)$$

with $f = (f_1, \cdots, f_j, \cdots)$, then (5.3.5) may be written as

$$S_{\Re} f = \frac{f - i \operatorname{sgn}(\omega) e^{\omega i x \cdot a} \widetilde{H}(e^{-\omega i x \cdot a} f)}{2}, \tag{5.3.6}$$

and so the result again follows in this case by Lemma 5.3.3.

Step 3: General n, but the rectangles ρ_j are the half-spaces $x_1 < a_j$, i.e., $\rho_j = \{x : x_1 < a_j\}$.

Let $S^{(1)}_{(-\infty, a_j)}$ denote the operator defined on $L^2(\mathbb{R}^n)$, which acts only on the x_1 variable, by the action given by $S_{(-\infty, a_j)}$. We claim that

$$S_{\rho_j} = S^{(1)}_{(-\infty, a_j)}. \tag{5.3.7}$$

This identity is obvious for L^2 functions of the product form

$$f'(x_1) f''(x_2, \cdots, x_n),$$

since their linear span is dense in L^2, the identity (5.3.7) is established.

We now use the L^p inequality, which is the result of the previous step for each fixed x_2, x_3, $\cdots$, x_n. We raise this inequality to the p^{th} power and integrate w.r.t. $x_2, \cdots, x_n$. This gives the desired result for the present case. Notice that the result holds as well if the half-space $\{x : x_1 < a_j\}_{j=1}^{\infty}$, is replaced by the half-space $\{x : x_1 > a_j\}_{j=1}^{\infty}$, or if the role of the x_1 axis is taken by the x_2 axis, etc.

Step 4: Observe that every general finite rectangle of the type considered is the intersection of $2n$ half-spaces, each half-space having its boundary hyperplane perpendicular to one of the axes of $\mathbb{R}^n$. Thus a $2n$-fold application of the result of the third step proves the theorem, where the family $\mathfrak{R}$ is made up of finite rectangles. Since the bounds obtained do not depend on the family $\mathfrak{R}$, we can pass to the general case where $\mathfrak{R}$ contains possibly infinite rectangles by an obvious limiting argument. ■

We state here the continuous analogue of Theorem 5.3.4. Let $(\Gamma, d\gamma)$ be a σ-finite measure space, and consider the Hilbert space $\mathcal{H}$ of square integrable functions on Γ, i.e., $\mathcal{H} = L^2(\Gamma, d\gamma)$. The elements

$$f \in L^p(\mathbb{R}^n, \mathcal{H})$$

are the complex-valued functions $f(x,\gamma) = f_\gamma(x)$ on $\mathbb{R}^n \times \Gamma$, which are jointly measurable, and for which $(\int_{\mathbb{R}^n}(\int_\Gamma |f(x,\gamma)|^2 d\gamma)^{p/2} dx)^{1/p} = \|f\|_p < \infty$, if $p < \infty$. Let $\mathfrak{R} = \{\rho_\gamma\}_{\gamma\in\Gamma}$, and suppose that the mapping $\gamma \to \rho_\gamma$ is a measurable function from Γ to rectangles; that is, the numerical-valued functions which assign to each γ the components of the vertices of ρ_γ are all measurable.

Suppose $f \in L^2(\mathbb{R}^n, \mathcal{H})$. Then we define $F = S_{\mathfrak{R}} f$ by the rule

$$F(x,\gamma) = S_{\rho_\gamma}(f_\gamma)(x), \quad (f_\gamma(x) = f(x,\gamma)).$$

Theorem 5.3.5.

It holds

$$\|S_{\mathfrak{R}} f\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{5.3.8}$$

for $f \in L^2(\mathbb{R}^n, \mathcal{H}) \cap L^p(\mathbb{R}^n, \mathcal{H})$, where the bound A_p does not depend on the measure space $(\Gamma, d\gamma)$, or on the function $\gamma \to \rho_\gamma$.

Proof. The proof of this theorem is an exact repetition of the argument given for Theorem 5.3.4. The reader may also obtain it from Theorem 5.3.4 by a limiting argument. ■

§ 5.4 The dyadic decomposition

We shall now consider a decomposition of $\mathbb{R}^n$ into rectangles.

First, in the case of $\mathbb{R}$, we decompose it as the union of the almost disjoint intervals (i.e., whose interiors are disjoint) $[2^k, 2^{k+1}]$, $-\infty < k < \infty$, and $[-2^{k+1}, -2^k]$, $-\infty < k < \infty$. This double collection of intervals, one collection for the positive half-line, the other for the negative half-line, will be the dyadic decomposition of $\mathbb{R}$.

Having obtained this decomposition of $\mathbb{R}$, we take the corresponding product decomposition for

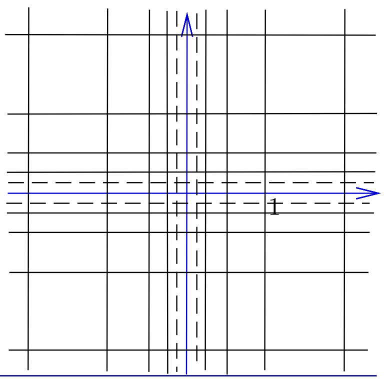


Figure 5.3: The dyadic decomposition

$\mathbb{R}^n$. Thus we write $\mathbb{R}^n$ as the union of almost disjoint rectangles which are products of the intervals occurring for the dyadic decomposition of each of the axes. This is the ***dyadic decomposition of*** $\mathbb{R}^n$.

The family of resulting rectangles will be denoted by Δ. We recall the partial sum operator S_ρ, defined in (5.3.1) for each rectangle. Now in an obvious sense, (e.g. L^2 convergence)

$$\sum_{\rho\in\Delta} S_\rho = \text{Identity}.$$

Also in the L^2 case, the different blocks, $S_\rho f$, $\rho \in \Delta$, behave as if they were independent; they are of course mutually orthogonal. To put the matter precisely: The L^2 norm of f can be given exactly in terms of the L^2 norms of $S_\rho f$, i.e.,

$$\sum_{\rho\in\Delta} \|S_\rho f\|_2^2 = \|f\|_2^2, \tag{5.4.1}$$

(and this is true for any decomposition of $\mathbb{R}^n$). For the general L^p case not as much can be hoped for, but the following important theorem can nevertheless be established.

Theorem 5.4.1: Littlewood-Paley square function theorem

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

$$\|(\sum_{\rho\in\Delta} |S_\rho f(x)|^2)^{1/2}\|_p \sim \|f\|_p.$$

The Rademacher functions provide a very useful device in the study of L^p norms in terms of quadratic expressions.

These functions, $r_0(t)$, $r_1(t)$, $\cdots$, $r_m(t)$, $\cdots$ are defined on the interval $(0,1)$ as follows:

$$r_0(t) = \begin{cases} 1, & 0 \leqslant t \leqslant 1/2, \\ -1, & 1/2 < t < 1, \end{cases}$$

r_0 is extended outside the unit interval by periodicity, i.e., $r_0(t+1) = r_0(t)$. In general, $r_m(t) = r_0(2^m t)$. The sequences of Rademacher functions are orthonormal (and in fact mutually independent) over $[0,1]$. In fact, for $m < k$, we have

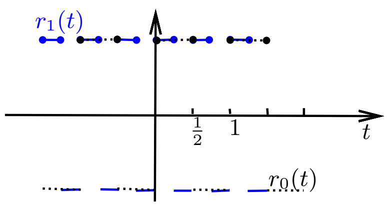


Figure 5.4: $r_0(t)$ with dotted line and $r_1(t)$ with solid line

$$\begin{aligned}
&\int_0^1 r_m(t) r_k(t)dt = \int_0^1 r_0(2^m t) r_0(2^k t)dt \\
=&2^{-m}\int_0^{2^m} r_0(s) r_0(2^{k-m}s)ds = \int_0^1 r_0(s) r_0(2^{k-m}s)ds \\
=&\int_0^{1/2} r_0(2^{k-m}s)ds - \int_{1/2}^1 r_0(2^{k-m}s)ds \\
=&2^{m-k}\left[\int_0^{2^{k-m-1}} r_0(t)dt - \int_{2^{k-m-1}}^{2^{k-m}} r_0(t)dt\right] \\
=&2^{-1}\left[\int_0^1 r_0(t)dt - \int_0^1 r_0(t)dt\right] = 0,
\end{aligned}$$

thus, they are orthogonal. It is clear that they are normal since $\int_0^1 (r_m(t))^2 dt = 1$.

For our purposes, their importance arises from the following fact.

Suppose $\sum_{m=0}^{\infty}|a_m|^2<\infty$ and set $F(t)=\sum_{m=0}^{\infty}a_mr_m(t)$. Then for every $1<p<\infty$, we can prove $F(t)\in L^p[0,1]$ and

$$A_p\|F\|_p\leqslant\|F\|_2=(\sum_{m=0}^{\infty}|a_m|^2)^{1/2}\leqslant B_p\|F\|_p, \tag{5.4.2}$$

for two positive constants A_p and B_p.

Thus, for functions which can be expanded in terms of the Rademacher functions, all the L^p norms, $1<p<\infty$, are comparable.

We shall also need the n-dimensional form of (5.4.2). We consider the unit cube $Q\subset\mathbb{R}^n$, $Q=\{t=(t_1,t_2,\cdots,t_n):0\leqslant t_j\leqslant 1\}$. Let m be an n-tuple of non-negative integers $m=(m_1,m_2,\cdots,m_n)$. Define $r_m(t)=r_{m_1}(t_1)r_{m_2}(t_2)\cdots r_{m_n}(t_n)$. Write $F(t)=\sum a_mr_m(t)$. With

$$\|F\|_p=\left(\int_Q|F(t)|^pdt\right)^{1/p},$$

we can show (5.4.2), whenever $\sum|a_m|^2<\infty$. We state these results as follows.

Lemma 5.4.2.

Let $F(t)=\sum a_mr_m(t)$ for $t\in\mathbb{R}^n$ and $m\in\mathbb{N}_0^n$. Suppose $\sum|a_m|^2<\infty$. Then it holds

$$\|F\|_p\sim\|F\|_2=\left(\sum_{|m|=0}^{\infty}|a_m|^2\right)^{1/2},\quad 1<p<\infty. \tag{5.4.3}$$

Proof. We split the proof into four steps.

Step 1: Let $n=1$ and μ, a_0, a_1, $\cdots$, a_N be real numbers. Then because the Rademacher functions are mutually independent, we have, in view of their definition,

$$\int_0^1 e^{\mu a_mr_m(t)}dt=\int_0^1 e^{\mu a_mr_0(2^mt)}dt=2^{-m}\int_0^{2^m}e^{\mu a_mr_0(s)}ds=\int_0^1e^{\mu a_mr_0(s)}ds$$
$$=2^{-1}(e^{\mu a_m}+e^{-\mu a_m})=\cosh\mu a_m.$$

and for $m<k$

$$\begin{aligned}&\int_0^1 e^{\mu a_mr_m(t)}e^{\mu a_kr_k(t)}dt=\int_0^1 e^{\mu a_mr_0(2^mt)}e^{\mu a_kr_0(2^kt)}dt\\
=&2^{-m}\int_0^{2^m}e^{\mu a_mr_0(s)}e^{\mu a_kr_0(2^{k-m}s)}ds=\int_0^1e^{\mu a_mr_0(s)}e^{\mu a_kr_0(2^{k-m}s)}ds\\
=&\int_0^{1/2}e^{\mu a_m}e^{\mu a_kr_0(2^{k-m}s)}ds+\int_{1/2}^1e^{-\mu a_m}e^{\mu a_kr_0(2^{k-m}s)}ds\\
=&2^{m-k}\left[\int_0^{2^{k-m-1}}e^{\mu a_m}e^{\mu a_kr_0(t)}dt+\int_{2^{k-m-1}}^{2^{k-m}}e^{-\mu a_m}e^{\mu a_kr_0(t)}dt\right]\\
=&2^{-1}(e^{\mu a_m}+e^{-\mu a_m})\int_0^1e^{\mu a_kr_0(t)}dt=\int_0^1e^{\mu a_mr_m(t)}dt\int_0^1e^{\mu a_kr_k(t)}dt.\end{aligned}$$

Thus, by induction, we can verify

$$\int_0^1e^{\mu\sum_{m=0}^Na_mr_m(t)}dt=\prod_{m=0}^N\int_0^1e^{\mu a_mr_m(t)}dt.$$

If we now make use of this simple inequality $\cosh x \leqslant e^{x^2}$ (since $\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \leqslant \sum_{k=0}^{\infty} \frac{x^{2k}}{k!} = e^{x^2}$ for $|x| < \infty$ by Taylor expansion), we obtain

$$\int_0^1 e^{\mu F(t)} dt = \prod_{m=0}^{N} \cosh \mu a_m \leqslant \prod_{m=0}^{N} e^{\mu^2 a_m^2} = e^{\mu^2 \sum_{m=0}^{N} a_m^2},$$

with $F(t) = \sum_{m=0}^{N} a_m r_m(t)$.

Step 2: $n = 1$. Let us make the normalizing assumption that $\sum_{m=0}^{N} a_m^2 = 1$. Then, since $e^{\mu|F|} \leqslant e^{\mu F} + e^{-\mu F}$, we have

$$\int_0^1 e^{\mu|F(t)|} dt \leqslant 2e^{\mu^2}.$$

Recall the distribution function $F_*(\alpha) = |\{t \in [0,1] : |F(t)| > \alpha\}|$. If we take $\mu = \alpha/2$ in the above inequality, we have

$$F_*(\alpha) = \int_{\{|F(t)|>\alpha\}\cap[0,1]} dt \leqslant e^{-\frac{\alpha^2}{2}} \int_{\{|F(t)|>\alpha\}\cap[0,1]} e^{\frac{\alpha}{2}|F(t)|} dt \leqslant e^{-\frac{\alpha^2}{2}} 2e^{\frac{\alpha^2}{4}} = 2e^{-\frac{\alpha^2}{4}}.$$

From Theorem 1.1.4, the above and changes of variables, it follows immediately that

$$\begin{aligned}
\|F\|_p = &\left(p \int_0^\infty \alpha^{p-1} F_*(\alpha) d\alpha\right)^{1/p} \leqslant \left(2p \int_0^\infty \alpha^{p-1} e^{-\frac{\alpha^2}{4}} d\alpha\right)^{1/p} \\
= &\left(2^p p \int_0^\infty s^{p/2-1} e^{-s} ds\right)^{1/p} \quad (\text{set } s = \alpha^2/4) \\
= &2(p\Gamma(p/2))^{1/p},
\end{aligned}$$

for $1 \leqslant p < \infty$, and so in general, we obtain

$$\|F\|_p \leqslant C_p \left(\sum_{m=0}^{\infty} |a_m|^2\right)^{1/2}, \quad 1 \leqslant p < \infty. \tag{5.4.4}$$

Step 3: We shall now extend the last inequality to several variables. The case of two variables is entirely of the inductive procedure used in the proof of the general case.

We can also limit ourselves to the situation when $p \geqslant 2$, since for the case $p < 2$ the desired inequality is a simple consequence of Hölder's inequality. (Indeed, for $p < 2$ and some $q \geqslant 2$, we have

$$\|F\|_{L^p(0,1)} \leqslant \|F\|_{L^q(0,1)} \|1\|_{L^{qp/(q-p)}(0,1)} \leqslant \|F\|_{L^q(0,1)}$$

by Hölder's inequality.)

We have

$$F(t_1, t_2) = \sum_{m_1=0}^{N} \sum_{m_2=0}^{N} a_{m_1 m_2} r_{m_1}(t_1) r_{m_2}(t_2) = \sum_{m_1=0}^{N} F_{m_1}(t_2) r_{m_1}(t_1),$$

where $F_{m_1}(t_2) = \sum_{m_2} a_{m_1 m_2} r_{m_2}(t_2)$. By(5.4.4), it follows

$$\int_0^1 |F(t_1, t_2)|^p dt_1 \leqslant C_p \left(\sum_{m_1} |F_{m_1}(t_2)|^2\right)^{p/2}.$$

Integrating this w.r.t. t_2, and using Minkowski's inequality with $p/2 \geqslant 1$, we have

$$\int_0^1 \left(\sum_{m_1} |F_{m_1}(t_2)|^2\right)^{p/2} dt_2 = \left\|\sum_{m_1} |F_{m_1}(t_2)|^2\right\|_{p/2}^{p/2} \leqslant \left(\sum_{m_1} \||F_{m_1}(t_2)|^2\|_{p/2}\right)^{p/2}$$

$$= \left(\sum_{m_1} \|F_{m_1}(t_2)\|_p^2 \right)^{p/2}.$$

However, $F_{m_1}(t_2) = \sum_{m_2} a_{m_1 m_2} r_{m_2}(t_2)$, and therefore the case already proved shows that

$$\|F_{m_1}(t_2)\|_p^2 \leqslant C_p \sum_{m_2} a_{m_1 m_2}^2.$$

Inserting this in the above gives

$$\int_0^1 \int_0^1 |F(t_1, t_2)|^p dt_1 dt_2 \leqslant C_p \left(\sum_{m_1} \sum_{m_2} a_{m_1 m_2}^2 \right)^{p/2},$$

which leads to the desired inequality

$$\|F\|_p \leqslant C_p \|F\|_2, \quad 2 \leqslant p < \infty.$$

Step 4: The converse inequality

$$\|F\|_2 \leqslant C_p \|F\|_p, \quad p > 1$$

is a simple consequence of the direct inequality.

In fact, for any $p > 1$, (here we may assume $p < 2$) by Hölder inequality

$$\|F\|_2 \leqslant \|F\|_p^{1/2} \|F\|_{p'}^{1/2}.$$

We already know that $\|F\|_{p'} \leqslant A'_{p'} \|F\|_2$, $p' > 2$. We therefore get

$$\|F\|_2 \leqslant C_{p'} \|F\|_p,$$

which is the required converse inequality. ■

Now, let us return to the proof of the Littlewood-Paley square function theorem.

Proof of Theorem 5.4.1. It will be presented in five steps.

Step 1: We show here that it suffices to prove the inequality

$$\left\| \left(\sum_{\rho \in \Delta} |S_\rho f(x)|^2 \right)^{1/2} \right\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{5.4.5}$$

for $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. To see this sufficiency, let $g \in L^2(\mathbb{R}^n) \cap L^{p'}(\mathbb{R}^n)$, and consider the identity

$$\sum_{\rho \in \Delta} \int_{\mathbb{R}^n} S_\rho f \overline{S_\rho g} dx = \int_{\mathbb{R}^n} f \bar{g} dx$$

which follows from (5.4.1) by polarization. By Schwarz's inequality and then Hölder's inequality, we get

$$\begin{aligned} \left| \int_{\mathbb{R}^n} f \bar{g} dx \right| \leqslant & \int_{\mathbb{R}^n} \left(\sum_\rho |S_\rho f|^2 \right)^{\frac{1}{2}} \left(\sum_\rho |S_\rho g|^2 \right)^{\frac{1}{2}} dx \\ \leqslant & \left\| \left(\sum_\rho |S_\rho f|^2 \right)^{\frac{1}{2}} \right\|_p \left\| \left(\sum_\rho |S_\rho g|^2 \right)^{\frac{1}{2}} \right\|_{p'}. \end{aligned}$$

Taking the supremum over all such g with the additional restriction that $\|g\|_{p'} \leqslant 1$, it gives $\|f\|_p$ for the l.h.s. of the above inequality. The r.h.s. is majorized by

$$A_{p'} \left\| \left(\sum |S_\rho f|^2 \right)^{1/2} \right\|_p,$$

since we assume (5.4.5) for all p. Thus, we have also

$$B_p\|f\|_p \leqslant \left\| \left(\sum_\rho |S_\rho f|^2 \right)^{1/2} \right\|_p . \tag{5.4.6}$$

To dispose of the additional assumption that $f \in L^2$, for $f \in L^p$ we take $f_j \in L^2 \cap L^p$ such that $\|f_j - f\|_p \to 0$ by density and use the inequality (5.4.5) and (5.4.6) for f_j and $f_j - f_{j'}$; after a simple limiting argument, we get (5.4.5) and (5.4.6) for f as well.

Step 2: Here we shall prove the inequality (5.4.5) for $n = 1$.

We shall need first to introduce a little more notations. In $\mathbb{R}$, let Δ_1 be the family of dyadic set $I_m = [2^m, 2^{m+1}] \cup [-2^{m+1}, -2^m]$ with $m \in \mathbb{Z}$. For each $I_m \in \Delta_1$, we consider the partial sum operator S_{I_m}, and a modification of it that we now define. Let $\varphi \in \mathscr{D}(\mathbb{R})$ be a fixed function with the following properties:

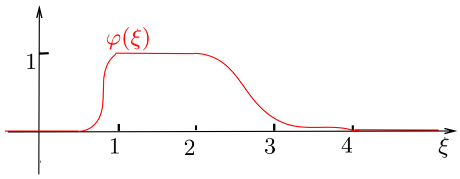


Figure 5.5: $\varphi(\xi)$

$$\varphi(\xi) = \begin{cases} 1, & 1 \leqslant |\xi| \leqslant 2, \\ 0, & |\xi| \leqslant 1/2, \text{ or } |\xi| \geqslant 4. \end{cases}$$

Define $\widetilde{S}_{I_m}$ by

$$\widehat{\widetilde{S}_{I_m} f}(\xi) = \varphi(2^{-m}\xi)\widehat{f}(\xi) = \varphi_m(\xi)\widehat{f}(\xi). \tag{5.4.7}$$

That is, $\widetilde{S}_{I_m}$, like S_{I_m}, is a multiplier transform where the multiplier is equal to one on the interval I_m; but unlike S_{I_m}, the multiplier of $\widetilde{S}_{I_m}$ is smooth. We observe that

$$S_{I_m}\widetilde{S}_{I_m} = S_{I_m}, \tag{5.4.8}$$

since S_{I_m} has the multiplier as the characteristic function of I_m.

Now for each $t \in [0, 1]$, consider the multiplier transform

$$\widetilde{T}_t = \sum_{m\in\mathbb{Z}} r_m(t)\widetilde{S}_{I_m}.$$

That is, for each t, $\widetilde{T}_t$ is the multiplier transform whose multiplier is $\widetilde{m}_t(\xi)$, with

$$\widetilde{m}_t(\xi) = \sum_{m\in\mathbb{Z}} r_m(t)\varphi_m(\xi). \tag{5.4.9}$$

By the definition of φ_m, it is clear that for any ξ at most five terms in the sum (5.4.9) can be non-zero. Moreover, we also see easily that

$$|\widetilde{m}_t(\xi)| \leqslant B, \quad \left| \frac{d\widetilde{m}_t}{d\xi}(\xi) \right| \leqslant \frac{B}{|\xi|}, \tag{5.4.10}$$

where B is independent of t. Thus, by the Mikhlin multiplier theorem (Theorem 5.2.1)

$$\|\widetilde{T}_t f\|_p \leqslant A_p\|f\|_p, \quad \text{for } f \in L^2 \cap L^p, \tag{5.4.11}$$

and with A_p independent of t. From this, it follows obviously that

$$\left(\int_0^1 \|\widetilde{T}_t f\|_p^p dt \right)^{1/p} \leqslant A_p\|f\|_p.$$

However, by Lemma 5.4.2 about the Rademacher functions,

$$\int_0^1 \|\widetilde{T}_t f\|_p^p dt = \int_0^1 \int_{\mathbb{R}} \left| \sum r_m(t)(\widetilde{S}_{I_m} f)(x) \right|^p dxdt$$

$$\geqslant A_p' \int_{\mathbb{R}} \left(\sum_m |\widetilde{S}_{I_m} f(x)|^2 \right)^{p/2} dx.$$

Thus, we have

$$\left\| \left(\sum_m |\widetilde{S}_{I_m}(f)|^2 \right)^{1/2} \right\|_p \leqslant B_p \|f\|_p. \tag{5.4.12}$$

Now using (5.4.8), applying the general theorem about partial sums, Theorem 5.3.4, with $\Re = \Delta_1$ here and (5.4.12), we get, for $F = (\widetilde{S}_{I_m} f)_{m\in\mathbb{Z}}$,

$$\left\| \left(\sum_m |S_{I_m} f|^2 \right)^{1/2} \right\|_p = \left\| \left(\sum_m |S_{I_m} \widetilde{S}_{I_m} f|^2 \right)^{1/2} \right\|_p = \|S_{\Delta_1} F\|_p$$

$$\leqslant A_p \|F\|_p = A_p \left\| \left(\sum_m |\widetilde{S}_{I_m} f|^2 \right)^{1/2} \right\|_p \leqslant A_p B_p \|f\|_p = C_p \|f\|_p, \tag{5.4.13}$$

which is the one-dimensional case of the inequality (5.4.5), and this is what we had set out to prove.

Step 3: We are still in the one-dimensional case, and we write T_t for the operator

$$T_t = \sum_m r_m(t) S_{I_m}.$$

Our claim is that

$$\|T_t f\|_p \leqslant A_p \|f\|_p, \quad 1 < p < \infty, \tag{5.4.14}$$

with A_p independent of t, and $f \in L^2 \cap L^p$.

Write $T_t^N = \sum_{|m|\leqslant N} r_m(t) S_{I_m}$, and it suffices to show that (5.4.14) holds, with T_t^N in place of T_t (and A_p independent of N and t). Since each S_{I_m} is a bounded operator on L^2 and L^p, we have that $T_t^N f \in L^2 \cap L^p$ and so we can apply Lemma 5.4.2 to it for $n = 1$. So

$$B_p \|T_t^N f\|_p \leqslant \left\| \left(\sum_{|m|\leqslant N} |S_{I_m} f|^2 \right)^{1/2} \right\|_p \leqslant C_p \|f\|_p,$$

by using (5.4.13). Letting $N \to \infty$, we get (5.4.14).

Step 4: We now turn to the n-dimensional case and define $T_{t_1}^{(1)}$, as the operator T_{t_1} acting only on the x_1 variable. Then, by the inequality (5.4.14), we get

$$\int_{\mathbb{R}} |T_{t_1}^{(1)} f(x_1, x_2, \cdots, x_n)|^p dx_1 \leqslant A_p^p \int_{\mathbb{R}} |f(x_1, \cdots, x_n)|^p dx_1, \tag{5.4.15}$$

for almost every fixed $x_2, x_3, \cdots, x_n$, since $x_1 \to f(x_1, x_2, \cdots, x_n) \in L^2_{x_1}(\mathbb{R}) \cap L^p_{x_1}(\mathbb{R})$ for almost every fixed $x_2, \cdots, x_n$, if $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. If we integrate (5.4.15) w.r.t. $x_2, \cdots, x_n$, we obtain

$$\|T_{t_1}^{(1)} f\|_p \leqslant A_p \|f\|_p, \quad f \in L^2 \cap L^p, \tag{5.4.16}$$

with A_p independent of t_1. The same inequality of course holds with x_1 replaced by x_2, or x_3, etc.

Step 5: We first describe the additional notation we shall need. With Δ representing the collection of dyadic rectangles in $\mathbb{R}^n$, we write any $\rho \in \Delta$, as $\rho = I_{m_1} \times I_{m_2} \times \cdots \times I_{m_n}$ where I_{m_j} represents the arbitrary dyadic set used above. Thus, if $m = (m_1, m_2, \cdots, m_n) \in \mathbb{Z}^n$, we write $\rho_m = I_{m_1} \times I_{m_2} \times \cdots \times I_{m_n}$.

We now apply the operator $T_{t_1}^{(1)}$ for the x_1 variable, and successively its analogues for x_2, x_3, etc. The result is

$$\|T_t f\|_p \leqslant A_p^n \|f\|_p. \tag{5.4.17}$$

Here

$$T_t = \sum_{\rho_m \in \Delta} r_m(t) S_{\rho_m}$$

with $r_m(t) = r_{m_1}(t_1) \cdots r_{m_n}(t_n)$ as described in the previous. The inequality holds uniformly for each $(t_1, t_2, \cdots, t_n)$ in the unit cube Q.

We raise this inequality to the p^{th} power and integrate it w.r.t. t, making use of the properties of the Rademacher functions, i.e., Lemma 5.4.2. We then get, as in the analogous proof of (5.4.12), that

$$\left\| \left(\sum_{\rho_m \in \Delta} |S_{\rho_m} f|^2 \right)^{1/2} \right\|_p \leqslant A_p \|f\|_p,$$

if $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. This together with the first step concludes the proof of Theorem 5.4.1. ■

§ 5.5 Marcinkiewicz multiplier theorem

We now present another multiplier theorem which is one of the most important results of the whole theory. For the sake of clarity, we state first the one-dimensional case whose proof is similar to the higher dimensional cases.

Theorem 5.5.1.

Let m be a bounded function on $\mathbb{R}$, which is of bounded variation on every finite interval not containing the origin. Suppose
(a) $|m(\xi)| \leqslant B$, $-\infty < \xi < \infty$,
(b) $\int_I |m(\xi)| d\xi \leqslant B$, for every dyadic interval I.
Then $m \in \mathcal{M}_p$, $1 < p < \infty$; and more precisely, for $f \in L^2 \cap L^p$,

$$\|T_m f\|_p \leqslant A_p \|f\|_p,$$

where A_p depends only on B and p.

To present the general theorem, we consider $\mathbb{R}$ as divided into its two half-lines, $\mathbb{R}^2$ as divided into its four quadrants, and generally $\mathbb{R}^n$ as divided into its 2^n "octants". Thus, the first octant in $\mathbb{R}^n$ will be the open "rectangle" of those ξ all of whose coordinates are strictly positive. We shall assume that $m(\xi)$ is defined on each such octant and is there continuous together with its partial derivatives up to and including order n. Thus m may be left undefined on the set of points where one or more coordinate variables vanishes.

For every $k \leqslant n$, we regard $\mathbb{R}^k$ embedded in $\mathbb{R}^n$ in the following obvious way: $\mathbb{R}^k$ is the subspace of all points of the form $(\xi_1, \xi_2, \cdots, \xi_k, 0, \cdots, 0)$.

Theorem 5.5.2: Marcinkiewicz' multiplier theorem

Let m be a bounded function on $\mathbb{R}^n$ that is $\mathscr{C}^n$ in all 2^n "octants". Suppose also
(a) $|m(\xi)| \leqslant B$,
(b) for each $0 < k \leqslant n$,

$$\sup_{\xi_{k+1},\cdots,\xi_n} \int_\rho \left| \frac{\partial^k m}{\partial\xi_1\partial\xi_2\cdots\partial\xi_k} \right| d\xi_1\cdots d\xi_k \leqslant B$$

as ρ ranges over dyadic rectangles of $\mathbb{R}^k$. (If $k = n$, the "sup" sign is omitted.)
(c) The condition analogous to (b) is valid for every one of the $n!$ permutations of the variables $\xi_1, \xi_2, \cdots, \xi_n$.
Then $m \in \mathcal{M}_p$, $1 < p < \infty$; and more precisely, for $f \in L^2 \cap L^p$, $\|T_m f\|_p \leqslant A_p \|f\|_p$, where A_p depends only on B, p and n.

Proof. It will be best to prove Theorem 5.5.2 in the case $n = 2$. This case is already completely typical of the general situation, and in doing only it we can avoid some notational complications.

Let $f \in L^2(\mathbb{R}^2) \cap L^p(\mathbb{R}^2)$ and $F = T_m f$, that is $\widehat{F}(\xi) = m(\xi)\widehat{f}(\xi)$.

Let Δ denote the family of dyadic rectangles, and for each $\rho \in \Delta$, we write $f_\rho = S_\rho f$, $F_\rho = S_\rho F$, thus $F_\rho = T_m f_\rho$.

In view of Theorem 5.4.1, it suffices to show that

$$\left\| \Big(\sum_{\rho\in\Delta} |F_\rho|^2 \Big)^{1/2} \right\|_p \leqslant C_p \left\| \Big(\sum_{\rho\in\Delta} |f_\rho|^2 \Big)^{1/2} \right\|_p. \tag{5.5.1}$$

The rectangles in Δ come from four sets, those in the first, the second, the third, and the fourth quadrants, respectively. In estimating the l.h.s. of (5.5.1), we consider the rectangles of each quadrant separately and assume from now on that the rectangles belong to the first quadrant.

We will express F_ρ in terms of an integral involving f_ρ and the partial sum operators. That this is possible is the essential idea of the proof.

Fix ρ and assume $\rho = \{(\xi_1, \xi_2) : 2^k \leqslant \xi_1 \leqslant 2^{k+1}, 2^l \leqslant \xi_2 \leqslant 2^{l+1}\}$. Then, for $(\xi_1, \xi_2) \in \rho$, it is easy to verify the identity by the fundamental theorem of calculus

$$\begin{aligned} m(\xi_1,\xi_2) = &\int_{2^l}^{\xi_2}\int_{2^k}^{\xi_1} \frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2} dt_1 dt_2 + \int_{2^k}^{\xi_1} \frac{\partial}{\partial t_1} m(t_1, 2^l) dt_1 \\ &+ \int_{2^l}^{\xi_2} \frac{\partial}{\partial t_2} m(2^k, t_2) dt_2 + m(2^k, 2^l). \end{aligned}$$

Now let S_t denote the multiplier transform corresponding to the rectangle $\{(\xi_1,\xi_2) : 2^{k+1} > \xi_1 > t_1,\ 2^{l+1} > \xi_2 > t_2\}$. Similarly, let $S^{(1)}_{t_1}$ denote the multiplier corresponding to the interval $\{\xi_1 : 2^{k+1} > \xi_1 > t_1\}$, similarly for $S^{(2)}_{t_2}$. Thus, in fact, $S_t = S^{(1)}_{t_1} \cdot S^{(2)}_{t_2}$. Multiplying both sides of the above equation by the function $\chi_\rho \widehat{f}$ and taking inverse Fourier transforms yields, by changing the order of integrals in view of Fubini's theorem and the fact that $S_\rho T_m f = F_\rho$, and $S^{(1)}_{t_1} S_\rho = S^{(1)}_{t_1}$, $S^{(2)}_{t_2} S_\rho = S^{(2)}_{t_2}$, $S_t S_\rho = S_t$, we have

$$\begin{aligned} F_\rho = &T_m S_\rho f = \mathscr{F}^{-1}(m\chi_\rho \widehat{f}) \\ = &\left(\frac{|\omega|}{2\pi}\right)^{n/2} \int_{\mathbb{R}^2} e^{\omega i x\cdot\xi} \Big[\int_{2^l}^{\xi_2}\int_{2^k}^{\xi_1} \frac{\partial^2 m(t_1,t_2)}{\partial t_1 \partial t_2} dt_1 dt_2 \chi_\rho(\xi)\widehat{f}(\xi) \Big] d\xi \end{aligned}$$

$$
\begin{aligned}
&+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\Big[\int_{2^k}^{\xi_1}\frac{\partial}{\partial t_1}m(t_1,2^l)dt_1\chi_\rho(\xi)\widehat{f}(\xi)\Big]d\xi\\
&+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\Big[\int_{2^l}^{\xi_2}\frac{\partial}{\partial t_2}m(2^k,t_2)dt_2\chi_\rho(\xi)\widehat{f}(\xi)\Big]d\xi\\
&+\mathscr{F}^{-1}[m(2^k,2^l)\chi_\rho(\xi)\widehat{f}(\xi)]\\
=&\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\int_{2^l}^{2^{l+1}}\int_{2^k}^{2^{k+1}}\frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}\chi_{[2^k,\xi_1]}(t_1)\chi_{[2^l,\xi_2]}(t_2)dt_1dt_2\\
&\qquad\qquad\cdot\chi_\rho(\xi)\widehat{f}(\xi)d\xi\\
&+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\int_{2^k}^{2^{k+1}}\frac{\partial}{\partial t_1}m(t_1,2^l)\chi_{[2^k,\xi_1]}(t_1)dt_1\chi_\rho(\xi)\widehat{f}(\xi)d\xi\\
&+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\int_{2^l}^{2^{l+1}}\frac{\partial}{\partial t_2}m(2^k,t_2)\chi_{[2^l,\xi_2]}(t_2)dt_2\chi_\rho(\xi)\widehat{f}(\xi)d\xi\\
&+m(2^k,2^l)f_\rho\\
=&\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{2^l}^{2^{l+1}}\int_{2^k}^{2^{k+1}}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\chi_{[t_1,2^{k+1}]}(\xi_1)\chi_{[t_2,2^{l+1}]}(\xi_2)\chi_\rho(\xi)\widehat{f}(\xi)d\xi\\
&\qquad\qquad\cdot\frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}dt_1dt_2\\
&+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{2^k}^{2^{k+1}}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\chi_{[t_1,2^{k+1}]}(\xi_1)\chi_\rho(\xi)\widehat{f}(\xi)d\xi\frac{\partial}{\partial t_1}m(t_1,2^l)dt_1\\
&+\left(\frac{|\omega|}{2\pi}\right)^{n/2}\int_{2^l}^{2^{l+1}}\int_{\mathbb{R}^2} e^{\omega i x\cdot\xi}\chi_{[t_2,2^{l+1}]}(\xi_2)\chi_\rho(\xi)\widehat{f}(\xi)d\xi\frac{\partial}{\partial t_2}m(2^k,t_2)dt_2\\
&+m(2^k,2^l)f_\rho\\
=&\int_\rho S_tf_\rho\frac{\partial^2 m(t_1,t_2)}{\partial t_1\partial t_2}dt_1dt_2+\int_{2^k}^{2^{k+1}}S^{(1)}_{t_1}f_\rho\frac{\partial}{\partial t_1}m(t_1,2^l)dt_1\\
&+\int_{2^l}^{2^{l+1}}S^{(2)}_{t_2}f_\rho\frac{\partial}{\partial t_2}m(2^k,t_2)dt_2+m(2^k,2^l)f_\rho.
\end{aligned}
$$

We apply the Cauchy-Schwarz inequality in the first three terms of the above w.r.t. the measures $|\partial_{t_1}\partial_{t_2}m(t_1,t_2)|dt_1dt_2$, $|\partial_{t_1}m(t_1,2^l)|dt_1$, $|\partial_{t_2}m(2^k,t_2)|dt_2$, respectively, and we use the assumptions of the theorem to deduce

$$
\begin{aligned}
|F_\rho|^2\lesssim&\Big(\int_\rho|S_tf_\rho|^2\left|\frac{\partial^2 m}{\partial t_1\partial t_2}\right|dt_1dt_2\Big)\Big(\int_\rho\left|\frac{\partial^2 m}{\partial t_1\partial t_2}\right|dt_1dt_2\Big)\\
&+\Big(\int_{2^k}^{2^{k+1}}|S^{(1)}_{t_1}f_\rho|^2\left|\frac{\partial}{\partial t_1}m(t_1,2^l)\right|dt_1\Big)\Big(\int_{2^k}^{2^{k+1}}\left|\frac{\partial}{\partial t_1}m(t_1,2^l)\right|dt_1\Big)\\
&+\Big(\int_{2^l}^{2^{l+1}}|S^{(2)}_{t_2}f_\rho|^2\left|\frac{\partial}{\partial t_2}m(2^k,t_2)\right|dt_2\Big)\Big(\int_{2^l}^{2^{l+1}}\left|\frac{\partial}{\partial t_2}m(2^k,t_2)\right|dt_2\Big)\\
&+|m(2^k,2^l)|^2|f_\rho|^2\\
\leqslant&B'\Big\{\int_\rho|S_tf_\rho|^2\left|\frac{\partial^2 m}{\partial t_1\partial t_2}\right|dt_1dt_2+\int_{I_1}|S^{(1)}_{t_1}f_\rho|^2\left|\frac{\partial m(t_1,2^l)}{\partial t_1}\right|dt_1\\
&\qquad+\int_{I_2}|S^{(2)}_{t_2}f_\rho|^2\left|\frac{\partial m(2^k,t_2)}{\partial t_2}\right|dt_2+|f_\rho|^2\Big\}\\
=&\Im^1_\rho+\Im^2_\rho+\Im^3_\rho+\Im^4_\rho,\ \text{with } \rho=I_1\times I_2.
\end{aligned}
$$

To estimate $\|(\sum_\rho|F_\rho|^2)^{1/2}\|_p$, we estimate separately the contributions of each of the four terms on the r.h.s. of the above inequality by the use of Theorem 5.3.5. To apply that theorem in the case of $\Im_\rho^1$ we take the first quadrant as Γ and $d\gamma = |\frac{\partial^2 m(t_1,t_2)}{\partial t_1 \partial t_2}| dt_1 dt_2$, the functions $\gamma \to \rho_\gamma$ are constant on the dyadic rectangles. Since for every rectangle, we have

$$\int_\rho d\gamma = \int_\rho \left| \frac{\partial^2 m(t_1, t_2)}{\partial t_1 \partial t_2} \right| dt_1 dt_2 \leqslant B,$$

then

$$\left\| \left(\sum_\rho |\Im_\rho^1| \right)^{1/2} \right\|_p \leqslant C_p \left\| \left(\sum_\rho |f_\rho|^2 \right)^{1/2} \right\|_p.$$

The similar argument for $\Im_\rho^2$, $\Im_\rho^3$ and $\Im_\rho^4$ concludes the proof. ■

References

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

VI

Smoothness and Function Spaces

§ 6.1 Riesz potentials and fractional integrals

The Laplacian satisfies the following identity for all $f \in \mathscr{S}(\mathbb{R}^n)$:

$$\widehat{-\Delta f}(\xi) = \omega^2 |\xi|^2 \widehat{f}(\xi). \tag{6.1.1}$$

From this, we replace the exponent 2 in $|\omega\xi|^2$ by a general exponent s, and thus to define (at least formally) the fractional power of the Laplacian by

$$(-\Delta)^{s/2} f = \mathscr{F}^{-1}((|\omega||\xi|)^s \widehat{f}(\xi)). \tag{6.1.2}$$

Of special significance will be the negative powers s in the range $-n < s < 0$. In general, with a slight change of notation, we can define

Definition 6.1.1.

Let $s > 0$. The ***Riesz potential*** of order s is the operator

$$I_s = (-\Delta)^{-s/2}. \tag{6.1.3}$$

For $0 < s < n$ and $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, I_s is actually given in the form

$$I_s f(x) = \frac{1}{\gamma(s)} \int_{\mathbb{R}^n} |x-y|^{-n+s} f(y) dy, \tag{6.1.4}$$

with

$$\gamma(s) = \left(\frac{|\omega|}{2}\right)^{\frac{n}{2}} 2^s \frac{\Gamma(s/2)}{\Gamma((n-s)/2)}.$$

The formal manipulations have a precise meaning.

Lemma 6.1.2.

Let $0 < s < n$.
(a) The Fourier transform of the function $|x|^{-n+s}$ is the function $\gamma(s)(|\omega||\xi|)^{-s}$, in the sense that

$$\int_{\mathbb{R}^n} |x|^{-n+s}\overline{\varphi(x)}dx = \gamma(s)\int_{\mathbb{R}^n} (|\omega||\xi|)^{-s}\overline{\widehat{\varphi}(\xi)}d\xi, \tag{6.1.5}$$

whenever $\varphi \in \mathscr{S}$.
(b) The identity $\widehat{I_s f}(\xi) = (|\omega||\xi|)^{-s}\widehat{f}(\xi)$ holds in the sense that

$$\int_{\mathbb{R}^n} I_s f(x)\overline{g(x)}dx = \int_{\mathbb{R}^n} \widehat{f}(\xi)(|\omega||\xi|)^{-s}\overline{\widehat{g}(\xi)}d\xi,$$

whenever $f, g \in \mathscr{S}$.

Proof. Part (a) is merely a restatement of Lemma 4.4.17 since $\gamma(s) = |\omega|^s\gamma_{0,s}$.

Part (b) follows immediately from part (a) by writing

$$\begin{aligned} I_s f(x) =& \frac{1}{\gamma(s)}\int_{\mathbb{R}^n} f(x-y)|y|^{-n+s}dy = \int_{\mathbb{R}^n} (|\omega||\xi|)^{-s}\overline{\widehat{\overline{f(x-\cdot)}}}d\xi \\ =& \int_{\mathbb{R}^n} (|\omega||\xi|)^{-s}\widehat{f}(\xi)e^{\omega i\xi\cdot x}d\xi = \int_{\mathbb{R}^n} (|\omega||\xi|)^{-s}\widehat{f}(\xi)\overline{e^{-\omega i\xi\cdot x}}d\xi, \end{aligned}$$

so

$$\begin{aligned} \int_{\mathbb{R}^n} I_s f(x)\overline{g(x)}dx =& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (|\omega||\xi|)^{-s}\widehat{f}(\xi)\overline{e^{-\omega i\xi\cdot x}}d\xi\overline{g(x)}dx \\ =& \int_{\mathbb{R}^n} (|\omega||\xi|)^{-s}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}d\xi. \end{aligned}$$

This completes the proof. ■

Now, we state two further identities which can be obtained from (6.1.2) or (6.1.3) and which reflect essential properties of the potentials I_s:

$$I_s(I_t f) = I_{s+t}f, \quad f \in \mathscr{S},\ s,t > 0,\ s+t < n; \tag{6.1.6}$$

$$\Delta(I_s f) = I_s(\Delta f) = -I_{s-2}f, \quad f \in \mathscr{S},\ n \geqslant 3,\ 2 \leqslant s \leqslant n. \tag{6.1.7}$$

The deduction of these two identities have no real difficulties, and these are best left to the interested reader to work out.

A simple consequence of (6.1.6) is the n-dimensional variant of the Beta function,

$$\int_{\mathbb{R}^n} |x-y|^{-n+s}|y|^{-n+t}dy = \frac{\gamma(s)\gamma(t)}{\gamma(s+t)}|x|^{-n+(s+t)} \quad \text{in } \mathscr{S}', \tag{6.1.8}$$

with $s, t > 0$ and $s + t < n$. Indeed, for any $\varphi \in \mathscr{S}$, we have, by the definition of Riesz potentials and (6.1.6), that

$$\begin{aligned} &\iint_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^{-n+s}|y|^{-n+t}dy\varphi(z-x)dx \\ =& \int_{\mathbb{R}^n} |y|^{-n+t}\int_{\mathbb{R}^n} |x-y|^{-n+s}\varphi(z-y-(x-y))dxdy \\ =& \int_{\mathbb{R}^n} |y|^{-n+t}\gamma(s)I_s\varphi(z-y)dy = \gamma(s)\gamma(t)I_t(I_s\varphi)(z) = \gamma(s)\gamma(t)I_{s+t}\varphi(z) \\ =& \frac{\gamma(s)\gamma(t)}{\gamma(s+t)}\int_{\mathbb{R}^n} |x|^{-n+(s+t)}\varphi(z-x)dx. \end{aligned}$$

We have considered the Riesz potentials formally and the operation for Schwartz functions. But since the Riesz potentials are integral operators, it is natural to inquire about their actions on the spaces $L^p(\mathbb{R}^n)$.

For this reason, we formulate the following problem. Given $s \in (0, n)$, for what pairs p and q, is the operator $f \to I_s f$ bounded from $L^p(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$? That is, when do we have the inequality

$$\|I_s f\|_q \leqslant A\|f\|_p? \tag{6.1.9}$$

There is a simple necessary condition, which is merely a reflection of the homogeneity of the kernel $(\gamma(s))^{-1}|y|^{-n+s}$. In fact, we have

Proposition 6.1.3.

If the inequality (6.1.9) 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

$$\begin{aligned}(I_s\delta^\varepsilon f)(x) =& \frac{1}{\gamma(s)}\int_{\mathbb{R}^n}|x-y|^{-n+s}f(\varepsilon y)dy\\ \overset{z=\varepsilon y}{=\!=\!=}&\varepsilon^{-n}\frac{1}{\gamma(s)}\int_{\mathbb{R}^n}|x-\varepsilon^{-1}z|^{-n+s}f(z)dz\\ =&\varepsilon^{-s}I_sf(\varepsilon x).\end{aligned} \tag{6.1.10}$$

Noticing that

$$\|\delta^\varepsilon f\|_p = \varepsilon^{-n/p}\|f\|_p, \quad \|\delta_{\varepsilon^{-1}} I_s f\|_q = \varepsilon^{n/q}\|I_s f\|_q, \tag{6.1.11}$$

by (6.1.9), we get

$$\begin{aligned}\|I_sf\|_q =&\varepsilon^s\|\delta^{\varepsilon^{-1}}I_s\delta^\varepsilon f\|_q = \varepsilon^{s+n/q}\|I_s\delta^\varepsilon f\|_q\\ \leqslant& A\varepsilon^{s+n/q}\|\delta^\varepsilon f\|_p = A\varepsilon^{s+n/q-n/p}\|f\|_p.\end{aligned}$$

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 non-negative, then $I_s f > 0$ everywhere and hence $\|I_s f\|_q > 0$, thus we can conclude the desired relations

$$1/q = 1/p - s/n. \tag{6.1.12}$$

■

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 6.1.4: Hardy-Littlewood-Sobolev theorem

Let $0 < s < n$, $1 \leqslant p < q < \infty$, $1/q = 1/p - s/n$.
(a) If $f \in L^p(\mathbb{R}^n)$, then the integral (6.1.4), defining $I_s f$, converges absolutely for almost every x.
(b) If, in addition, $p > 1$, then $\|I_s f\|_q \leqslant A_{p,q}\|f\|_p$.
(c) If $f \in L^1(\mathbb{R}^n)$, then $|\{x : |I_s f(x)| > \alpha\}| \leqslant (A\alpha^{-1}\|f\|_1)^q$, 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 (a) and (b). Let us write

$$\begin{aligned}\gamma(s)I_sf(x)=&\int_{B(x,\delta)}|x-y|^{-n+s}f(y)dy+\int_{\mathbb{R}^n\setminus B(x,\delta)}|x-y|^{-n+s}f(y)dy\\=:&L_\delta(x)+H_\delta(x).\end{aligned}$$

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

$$\begin{aligned}|L_\delta(x)|\leqslant&\left|\sum_{j=0}^\infty\int_{E_j}|x-y|^{-n+s}f(y)dy\right|\leqslant\sum_{j=0}^\infty\int_{E_j}|x-y|^{-n+s}|f(y)|dy\\\leqslant&\sum_{j=0}^\infty\int_{E_j}(2^{-(j+1)}\delta)^{-n+s}|f(y)|dy\\\leqslant&\sum_{j=0}^\infty\int_{B(x,2^{-j}\delta)}(2^{-(j+1)}\delta)^{-n+s}|f(y)|dy\\=&\sum_{j=0}^\infty\frac{(2^{-(j+1)}\delta)^{-n+s}|B(x,2^{-j}\delta)|}{|B(x,2^{-j}\delta)|}\int_{B(x,2^{-j}\delta)}|f(y)|dy\\=&\sum_{j=0}^\infty\frac{(2^{-(j+1)}\delta)^{-n+s}V_n(2^{-j}\delta)^n}{|B(x,2^{-j}\delta)|}\int_{B(x,2^{-j}\delta)}|f(y)|dy\\\leqslant&V_n\delta^s2^{n-s}\sum_{j=0}^\infty2^{-sj}Mf(x)=\frac{V_n\delta^s2^n}{2^s-1}Mf(x).\end{aligned}$$

Now, we derive an estimate for $H_\delta(x)$. By Hölder's inequality and the condition $1/p>s/n$ (i.e., $q<\infty$), we obtain

$$\begin{aligned}|H_\delta(x)|\leqslant&\|f\|_p\left(\int_{\mathbb{R}^n\setminus B(x,\delta)}|x-y|^{(-n+s)p'}dy\right)^{1/p'}\\=&\|f\|_p\left(\int_{S^{n-1}}\int_\delta^\infty r^{(-n+s)p'}r^{n-1}drd\sigma\right)^{1/p'}\\=&\omega_{n-1}^{1/p'}\|f\|_p\left(\int_\delta^\infty r^{(-n+s)p'+n-1}dr\right)^{1/p'}\\=&\left(\frac{\omega_{n-1}}{(n-s)p'-n}\right)^{1/p'}\delta^{n/p'-(n-s)}\|f\|_p=C(n,s,p)\delta^{s-n/p}\|f\|_p.\end{aligned}$$

By the above two inequalities, we have

$$|\gamma(s)I_sf(x)|\leqslant C(n,s)\delta^sMf(x)+C(n,s,p)\delta^{s-n/p}\|f\|_p=:F(\delta).$$

Choose $\delta=C(n,s,p)[\|f\|_p/Mf]^{p/n}$, such that the two terms of the r.h.s. of the above are equal, i.e., the minimizer of $F(\delta)$, to get

$$|\gamma(s)I_sf(x)|\leqslant C(Mf(x))^{1-ps/n}\|f\|_p^{ps/n}.$$

Therefore, by part (i) of Theorem 3.2.7 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 (a) of the theorem.

By part (iii) of Theorem 3.2.7, 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 (b).

Finally, we prove (c). Since we also have $|H_\delta(x)| \leqslant \|f\|_1\delta^{-n+s}$, taking $\alpha = \|f\|_1\delta^{-n+s}$, i.e., $\delta = (\|f\|_1/\alpha)^{1/(n-s)}$, by part (ii) of Theorem 3.2.7, we get

$$\begin{aligned}
&|\{x : |I_s f(x)| > 2(\gamma(s))^{-1}\alpha\}| \\
\leqslant &|\{x : |L_\delta(x)| > \alpha\}| + |\{x : |H_\delta(x)| > \alpha\}| \\
\leqslant &|\{x : |C\delta^s Mf(x)| > \alpha\}| + 0 \\
\leqslant &\frac{C}{\delta^{-s}\alpha}\|f\|_1 = C[\|f\|_1/\alpha]^{n/(n-s)} = C[\|f\|_1/\alpha]^q.
\end{aligned}$$

This completes the proof of part (c). ■

§ 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 $|x| \to \infty$ gets worse as s increases.

We can slightly adjust the Riesz potentials such that we maintain their essential behavior near zero but achieve exponential decay at infinity. The simplest way to achieve this is by replacing the "nonnegative" operator $-\Delta$ by the "strictly positive" operator $I - \Delta$, where $I =$ identity. Here the terms nonnegative and strictly positive, as one may have surmised, refer to the Fourier transforms of these expressions.

Definition 6.2.1.

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

$$J_s f = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \mathscr{F}^{-1}(\widehat{G_s}\widehat{f}) = G_s * f,$$

where

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

Now we give some properties of $G_s(x)$ and show why this adjustment yields exponential decay for G_s at infinity.

Proposition 6.2.2.

Let $s > 0$.

(a) $G_s(x) = \dfrac{1}{(4\pi)^{n/2}\Gamma(s/2)} \displaystyle\int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}$.

(b) $G_s(x) > 0, \quad \forall x \in \mathbb{R}^n$; and $G_s \in L^1(\mathbb{R}^n)$, precisely, $\int_{\mathbb{R}^n} G_s(x)dx = 1$.

(c) There exist two constants $0 < C(s,n), c(s,n) < \infty$ such that

$$G_s(x) \leqslant C(s,n)e^{-|x|/2}, \quad \text{when } |x| \geqslant 2,$$

and

$$\frac{1}{c(s,n)} \leqslant \frac{G_s(x)}{H_s(x)} \leqslant c(s,n), \quad \text{when } |x| \leqslant 2,$$

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 $|x| \to 0$.

(d) $G_s \in L^{p'}(\mathbb{R}^n)$ for any $1 \leqslant p \leqslant \infty$ and $s > n/p$.

Proof. (a) For $A, s > 0$, we have the Γ-function identity

$$A^{-s/2} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-tA} t^{s/2} \frac{dt}{t},$$

which we use to obtain

$$(1 + \omega^2|\xi|^2)^{-s/2} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} e^{-t|\omega\xi|^2} t^{s/2} \frac{dt}{t}.$$

Note that the above integral converges at both ends (as $|\xi| \to 0$, or ∞). Now taking the inverse Fourier transform in ξ and using Theorem 2.1.9, we obtain

$$\begin{aligned} G_s(x) &= \left(\frac{|\omega|}{2\pi}\right)^{n/2} \frac{1}{\Gamma(s/2)} \mathscr{F}_\xi^{-1} \int_0^\infty e^{-t} e^{-t|\omega\xi|^2} t^{s/2} \frac{dt}{t} \\ &= \left(\frac{|\omega|}{2\pi}\right)^{n/2} \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} \mathscr{F}_\xi^{-1} \left(e^{-t|\omega\xi|^2}\right) t^{s/2} \frac{dt}{t} \\ &= \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}. \end{aligned}$$

(b) We have easily[1] $\int_{\mathbb{R}^n} G_s(x)dx = \left(\frac{|\omega|}{2\pi}\right)^{-n/2} \mathscr{F} G_s(0) = 1$. Thus, $G_s \in L^1(\mathbb{R}^n)$.

(c) First, we suppose $|x| \geqslant 2$. Then $t + \frac{|x|^2}{4t} \geqslant t + \frac{1}{t}$ and also $t + \frac{|x|^2}{4t} \geqslant |x|$. This implies that

$$-t - \frac{|x|^2}{4t} \leqslant -\frac{t}{2} - \frac{1}{2t} - \frac{|x|}{2},$$

from which it follows that when $|x| \geqslant 2$

$$G_s(x) \leqslant \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-\frac{t}{2}} e^{-\frac{1}{2t}} t^{\frac{s-n}{2}} \frac{dt}{t} e^{-\frac{|x|}{2}} \leqslant C(s,n) e^{-\frac{|x|}{2}},$$

where $C(s,n) = \frac{2^{|s-n|/2}\Gamma(|s-n|/2)}{(4\pi)^{n/2}\Gamma(s/2)}$ for $s \neq n$, and $C(s,n) = \frac{4}{(4\pi)^{n/2}\Gamma(s/2)}$ for $s = n$ since

$$\int_0^\infty e^{-\frac{t}{2}} e^{-\frac{1}{2t}} \frac{dt}{t} \leqslant \int_0^1 e^{-\frac{1}{2t}} \frac{dt}{t} + \int_1^\infty e^{-\frac{t}{2}} dt = \int_{1/2}^\infty e^{-y} \frac{dy}{y} + 2e^{-1/2}$$

[1] Or use (a) to show it. From part (a), we know $G_s(x) > 0$. Since $\int_{\mathbb{R}^n} e^{-\pi|x|^2/t} dx = t^{n/2}$, by Fubini's theorem, we have

$$\begin{aligned} \int_{\mathbb{R}^n} G_s(x)dx &= \int_{\mathbb{R}^n} \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t} dx \\ &= \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} \int_{\mathbb{R}^n} e^{-\frac{|x|^2}{4t}} dx t^{\frac{s-n}{2}} \frac{dt}{t} \\ &= \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^\infty e^{-t} (4\pi t)^{n/2} t^{\frac{s-n}{2}} \frac{dt}{t} = \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-t} t^{\frac{s}{2}-1} dt = 1. \end{aligned}$$

$$\leqslant 2\int_{1/2}^{\infty} e^{-y}dy + 2 \leqslant 4.$$

Next, suppose that $|x| \leqslant 2$. Write $G_s(x) = G_s^1(x) + G_s^2(x) + G_s^3(x)$, where

$$G_s^1(x) = \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^{|x|^2} e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t},$$

$$G_s^2(x) = \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_{|x|^2}^{4} e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t},$$

$$G_s^3(x) = \frac{1}{(4\pi)^{n/2}\Gamma(s/2)} \int_4^{\infty} e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}.$$

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

$$\begin{aligned} G_s^1(x) =& \frac{|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^1 e^{-t|x|^2} e^{-\frac{1}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t} \\ =& \frac{|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^1 e^{-\frac{1}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t} + \frac{O(|x|^{s-n+2})}{(4\pi)^{n/2}\Gamma(s/2)} \int_0^1 e^{-\frac{1}{4t}} t^{\frac{s-n}{2}} dt \\ =& \frac{2^{n-s-2}|x|^{s-n}}{(4\pi)^{n/2}\Gamma(s/2)} \int_{1/4}^{\infty} e^{-y} y^{\frac{n-s}{2}} \frac{dy}{y} + \frac{2^{n-s-4}O(|x|^{s-n+2})}{(4\pi)^{n/2}\Gamma(s/2)} \int_{1/4}^{\infty} e^{-y} y^{\frac{n-s}{2}} \frac{dy}{y^2} \\ =& c_{s,n}^1 |x|^{s-n} + O(|x|^{s-n+2}), \qquad \text{as } |x| \to 0. \end{aligned}$$

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

$$G_s^2(x) \sim \int_{|x|^2}^4 t^{(s-n)/2} \frac{dt}{t} = \begin{cases} \frac{|x|^{s-n}}{n-s} - \frac{2^{s-n+1}}{n-s}, & s < n, \\ 2\ln\frac{2}{|x|}, & s = n, \\ \frac{2^{s-n+1}}{s-n}, & s > n. \end{cases}$$

Finally, we have $e^{-1/4} \leqslant e^{-\frac{|x|^2}{4t}} \leqslant 1$ in G_s^3, which yields that $G_s^3(x)$ is bounded above and below by fixed positive constants. Combining the estimates for $G_s^j(x)$, we obtain the desired conclusion.

(d) For $p = 1$ and so $p' = \infty$, by part (c), we have $\|G_s\|_\infty \leqslant C$ for $s > n$.

Next, we assume that $1 < p \leqslant \infty$ and so $1 \leqslant p' < \infty$. Again by part (c), we have, for $|x| \geqslant 2$, that $G_s^{p'} \leqslant Ce^{-p'|x|/2}$, and then the integration over this range $|x| \geqslant 2$ is clearly finite.

On the range $|x| \leqslant 2$, it is clear that $\int_{|x|\leqslant 2} G_s^{p'}(x)dx \leqslant C$ for $s > n$. For the case $s = n$ and $n \neq 1$, we also have $\int_{|x|\leqslant 2} G_s^{p'}(x)dx \leqslant C$ by noticing that

$$\int_{|x|\leqslant 2} \left(\ln\frac{2}{|x|}\right)^q dx = C\int_0^2 \left(\ln\frac{2}{r}\right)^q r^{n-1}dr \leqslant C$$

for any $q > 0$ since $\lim_{r\to 0} r^\varepsilon \ln(2/r) = 0$. For the case $s = n = 1$, we have

$$\begin{aligned} \int_{|x|\leqslant 2} (\ln\frac{2}{|x|})^q dx =& 2\int_0^2 (\ln 2/r)^q dr = 4\int_0^1 (\ln 1/r)^q dr \\ =& 4\int_0^\infty t^q e^{-t} dt = 4\Gamma(q+1) \end{aligned}$$

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. ■

We also have a result analogues to that of Riesz potentials for the operator J_s.

Theorem 6.2.3.

(a) For all $0 < s < \infty$, the operator J_s maps $L^r(\mathbb{R}^n)$ into itself with norm 1 for all $1 \leqslant r \leqslant \infty$.
(b) Let $0 < s < n$ and $1 < p < q < \infty$ satisfy $1/q = 1/p - s/n$. Then there exists a constant $C_{n,s,p} > 0$ such that for all $f \in L^p(\mathbb{R}^n)$, we have

$$\|J_s f\|_q \leqslant C_{n,s,p} \|f\|_p.$$

(c) 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 the result (a).

In the special case $0 < s < n$, we have, from the above proposition, that the kernel G_s of J_s satisfies

$$G_s(x) \sim \begin{cases} |x|^{-n+s}, & |x| \leqslant 2, \\ e^{-|x|/2}, & |x| \geqslant 2. \end{cases}$$

Then, we can write

$$\begin{aligned} J_s f(x) \leqslant & C_{n,s} \left[\int_{|y| \leqslant 2} |f(x-y)||y|^{-n+s} dy + \int_{|y| \geqslant 2} |f(x-y)| e^{-|y|/2} dy \right] \\ \leqslant & C_{n,s} \left[I_s(|f|)(x) + \int_{\mathbb{R}^n} |f(x-y)| e^{-|y|/2} dy \right]. \end{aligned}$$

We can use that the function $e^{-|y|/2} \in L^r$ for all $1 \leqslant r \leqslant \infty$, Young's inequality and Theorem 6.1.4 to complete the proofs of (b) and (c). ■

§ 6.3 Sobolev spaces

We start by weakening the notation of partial derivatives by the theory of distributions. The appropriate definition is stated in terms of the space $\mathscr{D}(\mathbb{R}^n)$.

Let ∂^α be a differential monomial, whose total order is $|\alpha|$. Suppose we are given two locally integrable functions on $\mathbb{R}^n$, f and g. Then we say that $\partial^\alpha f = g$ (in the weak sense), if

$$\int_{\mathbb{R}^n} f(x) \partial^\alpha \varphi(x) dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} g(x) \varphi(x) dx, \quad \forall \varphi \in \mathscr{D}. \tag{6.3.1}$$

Integration by parts shows us that this is indeed the relation that we would expect if f had continuous partial derivatives up to order $|\alpha|$, and $\partial^\alpha f = g$ had the usual meaning.

Of course, it is not true that every locally integrable function has partial derivatives in this sense: consider, for example, $f(x) = c^{i/|x|^n}$. However, when the partial derivatives exist, they are determined almost everywhere by the defining relation (6.3.1).

In this section, we study a quantitative way of measuring smoothness of functions. Sobolev spaces serve exactly this purpose. They measure the smoothness of a given function in terms of the integrability of its derivatives. We begin with the classical definition of Sobolev spaces.

Definition 6.3.1.

Let k be a nonnegative integer and let $1 \leqslant p \leqslant \infty$. The ***Sobolev space*** $W^{k,p}(\mathbb{R}^n)$ is defined as the space of functions f in $L^p(\mathbb{R}^n)$ all of whose distributional derivatives $\partial^\alpha f$ are also in $L^p(\mathbb{R}^n)$ for all multi-indices α that satisfies $|\alpha| \leqslant k$. This space is normed by the expression

$$\|f\|_{W^{k,p}} = \sum_{|\alpha| \leqslant k} \|\partial^\alpha f\|_p, \tag{6.3.2}$$

where $\partial^{(0,\dots,0)} f = f$.

The index k indicates the "degree" of smoothness of a given function in $W^{k,p}$. As k increases, the functions become smoother. Equivalently, these spaces form a decreasing sequence

$$L^p \supset W^{1,p} \supset W^{2,p} \supset \cdots$$

meaning that each $W^{k+1,p}(\mathbb{R}^n)$ is a subspace of $W^{k,p}(\mathbb{R}^n)$ in view of the Sobolev norms.

We next observe that ***the space*** $W^{k,p}(\mathbb{R}^n)$ ***is complete***. Indeed, if $\{f_m\}$ is a Cauchy sequence in $W^{k,p}$, then for each α, $\{\partial^\alpha f_m\}$ is a Cauchy sequence in L^p, $|\alpha| \leqslant k$. By the completeness of L^p, there exist functions $f^{(\alpha)}$ such that $f^{(\alpha)} = \lim_m \partial^\alpha f_m$ in L^p, then clearly

$$(-1)^{|\alpha|} \int_{\mathbb{R}^n} f_m \partial^\alpha \varphi dx = \int_{\mathbb{R}^n} \partial^\alpha f_m \varphi dx \to \int_{\mathbb{R}^n} f^{(\alpha)} \varphi dx,$$

for each $\varphi \in \mathscr{D}$. Since the first expression converges to

$$(-1)^{|\alpha|} \int_{\mathbb{R}^n} f \partial^\alpha \varphi dx,$$

it follows that the distributional derivative $\partial^\alpha f$ is $f^{(\alpha)}$. This implies that $f_j \to f$ in $W^{k,p}(\mathbb{R}^n)$ and proves the completeness of this space.

First, we generalize Riesz and Bessel potentials to any $s \in \mathbb{R}$ by

$$I^s f = \mathscr{F}^{-1}(|\omega\xi|^s \widehat{f}), \quad f \in \mathscr{S}'(\mathbb{R}^n),\ 0 \notin \operatorname{supp} \widehat{f}, \tag{6.3.3}$$

$$J^s f = \mathscr{F}^{-1}((1+|\omega\xi|^2)^{s/2} \widehat{f}), \quad f \in \mathscr{S}'(\mathbb{R}^n). \tag{6.3.4}$$

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 (6.3.3) 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 $g \in \mathscr{S}$

$$\begin{aligned}\int_{\mathbb{R}^n} (I^k x^\alpha) g(x) dx =& \int_{\mathbb{R}^n} x^\alpha \widehat{|\omega\xi|^k \check{g}}(x) dx \\ =& \int_{\mathbb{R}^n} e^{i\omega x \cdot 0} (\omega i)^{-|\alpha|} \widehat{\partial_\xi^\alpha (|\omega\xi|^k \check{g})}(x) dx\end{aligned}$$

$$= \left(\frac{|\omega|}{2\pi}\right)^{-n/2} (\omega i)^{-|\alpha|} \left[\partial_\xi^\alpha (|\omega\xi|^k \check{g})\right](0) = 0.$$

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.

Definition 6.3.2.

Define

$$\dot{\mathscr{S}}(\mathbb{R}^n) = \left\{ f \in \mathscr{S}(\mathbb{R}^n) : \int_{\mathbb{R}^n} x^\alpha f(x) dx = 0, \ \forall \alpha \in \mathbb{N}_0^n \right\}, \tag{6.3.5}$$

which is a subspace of $\mathscr{S}(\mathbb{R}^n)$ with the same topology.

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 6.3.3.

The dual space of $\dot{\mathscr{S}}(\mathbb{R}^n)$ under the topology inherited from $\mathscr{S}(\mathbb{R}^n)$ is

$$\dot{\mathscr{S}}'(\mathbb{R}^n) = \mathscr{S}'(\mathbb{R}^n)/\mathscr{P}(\mathbb{R}^n).$$

Proof. To identify the dual of $\dot{\mathscr{S}}(\mathbb{R}^n)$, we argue as follows. For each $u \in \mathscr{S}'(\mathbb{R}^n)$, let $J(u) = u|_{\dot{\mathscr{S}}(\mathbb{R}^n)}$ 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 $\dot{\mathscr{S}}'(\mathbb{R}^n)$.

Firstly, we claim that the kernel of J is exactly $\mathscr{P}(\mathbb{R}^n)$. 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 2.4.25. 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 6.3.3, we have that $u_j \to u$ in $\dot{\mathscr{S}}'$ if and only if 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 $\dot{\mathscr{S}}$ implies convergence in $\mathscr{S}$, 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 that $|\widehat{\phi}(\xi)| \leqslant C_m|\xi|^M$ for every $M \in \mathbb{Z}^+$, whenever 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 $\widehat{\phi}(\xi)$. But for each $|\beta| \leqslant |\alpha|$, we have $|\partial^\beta\widehat{\phi}(\xi)| \leqslant C_{M,\beta}|\xi|^M$ for all $M \in \mathbb{Z}^+$ 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 $\mathscr{F}^{-1}(|\xi|^s\hat{u}) \in \dot{\mathscr{S}}'(\mathbb{R}^n)$ by setting for all $\phi \in \dot{\mathscr{S}}(\mathbb{R}^n)$

$$\langle \mathscr{F}^{-1}(|\cdot|^s\hat{u}), \phi\rangle = \langle u, \widehat{|\cdot|^s\check{\phi}}\rangle.$$

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.

Next, we shall extend the spaces $W^{k,p}(\mathbb{R}^n)$ to the case where the number k is real.

Definition 6.3.4.

Let $s \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$. We write

$$\|f\|_{\dot{H}^s_p} = \|I^s f\|_p, \quad \|f\|_{H^s_p} = \|J^s f\|_p.$$

Then, the ***homogeneous Sobolev space*** $\dot{H}^s_p(\mathbb{R}^n)$ is defined by

$$\dot{H}^s_p(\mathbb{R}^n) = \left\{ f \in \dot{\mathscr{S}}'(\mathbb{R}^n) : \|f\|_{\dot{H}^s_p} < \infty \right\}, \tag{6.3.6}$$

and the ***non-homogeneous Sobolev space*** $H^s_p(\mathbb{R}^n)$ is defined by

$$H^s_p(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) : \|f\|_{H^s_p} < \infty \right\}. \tag{6.3.7}$$

If $p = 2$, we denote $\dot{H}^s_2(\mathbb{R}^n)$ by $\dot{H}^s(\mathbb{R}^n)$ and $H^s_2(\mathbb{R}^n)$ by $H^s(\mathbb{R}^n)$ for simplicity.

It is clear that the space $H^s_p(\mathbb{R}^n)$ is a normed linear space with the above norm. Moreover, it is complete and therefore Banach space. To prove the completeness, let $\{f_m\}$ be a Cauchy sequence in H^s_p. Then, by the completeness of L^p, there exists a $g \in L^p$ such that

$$\|f_m - J^{-s}g\|_{H^s_p} = \|J^s f_m - g\|_p \to 0, \quad \text{as } m \to \infty.$$

Clearly, $J^{-s}g \in \mathscr{S}'$ and thus H^s_p is complete.

We give some elementary results about Sobolev spaces.

Theorem 6.3.5.

Let $s \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$, then we have
(a) $\mathscr{S}$ is dense in H_p^s, $1 \leqslant p < \infty$.
(b) $H_p^{s+\varepsilon} \hookrightarrow H_p^s$, $\forall \varepsilon > 0$.
(c) $H_p^s \hookrightarrow L^\infty$, $\forall s > n/p$.
(d) Suppose $1 < p < \infty$ and $s \geqslant 1$. Then $f \in H_p^s(\mathbb{R}^n)$ if and only if $f \in H_p^{s-1}(\mathbb{R}^n)$ and for each j, $\frac{\partial f}{\partial x_j} \in H_p^{s-1}(\mathbb{R}^n)$. Moreover, the two norms are equivalent:

$$\|f\|_{H_p^s} \sim \|f\|_{H_p^{s-1}} + \sum_{j=1}^n \left\| \frac{\partial f}{\partial x_j} \right\|_{H_p^{s-1}}.$$

(e) $H_p^k(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$, $1 < p < \infty$, $\forall k \in \mathbb{N}$.

Proof. (a) 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 $J^{-s} g \in \mathscr{S}$, therefore $\mathscr{S}$ is dense in H_p^s.

(b) Suppose that $f \in H_p^{s+\varepsilon}$. By part (a) in Theorem 6.2.3, we see that J_ε maps L^p into L^p with norm 1 for $\varepsilon > 0$. Form this, we get the result since

$$\|f\|_{H_p^s} = \|J^s f\|_p = \|J^{-\varepsilon} J^{s+\varepsilon} f\|_p = \|J_\varepsilon J^{s+\varepsilon} f\|_p \leqslant \|J^{s+\varepsilon} f\|_p = \|f\|_{H_p^{s+\varepsilon}}.$$

(c) By Young's inequality, the definition of the kernel $G_s(x)$ and part (d) of Proposition 6.2.2, we get for $s > 0$

$$\begin{aligned}
\|f\|_\infty =& \|\mathscr{F}^{-1}[(1+|\omega\xi|^2)^{-s/2}(1+|\omega\xi|^2)^{s/2}\hat{f}]\|_\infty \\
=& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\mathscr{F}^{-1}[(1+|\omega\xi|^2)^{-s/2}] * J^s f\|_\infty \\
\leqslant& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\mathscr{F}^{-1}(1+|\omega\xi|^2)^{-s/2}\|_{p'} \|J^s f\|_p \\
=& \|G_s(x)\|_{p'} \|f\|_{H_p^s} \leqslant C\|f\|_{H_p^s}.
\end{aligned}$$

(d) From the Mikhlin multiplier theorem, we can get $(\omega\xi_j)(1+|\omega\xi|^2)^{-1/2} \in \mathcal{M}_p$ for $1 < p < \infty$, and thus

$$\begin{aligned}
\left\| \frac{\partial f}{\partial x_j} \right\|_{H_p^{s-1}} =& \|\mathscr{F}^{-1}[(1+|\omega\xi|^2)^{(s-1)/2}(\omega i\xi_j)\hat{f}]\|_p \\
=& \|\mathscr{F}^{-1}[(1+|\omega\xi|^2)^{-1/2}(\omega\xi_j)(1+|\omega\xi|^2)^{s/2}\hat{f}]\|_p \\
=& \left(\frac{|\omega|}{2\pi}\right)^{n/2} \|\mathscr{F}^{-1}[(1+|\omega\xi|^2)^{-1/2}(\omega\xi_j)] * J^s f\|_p \\
\leqslant& C\|J^s f\|_p = C\|f\|_{H_p^s}.
\end{aligned}$$

Combining with $\|f\|_{H_p^{s-1}} \leqslant \|f\|_{H_p^s}$, we get

$$\|f\|_{H_p^{s-1}} + \sum_{j=1}^n \left\| \frac{\partial f}{\partial x_j} \right\|_{H_p^{s-1}} \leqslant C\|f\|_{H_p^s}.$$

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

$$(1+|\omega\xi|^2)^{1/2}\left(1+\sum_{j=1}^{n}\chi(\xi_j)|\xi_j|\right)^{-1} \in \mathcal{M}_p,\ \chi(\xi_j)|\xi_j|\xi_j^{-1} \in \mathcal{M}_p,\ 1<p<\infty,$$

and then

$$\begin{aligned}\|f\|_{H^s_p} =&\|J^s f\|_p = \|\mathscr{F}^{-1}[(1+|\omega\xi|^2)^{1/2}\widehat{J^{s-1}f}]\|_p\\ \leqslant& C\|\mathscr{F}^{-1}[(1+\sum_{j=1}^{n}\chi(\xi_j)|\xi_j|)\widehat{J^{s-1}f}]\|_p\\ \leqslant& C\|f\|_{H^{s-1}_p} + C\sum_{j=1}^{n}\left\|\mathscr{F}^{-1}(\chi(\xi_j)|\xi_j|\xi_j^{-1}\widehat{J^{s-1}\tfrac{\partial f}{\partial x_j}})\right\|_p\\ \leqslant& C\|f\|_{H^{s-1}_p} + C\sum_{j=1}^{n}\left\|\frac{\partial f}{\partial x_j}\right\|_{H^{s-1}_p}.\end{aligned}$$

Thus, we have obtained the desired result.

(e) It is obvious that $W^{0,p} = H^0_p = L^p$ for $k = 0$. However, from part (d), if $k \geqslant 1$, then $f \in H^k_p$ if and only if f and $\frac{\partial f}{\partial x_j} \in H^{k-1}_p$, $j = 1, ..., n$. Thus, we can extends the identity of $W^{k,p} = H^k_p$ from $k = 0$ to $k = 1, 2, ...$ ■

We continue with the Sobolev embedding theorem.

Theorem 6.3.6: Sobolev embedding theorem

Let $1 < p \leqslant p_1 < \infty$ and $s, s_1 \in \mathbb{R}$. Assume that $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$. Then the following conclusions hold

$$H^s_p \hookrightarrow H^{s_1}_{p_1}, \quad \dot H^s_p \hookrightarrow \dot H^{s_1}_{p_1}.$$

Proof. It is trivial for the case $p = p_1$ since we also have $s = s_1$ in this case. Now, we assume that $p < p_1$. Since $\frac{1}{p_1} = \frac{1}{p} - \frac{s-s_1}{n}$, by part (b) of Theorem 6.2.3, we get

$$\|f\|_{H^{s_1}_{p_1}} = \|J^{s_1} f\|_{p_1} = \|J^{s_1-s}J^s f\|_{p_1} = \|J_{s-s_1}J^s f\|_{p_1} \leqslant C\|J^s f\|_p = C\|f\|_{H^s_p}.$$

Similarly, we can show the homogeneous case. ■

Theorem 6.3.7.

Let $s, \sigma \in \mathbb{R}$ and $1 \leqslant p \leqslant \infty$. Then J^σ is an isomorphism between H^s_p and $H^{s-\sigma}_p$.

Proof. It is clear from the definition. ■

Corollary 6.3.8.

Let $s \in \mathbb{R}$ and $1 \leqslant p < \infty$. Then

$$(H^s_p)' = H^{-s}_{p'}.$$

Proof. It follows from the above theorem and that $(L^p)' = L^{p'}$, if $1 \leqslant p < \infty$. ■

Finally, we give the connection between the homogeneous and the nonhomogeneous spaces, whose proof will be postponed to next section.

Theorem 6.3.9.

Suppose that $f \in \mathscr{S}'(\mathbb{R}^n)$ and $0 \notin \operatorname{supp} \widehat{f}$. Then

$$f \in \dot{H}^s_p \Leftrightarrow f \in H^s_p, \quad \forall s \in \mathbb{R},\ 1 \leqslant p \leqslant \infty.$$

Moreover, for $1 \leqslant p \leqslant \infty$, we have

$$\begin{aligned} H^s_p =& L^p \cap \dot{H}^s_p, \quad \forall s > 0, \\ H^s_p =& L^p + \dot{H}^s_p, \quad \forall s < 0, \\ H^0_p =& L^p = \dot{H}^0_p. \end{aligned}$$

§ 6.4 The smooth dyadic decomposition

For simplicity, let $\omega = 1$ in the definition of the Fourier transform and its inverse, and we will use the following forms of them:

$$\mathscr{F} f(\xi) = \widehat{f}(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) dx, \tag{6.4.1}$$

$$\mathscr{F}^{-1} g(x) = \check{g}(x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} g(\xi) d\xi. \tag{6.4.2}$$

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 do some operations such as differentiation, multiplier and so on. Therefore, we use a smooth form of this decomposition.

Throughout, we shall call a ***ball*** any set $\{\xi \in \mathbb{R}^n : |\xi| \leqslant R\}$ with $R > 0$ and an ***annulus*** any set $\{\xi \in \mathbb{R}^n : R_1 \leqslant |\xi| \leqslant R_2\}$ with $0 < R_1 < R_2$.

Now, we give the fundamental Bernstein inequalities.

Proposition 6.4.1: 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

$$\forall f \in L^p(\mathbb{R}^n) \text{ with } \operatorname{supp} \widehat{f} \subset \lambda B \quad \Longrightarrow \quad \sup_{|\alpha|=k} \|\partial^\alpha f\|_q \leqslant C^{k+1} \lambda^{k+n(\frac{1}{p}-\frac{1}{q})} \|f\|_p,$$

$$\forall f \in L^p(\mathbb{R}^n) \text{ with } \operatorname{supp} \widehat{f} \subset \lambda \mathbb{A} \quad \Longrightarrow$$

$$C^{-k-1} \lambda^k \|f\|_p \leqslant \sup_{|\alpha|=k} \|\partial^\alpha f\|_p \leqslant C^{k+1} \lambda^k \|f\|_p.$$

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 2.4.20. Then, it follows from Theorem 2.4.27 that $\widehat{\widehat{f}} \in C^\infty$ which implies that f coincides with a C^∞ 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 $\widehat{f}(\xi) = \phi_\lambda(\xi)\widehat{f}(\xi)$ point-wisely, we have

$$\partial^\alpha f = \partial^\alpha g_\lambda * f \quad \text{with} \quad g_\lambda = (2\pi)^{-n/2} \check{\phi_\lambda}.$$

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 get

$$\|\partial^\alpha f\|_q = \|\partial^\alpha g_\lambda * f\|_q \leqslant \|\partial^\alpha g_\lambda\|_r \|f\|_p$$

$$
\begin{aligned}
&=\lambda^{n+k}\|(\partial^\alpha g)(\lambda x)\|_r\|f\|_p=\lambda^{k+n/r'}\|\partial^\alpha g\|_r\|f\|_p\\
&=\lambda^{k+n(\frac{1}{p}-\frac{1}{q})}\|\partial^\alpha g\|_r\|f\|_p.
\end{aligned}
$$

The first assertion follows via

$$
\begin{aligned}
(2\pi)^{n/2}\|\partial^\alpha g\|_r \leqslant&\|\partial^\alpha g\|_\infty+\|\partial^\alpha g\|_1\\
\leqslant&\|\partial^\alpha g\|_\infty+\int_{\mathbb{R}^n}|\partial^\alpha g|(1+|x|^2)^n\frac{1}{(1+|x|^2)^n}dx\\
\leqslant&\|\partial^\alpha g\|_\infty+\|(1+|x|^2)^n\partial^\alpha g\|_\infty\int_{\mathbb{R}^n}\frac{1}{(1+|x|^2)^n}dx\\
\leqslant&C_n\|(1+|x|^2)^n\partial^\alpha g\|_\infty=C_n\|\mathscr{F}^{-1}\mathscr{F}((1+|x|^2)^n\partial^\alpha g)\|_\infty\\
\leqslant&C_n\|\mathscr{F}((1+|x|^2)^n\partial^\alpha g)\|_1=C_n\|(1-\Delta)^n((i\xi)^\alpha\phi(\xi))\|_1\\
=&C_n\left\|\sum_{j=0}^n C_n^j(-1)^j\Delta^j(\xi^\alpha\phi(\xi))\right\|_1\leqslant C_n\sum_{j=0}^n C_n^j\|\Delta^j(\xi^\alpha\phi(\xi))\|_1\\
\leqslant&C_n\sup_{0\leqslant|\beta|\leqslant|\alpha|,\,0\leqslant|\sigma|\leqslant 2n-|\beta|}\|\partial^\beta(\xi^\alpha)\partial^\sigma\phi\|_1\\
\leqslant&C_n\sup_{0\leqslant|\beta|\leqslant|\alpha|,\,0\leqslant|\sigma|\leqslant 2n-|\beta|}\|\xi^\beta\partial^\sigma\phi\|_1\\
\leqslant&C_nC^k\sup_{0\leqslant|\sigma|\leqslant 2n}\|\partial^\sigma\phi\|_1\quad\text{(since ϕ is compactly supported)}\\
\leqslant&C_n^{k+1}.
\end{aligned}
$$

To prove the second assertion, we consider a function $\widetilde{\phi}\in\mathscr{D}(\mathbb{R}^n\setminus\{0\})$ 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 $\widehat{f}=\widetilde{\phi}\widehat{f}$, we deduce that there exists a family of integers $(a_\alpha)_{\alpha\in\mathbb{N}_0^n}$ such that

$$
f=\sum_{|\alpha|=k}h_\alpha*\partial^\alpha f,\quad\text{with}\quad h_\alpha:=(2\pi)^{-n/2}a_\alpha\mathscr{F}^{-1}\left((-i\xi)^\alpha|\xi|^{-2k}\widetilde{\phi}(\xi)\right)\in\mathscr{S}\hookrightarrow L^1.
$$

For $\lambda>0$, we have

$$
\widehat{f}(\xi)=\sum_{|\alpha|=k}a_\alpha\frac{(-i\xi)^\alpha}{|\xi|^{2k}}\widetilde{\phi}(\xi/\lambda)(i\xi)^\alpha\widehat{f}(\xi)=\lambda^{-k}\sum_{|\alpha|=k}a_\alpha\frac{(-i\xi/\lambda)^\alpha}{|\xi/\lambda|^{2k}}\widetilde{\phi}(\xi/\lambda)(i\xi)^\alpha\widehat{f}(\xi),
$$

which implies that

$$
f=\lambda^{-k}\sum_{|\alpha|=k}\lambda^n h_\alpha(\lambda\cdot)*\partial^\alpha f.
$$

Then by Young's inequality we get

$$
\|f\|_p\leqslant\lambda^{-k}\sum_{|\alpha|=k}\|h_\alpha\|_1\|\partial^\alpha f\|_p\leqslant C^{k+1}\lambda^{-k}\sum_{|\alpha|=k}\|\partial^\alpha f\|_p,
$$

and the result follows from the first inequality. ■

Remark 6.4.2. 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 6.4.3.

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, independent of λ, such that for any $p \in [1, \infty]$ and any $\lambda > 0$, we have, for any function $f \in L^p$ with $\operatorname{supp} \widehat{f} \subset \lambda\mathbb{A}$,

$$\|\sigma(D)f\|_p \leqslant C\lambda^m \|f\|_p, \quad \text{with } \sigma(D)f := \mathscr{F}^{-1}(\sigma\widehat{f}).$$

Proof. Consider a smooth function θ supported in an annulus and such that $\theta = 1$ on $\mathbb{A}$. It is clear that we have

$$\sigma(D)f = (2\pi)^{-n/2}\check{\sigma} * f = (2\pi)^{-n/2}\mathscr{F}^{-1}(\theta(\xi/\lambda)\sigma(\xi)) * f.$$

Thus, we only need to prove $\theta(\xi)\sigma(\lambda\xi) \in \mathcal{M}_p(\mathbb{R}^n)$, or equivalently, $\theta(\xi)\sigma(\lambda\xi) \in \mathcal{M}_p(\mathbb{R}^n)$. We can use the Bernstein multiplier theorem (i.e., Theorem 2.6.5) to prove it. In fact, we have

$$\|\theta(\xi)\sigma(\lambda\xi)\|_2 \leqslant C_0 \|\theta(\xi)|\lambda\xi|^m\|_2 = C_0\lambda^m \|\theta(\xi)|\xi|^m\|_2 \leqslant C\lambda^m,$$

and by Leibniz's rule

$$\begin{aligned}
\|\partial_{\xi_j}^k(\theta(\xi)\sigma(\lambda\xi))\|_2 \leqslant & \sum_{\ell=0}^k C_k^\ell \|\partial_{\xi_j}^{k-\ell}\theta(\xi)\lambda^\ell(\partial_{\xi_j}^\ell\sigma)(\lambda\xi)\|_2 \\
\leqslant & C_k\lambda^\ell \sum_{\ell=0}^k \|\partial_{\xi_j}^{k-\ell}\theta(\xi)|\lambda\xi|^{m-\ell}\|_2 \\
\leqslant & C_k\lambda^m \sum_{\ell=0}^k \|\partial_{\xi_j}^{k-\ell}\theta(\xi)|\xi|^{m-\ell}\|_2 \\
\leqslant & C\lambda^m.
\end{aligned}$$

Thus, we have $\|\theta(\xi)\sigma(\lambda\xi)\|_{\mathcal{M}_p} \leqslant C\lambda^m$ by the Bernstein multiplier theorem for any $p \in [1, \infty]$. Then, we obtain 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.,

$$\psi(\xi) = \begin{cases} 1, & |\xi| \leqslant \alpha^{-1}, \\ \text{smooth}, & \alpha^{-1} < |\xi| < \alpha, \\ 0, & |\xi| \geqslant \alpha. \end{cases} \tag{6.4.3}$$

Let $\varphi(\xi)$ be the function

$$\varphi(\xi) := \psi(\xi/2) - \psi(\xi). \tag{6.4.4}$$

Thus, φ is a bump function supported on the annulus

$$\mathbb{A} = \left\{\xi : \alpha^{-1} \leqslant |\xi| \leqslant 2\alpha\right\}. \tag{6.4.5}$$

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

$$\begin{cases} \psi_k(\xi) = \psi(2^{-k}\xi), & k \in \mathbb{Z}, \\ \varphi_k(\xi) = \varphi(2^{-k}\xi) = \psi_{k+1}(\xi) - \psi_k(\xi), & k \in \mathbb{Z}. \end{cases} \tag{6.4.6}$$

Since $\operatorname{supp}\varphi \subset \mathbb{A}$, we have

$$\begin{aligned} &\operatorname{supp}\varphi_k \subset 2^k\mathbb{A} := \left\{\xi : 2^k\alpha^{-1} \leqslant |\xi| \leqslant 2^{k+1}\alpha\right\}, \quad k \in \mathbb{Z}, \\ &\operatorname{supp}\psi_k \subset \left\{\xi : |\xi| \leqslant 2^k\alpha\right\}, \quad k \in \mathbb{Z}. \end{aligned} \tag{6.4.7}$$

We now define the ***k-th homogeneous dyadic blocks*** $\dot{\Delta}_k$ and the ***homogeneous low-frequency cut-off operators*** $\dot{S}_k$ by

$$\dot{\Delta}_k f = \mathscr{F}^{-1}(\varphi_k\widehat{f}), \quad \dot{S}_k f = \mathscr{F}^{-1}(\psi_k\widehat{f}) = \sum_{j \leqslant k-1} \dot{\Delta}_j f, \quad k \in \mathbb{Z}. \tag{6.4.8}$$

Informally, $\dot{\Delta}_k$ is a frequency projection to the annulus $\{\xi : 2^k\alpha^{-1} \leqslant |\xi| \leqslant 2^{k+1}\alpha\}$, while $\dot{S}_k$ is a frequency projection to the ball $\{\xi : |\xi| \leqslant 2^k\alpha\}$. The non-homogeneous 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 non-homogeneous 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 (6.4.6). Also, if f is an L^2 function, then $\dot{S}_k f \to 0$ in L^2 as $k \to -\infty$, and $\dot{S}_k f \to f$ in L^2 as $k \to +\infty$ (this is an easy consequence of Parseval's theorem). By telescoping the series, we thus can write the following (formal) Littlewood-Paley (or dyadic) decomposition

$$\mathrm{Id} = \sum_{k \in \mathbb{Z}} \dot{\Delta}_k \quad \text{and} \quad \mathrm{Id} = \sum_{k \in \mathbb{Z}} \Delta_k. \tag{6.4.9}$$

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 frequency of magnitude roughly 2^k. Lower values of k represent low frequency components of f; higher values represent high frequency components.

Both decompositions have advantages and drawbacks. The non-homogeneous 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 non-homogeneous cases, the above decomposition makes sense in $\mathscr{S}'(\mathbb{R}^n)$.

Proposition 6.4.4.

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 $\psi(2^{-k}\cdot)\hat{g}$ tends to $\hat{g}$ in $\mathscr{S}(\mathbb{R}^n)$. This can easily be verified, so we left it to the interested reader. ■

We now state another result of convergence.

Proposition 6.4.5.

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 $N \in \mathbb{N}$

$$\|u_j\|_\infty \leqslant C2^{jN}, \quad \forall j \in \mathbb{N}, \tag{6.4.10}$$

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 $\widetilde{\mathbb{A}}$ 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,

$$u_0 = \sum_{|\alpha|=k} g_\alpha * \partial^\alpha u_0, \quad g_\alpha = (2\pi)^{-n/2}a_\alpha\mathscr{F}^{-1}\left[\frac{(-i\xi)^\alpha}{|\xi|^{2k}}\phi(\xi)\right].$$

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,

$$u_j = 2^{-jk}\sum_{|\alpha|=k} 2^{jn}g_\alpha(2^j\cdot) * \partial^\alpha u_j. \tag{6.4.11}$$

For any $f \in \mathscr{S}$, we get

$$\begin{aligned}|\langle u_j, f\rangle| =& 2^{-jk}\left|\sum_{|\alpha|=k}\langle u_j, 2^{jn}g_\alpha(-2^j\cdot) * (-\partial)^\alpha f\rangle\right| \\ \leqslant& 2^{-jk}\sum_{|\alpha|=k}\|u_j\|_\infty\|2^{jn}g_\alpha(-2^j\cdot) * \partial^\alpha f\|_1 \\ \leqslant& C2^{-jk}\sum_{|\alpha|=k}2^{jN}\|\partial^\alpha f\|_1.\end{aligned}$$

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 the series converges in $\mathscr{S}'$ by the equivalent conditions of $\mathscr{S}'$. Thus, the convergent series

$$\langle u, f\rangle := \lim_{j\infty}\sum_{j'\leqslant j}\langle u_{j'}, f\rangle$$

defines a tempered distribution. ■

For the operators $\dot\Delta_k$ and $\dot S_k$, we can easily verify the following result:

Proposition 6.4.6.

Let $\alpha \in (1, \sqrt{2})$, $k, l \in \mathbb{Z}$, and $\dot\Delta_k, \dot S_k$ be defined as in (6.4.8). For any $f \in \mathscr{S}'(\mathbb{R}^n)$, we have the following properties:

$$\dot S_k \dot\Delta_{k+l} f \equiv 0, \quad \text{if } l \geqslant 1, \tag{6.4.12}$$

$$\dot\Delta_k \dot\Delta_l f \equiv 0, \quad \text{if } |k-l| \geqslant 2. \tag{6.4.13}$$

Remark 6.4.7. In these properties, we need the condition $\alpha^2 < 2$ which is the reason that we requires $\alpha < \sqrt{2}$ in the beginning of the section.

When dealing with the Littlewood-Paley decomposition, it is convenient to introduce the functions

$$\widetilde\psi(\xi) = \psi(\xi/2), \quad \widetilde\varphi(\xi) = \varphi_{-1}(\xi) + \varphi_0(\xi) + \varphi_1(\xi) = \psi(\xi/4) - \psi(2\xi).$$

as well as the operators

$$\widetilde S_k = \mathscr{F}^{-1}\widetilde\psi(2^{-k}\xi)\mathscr{F} = \dot S_{k+1}, \quad \widetilde\triangle_k = \mathscr{F}^{-1}\widetilde\varphi(2^{-k}\xi)\mathscr{F}.$$

It is clear that $\dot S_k = \widetilde S_k \dot S_k$, and $\dot\Delta_k = \widetilde\triangle_k \dot\Delta_k$ from Proposition 6.4.6.

By the Bernstein multiplier theorem, we can easily prove the following crucial properties of the operators $\dot\Delta_k$ and $\dot S_k$:

Proposition 6.4.8: Boundedness of the operators

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 $\dot\Delta_k f$ relate to the L^p behavior of f. One can already see this when $p = 2$, in which case we have

$$\|f\|_2 \sim \left(\sum_{k\in\mathbb{Z}} \|\dot\Delta_k f\|_2^2\right)^{1/2}. \tag{6.4.14}$$

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 $\varphi_k(\xi)$ which does not vanish. That is, for $\xi \in \operatorname{supp}\varphi_\ell$,

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} |\varphi_k(\xi)|^2 =& \varphi_{\ell-1}^2(\xi) + \varphi_\ell^2(\xi) + \varphi_{\ell+1}^2(\xi) \\
=& (\varphi_{\ell-1}(\xi) + \varphi_\ell(\xi) + \varphi_{\ell+1}(\xi))^2 \\
& - 2(\varphi_{\ell-1}(\xi)\varphi_\ell(\xi) + \varphi_{\ell-1}(\xi)\varphi_{\ell+1}(\xi) + \varphi_\ell(\xi)\varphi_{\ell+1}(\xi)) \\
=& 1 - 2(\varphi_{\ell-1}(\xi) + \varphi_{\ell+1}(\xi))\varphi_\ell(\xi)
\end{aligned}$$

$$=1-2(1-\varphi_\ell(\xi))\varphi_\ell(\xi)$$
$$=1-2\varphi_\ell(\xi)+2\varphi_\ell^2(\xi)$$
$$=\frac{1}{2}+2\left(\frac{1}{2}-\varphi_\ell(\xi)\right)^2,$$

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 (6.4.14) is

$$\|f\|_2\sim\left\|\left(\sum_{k\in\mathbb{Z}}|\dot{\Delta}_k f|^2\right)^{1/2}\right\|_2. \tag{6.4.15}$$

The quantity $\left(\sum_{k\in\mathbb{Z}}|\dot{\Delta}_k f|^2\right)^{1/2}$ is also known as the ***Littlewood-Paley square function***. More generally, the Littlewood-Paley square function theorem is valid for this smooth type decomposition:

Theorem 6.4.9: Littlewood-Paley square function theorem

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.

The proof of this theorem is very similar to that of Theorem 5.4.1, so we remain it to the interested reader.

Now, we can give the proof of Theorem 6.3.9.

Proof of Theorem 6.3.9. Since $0\notin\operatorname{supp}\widehat{f}$, we have $\widehat{f}(\xi)=0$ in a neighborhood of $\xi=0$. Then there is some integer k_0 such that $f=\sum_{k\geqslant k_0}\dot{\Delta}_k f$. Noting that

$$(1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\sum_{k\geqslant k_0}\varphi_k(\xi)\in\mathcal{M}_p$$

by the Bernstein multiplier theorem, we see that for $f\in\dot{H}^s_p(\mathbb{R}^n)$

$$\|f\|_{H^s_p}=\left\|\mathscr{F}^{-1}\left((1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\sum_{k\geqslant k_0}\varphi_k(\xi)\widehat{I^s f}\right)\right\|_p\leqslant C\|f\|_{\dot{H}^s_p}.$$

Conversely, if $f\in H^s_p$, then we note that $|\omega\xi|^s(1+|\omega\xi|^2)^{-s/2}\sum_{k\geqslant k_0}\varphi_k(\xi)\in\mathcal{M}_p$ in view of Bernstein multiplier theorem. Thus,

$$\|f\|_{\dot{H}^s_p}=\left\|\mathscr{F}^{-1}\left(|\omega\xi|^s(1+|\omega\xi|^2)^{-s/2}\sum_{k\geqslant k_0}\varphi_k(\xi)\widehat{J^s f}\right)\right\|_p\leqslant C\|f\|_{H^s_p}.$$

We consider the case $s>0$. If $f\in L^p\cap\dot{H}^s_p$, then we obtain as above

$$\|f\|_{H^s_p}\leqslant\left\|\mathscr{F}^{-1}\left((1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\sum_{k\geqslant 0}\varphi_k(\xi)\widehat{I^s f}\right)\right\|_p$$

$$+\left\|\mathscr{F}^{-1}\left((1+|\omega\xi|^2)^{s/2}\sum_{k<0}\varphi_k(\xi)\widehat{f}\right)\right\|_p$$
$$\leqslant C(\|f\|_{\dot{H}_p^s}+\|f\|_p).$$

Conversely, if $f\in H_p^s$, then clearly $f\in L^p$ and

$$\|f\|_{\dot{H}_p^s}\leqslant\left\|\mathscr{F}^{-1}\left(|\omega\xi|^s(1+|\omega\xi|^2)^{-s/2}\sum_{k\geqslant 0}\varphi_k(\xi)\widehat{J^sf}\right)\right\|_p$$
$$+|\omega|^s\sum_{k<0}2^{ks}\left\|\mathscr{F}^{-1}\left((2^{-k}|\xi|)^s\varphi(2^{-k}\xi)\widehat{f}\right)\right\|_p$$
$$\leqslant C(\|f\|_{H_p^s}+\|f\|_p)\leqslant C\|f\|_{H_p^s}.$$

Now, we consider the case $s<0$. If $f\in L^p+\dot{H}_p^s$, i.e., $f=f_1+f_2$ for some $f_1\in L^p$ and $f_2\in\dot{H}_p^s$ with $0\notin\operatorname{supp}\widehat{f_2}$, then

$$\|f\|_{H_p^s}\leqslant\|f_1\|_{H_p^s}+\|f_2\|_{H_p^s}$$
$$=\|J^sf_1\|_p+\left\|\mathscr{F}^{-1}\left((1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\sum_{k\geqslant k_0}\varphi_k(\xi)\widehat{I_sf_2}\right)\right\|_p$$
$$\leqslant\|f_1\|_p+C\|f_2\|_{\dot{H}_p^s},$$

by Theorem 6.2.3 and the fact that $(1+|\omega\xi|^2)^{s/2}|\omega\xi|^{-s}\sum_{k\geqslant k_0}\varphi_k(\xi)\in\mathcal{M}_p$ for $s<0$ by the Bernstein multiplier theorem. Conversely, if $f\in H_p^s$, then $f=\sum_{k<0}\dot{\Delta}_kf+\sum_{k\geqslant 0}\dot{\Delta}_kf$ where $\|\sum_{k<0}\dot{\Delta}_kf\|_p\leqslant\|f\|_p$ and $\|\sum_{k\geqslant 0}\dot{\Delta}_kf\|_{\dot{H}_p^s}\leqslant\|f\|_{H_p^s}$ by the first conclusion since $0\notin\operatorname{supp}\mathscr{F}(\sum_{k\geqslant 0}\dot{\Delta}_kf)$.

For the case $s=0$, it is obviously from the definitions. ■

§ 6.5 Besov spaces and Triebel-Lizorkin spaces

The Littlewood-Paley decomposition is very useful. For example, we can define (independently of the choice of the initial function ψ) the following notations.

Definition 6.5.1.

Let $s\in\mathbb{R}$, $1\leqslant p,\ r\leqslant\infty$. For $f\in\mathscr{S}'(\mathbb{R}^n)$, we write

$$\|f\|_{\dot{B}^s_{p,r}}=\left(\sum_{k=-\infty}^{\infty}\left(2^{sk}\|\dot{\Delta}_kf\|_p\right)^r\right)^{\frac{1}{r}},\tag{6.5.1}$$

$$\|f\|_{B^s_{p,r}}=\|S_0f\|_p+\left(\sum_{k=0}^{\infty}\left(2^{sk}\|\Delta_kf\|_p\right)^r\right)^{\frac{1}{r}}.\tag{6.5.2}$$

Observe that (6.5.1) does not satisfy the condition of the norms, since we have $\dot{\Delta}_kP(x)=0$ in $\mathscr{S}'$ for any $P\in\mathscr{P}$. In fact,

$$\dot{\Delta}_kP(x)=0\text{ in }\mathscr{S}'\iff\langle\dot{\Delta}_kP,g\rangle=0,\ \forall g\in\mathscr{S}.$$

It follows from $0\notin\operatorname{supp}\varphi_k$ for any $k\in\mathbb{Z}$ that for any $\alpha\in\mathbb{N}_0^n$

$$\int_{\mathbb{R}^n}x^\alpha\dot{\Delta}_kg(x)dx=\int_{\mathbb{R}^n}x^\alpha\widecheck{\widehat{\dot{\Delta}_kg}}(x)dx=\int_{\mathbb{R}^n}e^{-ix\cdot 0}i^{|\alpha|}\widecheck{\partial_\xi^\alpha\widehat{\dot{\Delta}_kg}}(x)dx$$

$$=(2\pi)^{n/2} i^{|\alpha|} \left[\partial_\xi^\alpha \widehat{\dot{\Delta}_k g}\right](0) = (2\pi)^{n/2} i^{|\alpha|} \left(\frac{i}{\omega}\right)^{|\alpha|} \left[\partial_\xi^\alpha (\varphi_k \hat{g})\right](0) = 0.$$

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 6.5.2.

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 non-homogeneous 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 6.5.3.

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 non-homogeneous 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 6.5.4. It is easy to see that the above quantities define a quasi-norm and a norm in general, with the usual convention that $r = \infty$ in both cases corresponds to the usual L^∞ norm. On the other hand, we have not included the case $r = \infty$ in the definition of Triebel-Lizorkin space because the L^∞ norm has to be replaced here by a more complicated Carleson measure.

Besov space and Triebel-Lizorkin space were constructed between 1960's and 1980's. Recently, they are widely applied to study PDEs. Roughly speaking, these spaces are products of the function spaces $\ell^r(L^p)$ or $L^p(\ell^r)$ by combining the Littlewood-Paley decomposition of phase space. The index s in the definition, describes the regularity of the space.

From Theorem 6.4.9, we immediately have the following relations involving Sobolev spaces and Triebel-Lizorkin spaces:

Theorem 6.5.5.

Let $s \in \mathbb{R}$ and $1 < p < \infty$. Then

$$H_p^s = F_{p,2}^s, \quad \dot{H}_p^s = \dot{F}_{p,2}^s, \tag{6.5.3}$$

with equivalent norms.

For simplicity, we use X to denote B or F in the spaces, that is, $X_{p,r}^s$ ($\dot{X}_{p,r}^s$, resp.) denotes $B_{p,r}^s$ ($\dot{B}_{p,r}^s$, resp.) or $F_{p,r}^s$ ($\dot{F}_{p,r}^s$, resp.). But it will denote only one of them in the same formula. We always assume that $1 \leqslant p \leqslant \infty$ for $B_{p,r}^s$ ($\dot{B}_{p,r}^s$, resp.) and $1 \leqslant p < \infty$ for $F_{p,r}^s$ ($\dot{F}_{p,r}^s$, resp.) if no other statement is declared. We have the following embedding relations:

Theorem 6.5.6.

Let X denote B or F. Then, we have the following embedding:

$$\begin{aligned} X_{p,r_1}^s \hookrightarrow X_{p,r_2}^s, \quad \dot{X}_{p,r_1}^s &\hookrightarrow \dot{X}_{p,r_2}^s, && \text{if } r_1 \leqslant r_2, \\ X_{p,r_1}^{s+\varepsilon} &\hookrightarrow X_{p,r_2}^s, && \text{if } \varepsilon > 0, \\ B_{p,\min(p,r)}^s \hookrightarrow F_{p,r}^s &\hookrightarrow B_{p,\max(p,r)}^s, && \text{if } 1 \leqslant p < \infty, \\ \dot{B}_{p,\min(p,r)}^s \hookrightarrow \dot{F}_{p,r}^s &\hookrightarrow \dot{B}_{p,\max(p,r)}^s, && \text{if } 1 \leqslant p < \infty. \end{aligned}$$

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

$$\left(\sum_{k=0}^{\infty} 2^{skr_2}|a_k|^{r_2}\right)^{\frac{1}{r_2}} \leqslant \sup_{k\geqslant 0} 2^{(s+\varepsilon)k}|a_k| \left(\sum_{k=0}^{\infty} 2^{-\varepsilon kr_2}\right)^{\frac{1}{r_2}} \lesssim \sup_{k\geqslant 0} 2^{(s+\varepsilon)k}|a_k|.$$

Taking $a_k = \|\Delta_k f\|_p$ or $a_k = |\Delta_k f|$, we can get

$$X_{p,\infty}^{s+\varepsilon} \hookrightarrow X_{p,r_2}^s,$$

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, we get

$$\left\|\left(\sum_{k=j}^{\infty} b_k^r\right)^{\frac{1}{r}}\right\|_p^r = \left\|\sum_{k=j}^{\infty} b_k^r\right\|_{\frac{p}{r}} \leqslant \sum_{k=j}^{\infty} \|b_k^r\|_{\frac{p}{r}} = \sum_{k=j}^{\infty} \|b_k\|_p^r,$$

which yields the first parts of embedding relations.

Case II: $p < r$. By Minkowski's inequality, we have

$$\sum_{k=j}^{\infty} \|b_k\|_p^r = \sum_{k=j}^{\infty} \|b_k^r\|_{\frac{p}{r}} \leqslant \left\|\sum_{k=j}^{\infty} b_k^r\right\|_{\frac{p}{r}} = \left\|\left(\sum_{k=j}^{\infty} b_k^r\right)^{\frac{1}{r}}\right\|_p^r,$$

which yields the second parts of embedding relations.

In this case, we have $\ell^p \hookrightarrow \ell^r$ 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. ■

From Theorems 6.5.5 and 6.5.6, we can get the following corollary.

Corollary 6.5.7.

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 6.5.5 and 6.5.6 except the endpoint cases $p = 1$ or ∞ in ii). For the proof of the endpoint cases, one can see [BL76, Chapter 6]. ■

Theorem 6.5.8.

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 non-homogeneous 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 proved in the final. Let's 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 $\sigma = \max(s, 0)$ and sufficiently large[2] $L \in \mathbb{N}_0$, we have for any $f \in \mathscr{S}$, from Proposition 6.4.1 and 6.4.8, that

$$\begin{aligned}\|f\|_{B^s_{p,\infty}} =& \|S_0 f\|_p + \sup_{k\geqslant 0} 2^{sk} \|\Delta_k f\|_p \\ \leqslant& C\|f\|_p + \sup_{k\geqslant 0} 2^{sk} \|(\sqrt{-\Delta})^{-\sigma} \Delta_k (\sqrt{-\Delta})^{\sigma} f\|_p\end{aligned}$$

[2] 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 $2pL > n$.

$$\lesssim \sum_{\alpha,\beta} |f|_{\alpha,\beta} + \sup_{k\geqslant 0} 2^{sk}2^{-\sigma k}\|(\sqrt{-\Delta})^\sigma f\|_p$$
$$\lesssim \sum_{\alpha,\beta} |f|_{\alpha,\beta} + \|(1+|x|^2)^L(\sqrt{-\Delta})^\sigma f\|_\infty \lesssim \sum_{\alpha,\beta} |f|_{\alpha,\beta}.$$

where $|f|_{\alpha,\beta}$ is one of the semi-norm sequence of $\mathscr{S}$. Thus, we obtain the result.

Step 2: To prove $\mathscr{S} \hookrightarrow X^s_{p,r}$. From Step 1, we know $\mathscr{S} \hookrightarrow B^{s+\varepsilon}_{p,\infty}$ for any $\varepsilon > 0$. From Theorem 6.5.6, we get $B^{s+\varepsilon}_{p,\infty} \hookrightarrow B^s_{p,\min(p,r)} \hookrightarrow B^s_{p,r} \cap F^s_{p,r}$. Therefore, $\mathscr{S} \hookrightarrow X^s_{p,r}$.

Step 3: To prove $B^s_{p,\infty} \hookrightarrow \mathscr{S}'$. For simplicity, we denote $\Delta_{-1} \equiv 0$ temporarily. For any $f \in B^s_{p,\infty}$ and $\alpha \in \mathscr{S}$, we have, from Schwarz' inequality, Proposition 6.4.8 and the result in Step 1, that

$$\begin{aligned}
|\langle f, \alpha\rangle| =& \left| \left\langle (S_0 + \sum_{k=0}^\infty \Delta_k) f, (S_0 + \sum_{l=0}^\infty \Delta_l)\alpha \right\rangle \right| \\
\leqslant& |\langle S_0 f, S_0\alpha\rangle| + |\langle S_0 f, \Delta_0\alpha\rangle| + |\langle \Delta_0 f, S_0\alpha\rangle| \\
&+ \sum_{k=0}^\infty \sum_{l=-1}^1 |\langle \Delta_k f, \Delta_{k+l}\alpha\rangle| \\
\lesssim& \|f\|_p \|\alpha\|_{L^{p'}} + \sum_{k=0}^\infty \sum_{l=-1}^1 \|\Delta_k f\|_p \|\Delta_{k+l}\alpha\|_{L^{p'}} \\
\lesssim& \|f\|_p \|\alpha\|_{L^{p'}} + \sum_{k=0}^\infty \sum_{l=-1}^1 2^{sk}\|\Delta_k f\|_p 2^{-sk}\|\Delta_{k+l}\alpha\|_{L^{p'}} \\
\lesssim& \|f\|_p \|\alpha\|_{L^{p'}} + \sup_{k\geqslant 0} 2^{sk}\|\Delta_k f\|_p \sum_{k=0}^\infty 2^{-sk}\|\Delta_k \alpha\|_{L^{p'}} \\
\lesssim& \|f\|_{B^s_{p,\infty}} \|\alpha\|_{B^{-s+\varepsilon}_{p',\infty}} \\
\lesssim& \|f\|_{B^s_{p,\infty}} p_N(\alpha).
\end{aligned}$$

Here, we can take α over a bounded set B of $\mathscr{S}$, then $p_N(\alpha) \leqslant C$. Thus, we have proved the result.

Step 4: To prove $X^s_{p,r} \hookrightarrow \mathscr{S}'$. From Theorem 6.5.6, we have $X^s_{p,r} \hookrightarrow B^s_{p,\max(p,q)} \hookrightarrow B^s_{p,\infty} \hookrightarrow \mathscr{S}'$.

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

$$\|\Delta_k f_l - g_k\|_p \to 0, \quad l \to \infty. \tag{6.5.4}$$

Since $L^p \hookrightarrow \mathscr{S}'$ and $\Delta_k f_l \to \Delta_k f$ as $l \to \infty$ in $\mathscr{S}'$, we get $g_k = \Delta_k f$. Hence, (6.5.4) 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. ■

§ 6.6 Embedding and interpolation of spaces

Theorem 6.6.1: The embedding theorem

Let $1 \leqslant p, p_1, r, r_1 \leqslant \infty$ and $s, s_1 \in \mathbb{R}$. Assume that $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$. Then the following conclusions hold

$$B^s_{p,r} \hookrightarrow B^{s_1}_{p_1,r_1}, \quad \dot{B}^s_{p,r} \hookrightarrow \dot{B}^{s_1}_{p_1,r_1}, \quad \forall p \leqslant p_1 \text{ and } r \leqslant r_1;$$
$$F^s_{p,r} \hookrightarrow F^{s_1}_{p_1,r_1}, \quad \dot{F}^s_{p,r} \hookrightarrow \dot{F}^{s_1}_{p_1,r_1}, \quad \forall p < p_1 < \infty.$$

Proof. We only give the proof of the non-homogeneous cases, the homogeneous cases can be treated in a similar way.

Let us prove the first conclusion. From the Bernstein inequality in Proposition 6.4.1, we immediately have

$$\|\Delta_k f\|_{p_1} \lesssim 2^{kn(\frac{1}{p}-\frac{1}{p_1})}\|\Delta_k f\|_p, \quad \|S_0 f\|_{p_1} \lesssim \|S_0 f\|_p, \tag{6.6.1}$$

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 6.5.6, we get

$$\begin{aligned}\|f\|_{B^{s_1}_{p_1,r_1}} =& \|S_0 f\|_{p_1} + \left(\sum_{k=0}^{\infty}\left(2^{s_1 k}\|\Delta_k f\|_{p_1}\right)^{r_1}\right)^{\frac{1}{r_1}}\\ \lesssim& \|S_0 f\|_p + \left(\sum_{k=0}^{\infty}\left(2^{sk}\|\Delta_k f\|_p\right)^{r_1}\right)^{\frac{1}{r_1}} = \|f\|_{B^s_{p,r_1}} \lesssim \|f\|_{B^s_{p,r}}.\end{aligned}$$

This gives the first conclusion.

Next, we prove the second conclusion. In view of Theorem 6.5.6, 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 1.1.4) 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

$$\begin{aligned}\left\|\sum_{k=0}^{\infty} 2^{s_1 k}|\Delta_k f|\right\|_{p_1}^{p_1} =& p_1\int_0^A t^{p_1-1}\left|\left\{x : \sum_{k=0}^{\infty} 2^{s_1 k}|\Delta_k f(x)| > t\right\}\right| dt\\ &+ p_1\int_A^\infty t^{p_1-1}\left|\left\{x : \sum_{k=0}^{\infty} 2^{s_1 k}|\Delta_k f(x)| > t\right\}\right| dt\\ =:& I + II,\end{aligned}$$

where $A \gg 1$ is a constant which can be chosen as below. Noticing that $p < p_1$ and $s - \frac{n}{p} = s_1 - \frac{n}{p_1}$ imply $s > s_1$, we have

$$\sum_{k=K}^{\infty} 2^{s_1 k}|\Delta_k f| \lesssim 2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk}|\Delta_k f|, \quad \forall K \in \mathbb{N}_0. \tag{6.6.2}$$

By taking $K = 0$ and noticing $p < p_1$ (which implies that $t^{p_1-1} \leqslant A^{p_1-p}t^{p-1}$ for $t \leqslant A$), we get

$$\begin{aligned}
I \lesssim & \int_0^A t^{p_1-1}\left|\left\{x : \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > ct\right\}\right| dt \\
\lesssim & \int_0^{cA} \tau^{p-1}\left|\left\{x : \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > \tau\right\}\right| d\tau \lesssim \left\|\sup_{k\geqslant 0} 2^{sk}|\Delta_k f|\right\|_p^p \lesssim 1,
\end{aligned}$$

where the implicit constant depends on A, but it is a fixed constant.

Now we estimate II. By the Bernstein inequality in Proposition 6.4.1, 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

$$\begin{aligned}
\sum_{k=0}^{K-1} 2^{s_1 k}|\Delta_k f| \lesssim & \sum_{k=0}^{K-1} 2^{k(s_1-s+n/p)}\left\|\sup_{k\geqslant 0} 2^{sk}|\Delta_k f|\right\|_p \\
\lesssim & 2^{Kn/p_1}\left\|\sup_{k\geqslant 0} 2^{sk}|\Delta_k f|\right\|_p \lesssim 2^{Kn/p_1}.
\end{aligned} \tag{6.6.3}$$

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 $\sum_{k=0}^{\infty} 2^{s_1 k}|(\Delta_k f)(x)| \geqslant t$, we have, from (6.6.2) and (6.6.3), that

$$C2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk}|\Delta_k f| \geqslant \sum_{k=K}^{\infty} 2^{s_1 k}|\Delta_k f| > t/2. \tag{6.6.4}$$

Hence, from (6.6.3) and (6.6.4), we get

$$\begin{aligned}
II =& p_1\int_A^\infty t^{p_1-1}\left|\left\{x : \sum_{k=0}^{\infty} 2^{s_1 k}|\Delta_k f(x)| > t\right\}\right| dt \\
\lesssim & \int_A^\infty t^{p_1-1}\left|\left\{x : \sum_{k=0}^{K-1} 2^{s_1 k}|\Delta_k f(x)| > t/2\right\}\right| dt \\
& + \int_A^\infty t^{p_1-1}\left|\left\{x : \sum_{k=K}^{\infty} 2^{s_1 k}|\Delta_k f(x)| > t/2\right\}\right| dt \\
\lesssim & \int_A^\infty t^{p_1-1}\left|\left\{x : C2^{Kn/p_1} > t/2\right\}\right| dt \\
& + \int_A^\infty t^{p_1-1}\left|\left\{x : C2^{K(s_1-s)} \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > t/2\right\}\right| dt \\
\lesssim & \int_A^\infty t^{p_1-1}\left|\left\{x : \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > ct^{p_1/p}\right\}\right| dt \\
\lesssim & \int_{A'}^\infty \tau^{p-1}\left|\left\{x : \sup_{k\geqslant 0} 2^{sk}|\Delta_k f(x)| > \tau\right\}\right| d\tau
\end{aligned}$$

$$\lesssim \| \sup_{k\geqslant 0} 2^{sk}|\Delta_k f|\|_p^p \lesssim 1.$$

That is,

$$\left\| \sum_{k=0}^{\infty} 2^{s_1 k}|\Delta_k f| \right\|_{p_1} \lesssim 1.$$

On the other hand, from (6.6.1), we have $\|S_0 f\|_{p_1} \lesssim 1$. Therefore, we have obtained $\|f\|_{B^{s_1}_{p_1,1}} \lesssim 1$ under the assumption $\|f\|_{B^s_{p,\infty}} = 1$. This completes the proof. ■

Theorem 6.6.2.

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 6.5.6, 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 6.6.3.

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:

$$s - \frac{n}{p} = (1-\theta)\left(s_0 - \frac{n}{p_0}\right) + \theta\left(s_1 - \frac{n}{p_1}\right), \tag{6.6.5}$$

$$s \leqslant (1-\theta)s_0 + \theta s_1, \tag{6.6.6}$$

$$\frac{1}{r} \leqslant \frac{1-\theta}{r_0} + \frac{\theta}{r_1}. \tag{6.6.7}$$

Then the fractional GN inequality of the following type

$$\|u\|_{\dot{B}^s_{p,r}} \lesssim \|u\|^{1-\theta}_{\dot{B}^{s_0}_{p_0,r_0}} \|u\|^{\theta}_{\dot{B}^{s_1}_{p_1,r_1}} \tag{6.6.8}$$

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 (6.6.6), we have $s \leqslant s^*$ and $r^* \leqslant r$. Applying the convexity Hölder inequality, we have

$$\|f\|_{\dot{B}^{s^*}_{p^*,r^*}} \leqslant \|f\|^{1-\theta}_{\dot{B}^{s_0}_{p_0,r_0}} \|f\|^{\theta}_{\dot{B}^{s_1}_{p_1,r_1}}. \tag{6.6.9}$$

Using the embedding $\dot{B}^{s^*}_{p^*,r^*} \hookrightarrow \dot{B}^s_{p,r}$, we get the conclusion. ■

For the most general case, we give some fractional Gagliardo-Nirenberg inequalities in homogeneous Besov, Triebel-Lizorkin and Sobolev spaces without proofs (cf. [HMOW11]).

Theorem 6.6.4.

Let $1 \leqslant p, p_0, p_1, r, r_0, r_1 \leqslant \infty$, $s, s_0, s_1 \in \mathbb{R}$, $0 \leqslant \theta \leqslant 1$. Assume that
$$s - \frac{n}{p} = (1-\theta)\left(s_0 - \frac{n}{p_0}\right) + \theta\left(s_1 - \frac{n}{p_1}\right),$$
then the fractional Gagliardo-Nirenberg inequality of the following type
$$\|f\|_{\dot{B}^s_{p,r}} \lesssim \|f\|^{1-\theta}_{\dot{B}^{s_0}_{p_0,r_0}} \|f\|^{\theta}_{\dot{B}^{s_1}_{p_1,r_1}}$$
holds for all $f \in \dot{B}^{s_0}_{p_0,r_0} \cap \dot{B}^{s_1}_{p_1,r_1}$ if and only if one of the following conditions holds:
i) $s \leqslant (1-\theta)s_0 + \theta s_1$ and $\frac{1}{r} \leqslant \frac{1-\theta}{r_0} + \frac{\theta}{r_1}$;
ii) $p_0 = p_1$ and $s = (1-\theta)s_0 + \theta s_1$ but $s_0 \neq s_1$;
iii) $s_0 - \frac{n}{p_0} \neq s - \frac{n}{p}$ and $s < (1-\theta)s_0 + \theta s_1$.

Theorem 6.6.5.

Let $1 \leqslant p, p_i, r < \infty$, $s, s_0, s_1 \in \mathbb{R}$, $0 < \theta < 1$. Then the fractional Gagliardo-Nirenberg inequality of the following type
$$\|f\|_{\dot{F}^s_{p,r}} \lesssim \|f\|^{1-\theta}_{\dot{F}^{s_0}_{p_0,\infty}} \|f\|^{\theta}_{\dot{F}^{s_1}_{p_1,\infty}}$$
holds if and only if
$$s - \frac{n}{p} = (1-\theta)\left(s_0 - \frac{n}{p_0}\right) + \theta\left(s_1 - \frac{n}{p_1}\right),$$
$$s \leqslant (1-\theta)s_0 + \theta s_1,$$
$$s_0 \neq s_1 \text{ if } s = (1-\theta)s_0 + \theta s_1.$$

Corollary 6.6.6.

Let $1 < p, p_0, p_1 < \infty$, $s, \ s_1 \in \mathbb{R}$, $0 \leqslant \theta \leqslant 1$. Then the fractional Gagliardo-Nirenberg inequality of the following type
$$\|f\|_{\dot{H}^s_p} \lesssim \|f\|^{1-\theta}_{L^{p_0}} \|f\|^{\theta}_{\dot{H}^{s_1}_{p_1}}$$
holds if and only if
$$s - \frac{n}{p} = (1-\theta)\frac{n}{p_0} + \theta\left(s_1 - \frac{n}{p_1}\right), \quad s \leqslant \theta s_1.$$

Now, we give the duality theorem:

Theorem 6.6.7: The duality theorem

Let $s \in \mathbb{R}$. Then we have
i) $(B^s_{p,r})' = B^{-s}_{p',r'}$, if $1 \leqslant p, \ r < \infty$.
ii) $(F^s_{p,r})' = F^{-s}_{p',r'}$, if $1 < p, \ r < \infty$.

Proof. Please read [BL76, Tri83] for details. ■

§ 6.7 Differential-difference norm on Besov spaces

The next theorem points to an alternative definition of the Besov spaces $B^s_{p,r}$ ($s > 0$) in terms of derivatives and moduli of continuity. The modulus of continuity is defined by

$$\omega_p^m(t, f) = \sup_{|y|<t} \| \triangle_y^m f\|_p,$$

where $\triangle_y^m$ is the m-th order difference operator:

$$\triangle_y^m f(x) = \sum_{k=0}^{m} C_m^k (-1)^k f(x + ky).$$

Theorem 6.7.1.

Assume that $s > 0$, and let m and N be integers, such that $m + N > s$ and $0 \leqslant N < s$. Then, with $1 \leqslant p, r \leqslant \infty$,

$$\|f\|_{B^s_{p,r}} \sim \|f\|_p + \sum_{j=1}^{n} \left(\int_0^\infty \left(t^{N-s} \omega_p^m \left(t, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \frac{dt}{t} \right)^{1/r}.$$

Proof. We note that ω_p^m is an increasing function of t. Therefore, it is sufficient to prove that

$$\|f\|_{B^s_{p,r}} \sim \|f\|_p + \sum_{j=1}^{n} \left(\sum_{\ell=-\infty}^{\infty} \left(2^{\ell(s-N)} \omega_p^m \left(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N} \right) \right)^r \right)^{1/r}.$$

First, we assume that $f \in B^s_{p,r}$. It is clear that

$$\begin{aligned}
\omega_p^m(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N}) &= \sup_{|y|<2^{-\ell}} \left\| \triangle_y^m \frac{\partial^N f}{\partial x_j^N} \right\|_p = \sup_{|y|<2^{-\ell}} \left\| \sum_{k=0}^{m} C_m^k (-1)^k \frac{\partial^N f}{\partial x_j^N}(x + ky) \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \left(\sum_{k=0}^{m} C_m^k (-1)^k f(x + ky) \right) \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \left(\sum_{k=0}^{m} C_m^k (-1)^k \mathscr{F}^{-1}(e^{iky\cdot\xi} \widehat{f}) \right) \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \mathscr{F}^{-1} \left(\sum_{k=0}^{m} C_m^k (-1)^k e^{iky\cdot\xi} \widehat{f} \right) \right\|_p \\
&= \sup_{|y|<2^{-\ell}} \left\| \frac{\partial^N}{\partial x_j^N} \mathscr{F}^{-1} \left((1 - e^{iy\cdot\xi})^m \widehat{f} \right) \right\|_p.
\end{aligned}$$

Denote $\rho_y(\xi) = (1 - e^{iy\cdot\xi})^m$. By the Littlewood-Paley decomposition and the Bernstein inequalities, we have

$$\begin{aligned}
&\omega_p^m(2^{-\ell}, \frac{\partial^N f}{\partial x_j^N}) \\
= &\sup_{|y|<2^{-\ell}} \left\| \left(S_0 + \sum_{k=0}^{\infty} \Delta_k \right) \frac{\partial^N}{\partial x_j^N} \mathscr{F}^{-1} \left(\rho_y(\xi) \widehat{f} \right) \right\|_p \\
\lesssim &\sup_{|y|<2^{-\ell}} \left\| (2\pi)^{-n/2} \widecheck{\rho_y} * S_0 f \right\|_p + \sup_{|y|<2^{-\ell}} \sum_{k=0}^{\infty} 2^{kN} \left\| (2\pi)^{-n/2} \widecheck{\rho_y} * \Delta_k f \right\|_p.
\end{aligned}$$

Hence, it follows

$$\sum_{j=1}^{n}\left(\sum_{\ell=-\infty}^{\infty}\left(2^{\ell(s-N)}\omega_p^m\left(2^{-\ell},\frac{\partial^N f}{\partial x_j^N}\right)\right)^r\right)^{1/r}$$
$$\lesssim\Big(\sum_{\ell=-\infty}^{\infty}\Big(2^{\ell(s-N)}\sup_{|y|<2^{-\ell}}\|(2\pi)^{-n/2}\widetilde{\rho_y}*S_0f\|_p$$
$$+\sup_{|y|<2^{-\ell}}\sum_{k=0}^{\infty}2^{(\ell-k)(s-N)}2^{ks}\|(2\pi)^{-n/2}\widetilde{\rho_y}*\Delta_kf\|_p\Big)^r\Big)^{1/r}.$$

If we can prove that for all integers k

$$\|(2\pi)^{-n/2}\widetilde{\rho_y}*S_0f\|_p\lesssim\min(1,|y|^m)\|S_0f\|_p, \tag{6.7.1}$$

and

$$\|(2\pi)^{-n/2}\widetilde{\rho_y}*\Delta_kf\|_p\lesssim\min(1,|y|^m2^{mk})\|\Delta_kf\|_p. \tag{6.7.2}$$

Then, we can obtain

$$\sum_{j=1}^{n}\left(\sum_{\ell=-\infty}^{\infty}\left(2^{\ell(s-N)}\omega_p^m\left(2^{-\ell},\frac{\partial^N f}{\partial x_j^N}\right)\right)^r\right)^{1/r}$$
$$\lesssim\Big(\sum_{\ell=-\infty}^{\infty}\Big(2^{\ell(s-N)}\sup_{|y|<2^{-\ell}}\min(1,|y|^m)\|S_0f\|_p$$
$$+\sup_{|y|<2^{-\ell}}\sum_{k=0}^{\infty}2^{(\ell-k)(s-N)}2^{ks}\min(1,|y|^m2^{mk})\|\Delta_kf\|_p\Big)^r\Big)^{1/r}$$
$$\lesssim\Big(\sum_{\ell=-\infty}^{\infty}\Big(2^{\ell(s-N)}\min(1,2^{-\ell m})\|S_0f\|_p$$
$$+\sum_{k=0}^{\infty}2^{(\ell-k)(s-N)}\min(1,2^{-(\ell-k)m})2^{ks}\|\Delta_kf\|_p\Big)^r\Big)^{1/r}$$
$$\lesssim\|(2^{k(s-N)}\min(1,2^{-km}))*(\alpha_k)\|_{\ell^r}$$
$$\lesssim\|(2^{k(s-N)}\min(1,2^{-km}))\|_{\ell^1}\|(\alpha_k)\|_{\ell^r}\lesssim\|f\|_{B^s_{p,r}},$$

where the sequence $(\alpha_k)_{k=-\infty}^{\infty}$ with $\alpha_k=2^{sk}\|\Delta_kf\|_p$ if $k\geqslant 0$, $\alpha_{-1}=\|S_0f\|_p$ and $\alpha_k=0$ if $k<-1$, and we have used the Young inequality for a convolution of two sequences. In addition, we have

$$\|f\|_p\lesssim\|S_0f\|_p+\sum_{k=0}^{\infty}\|\Delta_kf\|_p$$
$$\lesssim\|S_0f\|_p+\left(\sum_{k=0}^{\infty}2^{-skr'}\right)^{1/r'}\left(\sum_{k=0}^{\infty}(2^{sk}\|\Delta_kf\|_p)^r\right)^{1/r}\lesssim\|f\|_{B^s_{p,r}},$$

which implies the desired conclusion.

Now, we turn to prove (6.7.1) and (6.7.2). We only need to show $\rho_y\in\mathcal{M}_p$ and $\rho_y(\cdot)\langle y,\cdot\rangle^{-m}\in\mathcal{M}_p$ for $p\in[1,\infty]$ and

$$\|\rho_y\|_{\mathcal{M}_p}\leqslant C,\quad\|\rho_y(\cdot)\langle y,\cdot\rangle^{-m}\|_{\mathcal{M}_p}\leqslant C,\quad\forall y\neq 0. \tag{6.7.3}$$

In fact, from the definition of ρ_y, we get

$$\|\rho_y\|_{\mathcal{M}_p}=(2\pi)^{-n/2}\sup_{f\in\mathscr{S}}\frac{\|\widetilde{\rho_y}*f\|_p}{\|f\|_p}=\sup_{f\in\mathscr{S}}\frac{\|\sum_{k=0}^m C_m^k(-1)^kf(x+ky)\|_p}{\|f\|_p}$$

$$\leqslant \sum_{k=0}^{m} C_m^k = 2^m.$$

By Theorem 2.6.4, 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 2.6.5), 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 get $|(1-e^{i\eta})/\eta| \leqslant 1$; while for its first order derivative, we can use Taylor's expansion $e^z = \sum\limits_{k=0}^{\infty} \frac{z^k}{k!}$ whenever $|z| < \infty$ ($z \in \mathbb{C}$) to get

$$\begin{aligned}\partial_\eta((1-e^{i\eta})/\eta) =& -\eta^{-2}(i\eta e^{i\eta} + 1 - e^{i\eta})\\ =& -\eta^{-2}\left(i\eta\sum_{k=0}^{\infty}\frac{(i\eta)^k}{k!} - \sum_{k=1}^{\infty}\frac{(i\eta)^k}{k!}\right)\\ =& -\eta^{-2}\left(\sum_{k=0}^{\infty}\frac{(i\eta)^{k+1}}{k!} - \sum_{k=0}^{\infty}\frac{(i\eta)^{k+1}}{(k+1)!}\right)\\ =&\sum_{k=1}^{\infty}\frac{k(i\eta)^{k-1}}{(k+1)!}\\ =&\sum_{k=1}^{\infty}\frac{(i\eta)^{k-1}}{(k-1)!}\frac{1}{(k+1)},\end{aligned}$$

which implies $|\partial_\eta((1-e^{i\eta})/\eta)| \leqslant \frac{1}{2}e^{|\eta|}$. Then it is easy to get the bound of the L^2 integral. Thus, $\|((1-e^{i\eta})/\eta)\|_{\mathcal{M}_p(\mathbb{R})} \leqslant C$ by Theorem 2.6.5, which completes the proof of (6.7.3).

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

$$\|\langle y,\cdot\rangle^m\widetilde{\psi}(\cdot)\|_{\mathcal{M}_p} \leqslant C|y|^m, \quad \|\langle y,\cdot\rangle^m\widetilde{\varphi}(2^{-k}\cdot)\|_{\mathcal{M}_p} \leqslant C|y|^m 2^{mk}.$$

Thus, we get

$$\begin{aligned}&\|(2\pi)^{-n/2}\widecheck{\rho_y} * S_0 f\|_p \lesssim\|S_0 f\|_p,\\ &\|(2\pi)^{-n/2}\widecheck{\rho_y} * S_0 f\|_p =\|(2\pi)^{-n}\mathscr{F}^{-1}(\rho_y(\xi)\langle y,\xi\rangle^{-m}) * \mathscr{F}^{-1}(\langle y,\xi\rangle^m\widetilde{\psi}(\xi)) * S_0 f\|_p\\ &\qquad\lesssim|y|^m\|S_0 f\|_p,\end{aligned}$$

which yields (6.7.1). In the same way, we have

$$\begin{aligned}&\|(2\pi)^{-n/2}\widecheck{\rho_y} * \Delta_k f\|_p \lesssim\|\Delta_k f\|_p,\\ &\|(2\pi)^{-n/2}\widecheck{\rho_y} * \Delta_k f\|_p =\|(2\pi)^{-n}(\mathscr{F}^{-1}\rho_y(\xi)\langle y,\xi\rangle^{-m}) * \mathscr{F}^{-1}(\langle y,\xi\rangle^m\widetilde{\varphi}(2^{-k}\xi)) * \Delta_k f\|_p\\ &\qquad\lesssim|y|^m 2^{mk}\|\Delta_k f\|_p,\end{aligned}$$

which yields (6.7.2).

The converse inequality will follow if we can prove the estimate

$$\|\Delta_k f\|_p \leqslant C2^{-Nk} \sum_{j=1}^{n} \left\| \widecheck{\rho_{jk}} * \frac{\partial^N f}{\partial x_j^N} \right\|_p, \tag{6.7.4}$$

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 (6.7.4) is valid, we have, by noting $\psi \in \mathcal{M}_1$, that

$$\begin{aligned}\|f\|_{B^s_{p,r}} \lesssim &\|f\|_p + \left(\sum_{k=0}^{\infty} \left(2^{k(s-N)} \sum_{j=1}^{n} \left\| \widecheck{\rho_{jk}} * \frac{\partial^N f}{\partial x_j^N} \right\|_p \right)^r \right)^{1/r} \\ \lesssim &\|f\|_p + \sum_{j=1}^{n} \left(\sum_{k=0}^{\infty} \left(2^{k(s-N)} \omega_p^m \left(2^{-k}, \frac{\partial^N f}{\partial x_j^N}\right)\right)^r \right)^{1/r},\end{aligned}$$

which implies the desired inequality.

In order to prove (6.7.4), we need the following lemma.

Lemma 6.7.2.

Assume that $n \geqslant 2$ and take φ as in (6.4.4). Then there exist functions $\chi_j \in \mathscr{S}(\mathbb{R}^n)$ $(1 \leqslant j \leqslant n)$, such that

$$\begin{aligned}&\sum_{j=1}^{n} \chi_j = 1 \quad \text{on } \operatorname{supp} \varphi, \\ &\operatorname{supp} \chi_j \subset \left\{\xi \in \mathbb{R}^n : |\xi_j| \geqslant (3\sqrt{n})^{-1}\right\}, \quad 1 \leqslant j \leqslant n.\end{aligned}$$

Proof. Choose $\kappa \in \mathscr{S}(\mathbb{R})$ with $\operatorname{supp} \kappa = \{\xi \in \mathbb{R} : |\xi| \geqslant (3\sqrt{n})^{-1}\}$ and with positive values in the interior of $\operatorname{supp} \kappa$. Moreover, choose $\sigma \in \mathscr{S}(\mathbb{R}^{n-1})$ with $\operatorname{supp} \sigma = \left\{\xi \in \mathbb{R}^{n-1} : |\xi| \leqslant 3\right\}$ and positive in the interior. Writing

$$\bar{\xi^j} = (\xi_1, \cdots, \xi_{j-1}, \xi_{j+1}, \cdots, \xi_n)$$

and

$$\chi_j(\xi) = \kappa(\xi_j)\sigma(\bar{\xi^j}) / \sum_{j=1}^{n} \kappa(\xi_j)\sigma(\bar{\xi^j}), \quad 1 \leqslant j \leqslant n,$$

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 (6.7.4). By the previous lemma, we obtain the formula

$$\begin{aligned}\|\Delta_k f\|_p \lesssim &\sum_{j=1}^{n} \left\| \mathscr{F}^{-1}(\rho_{jk}^{-m} \chi_j(2^{-k}\cdot)\xi_j^{-N} \varphi(2^{-k}\cdot)) * \mathscr{F}^{-1}(\rho_{jk}^m \widehat{\tfrac{\partial^N f}{\partial x_j^N}}) \right\|_p \\ \lesssim &2^{-kN} \sum_{j=1}^{n} \left\| \mathscr{F}^{-1}(\rho_{jk}^{-m} \chi_j(2^{-k}\cdot)(2^{-k}\xi_j)^{-N} \varphi(2^{-k}\cdot)) * \mathscr{F}^{-1}(\rho_{jk}^m \widehat{\tfrac{\partial^N f}{\partial x_j^N}}) \right\|_p \\ \lesssim &2^{-kN} \sum_{j=1}^{n} \left\| \rho_{j0}^{-m} \chi_j \xi_j^{-N} \varphi \right\|_{\mathcal{M}_p(\mathbb{R}^n)} \left\| \widecheck{\rho_{jk}^m} * \frac{\partial^N f}{\partial x_j^N} \right\|_p \\ \lesssim &2^{-kN} \sum_{j=1}^{n} \left\| \widecheck{\rho_{jk}^m} * \frac{\partial^N f}{\partial x_j^N} \right\|_p,\end{aligned}$$

since, by Theorem 2.6.4 and 2.6.5, 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 which is very convenient for nonlinear estimates in PDEs.

Corollary 6.7.3.

Assume that $s>0$ and $s\notin\mathbb{N}$. Let $1\leqslant p,\ r\leqslant\infty$, then

$$\|f\|_{B^s_{p,r}}\sim\|f\|_p+\sum_{j=1}^n\left(\int_0^\infty\left(t^{[s]-s}\sup_{|h|\leqslant t}\left\|\triangle_h\,\partial^{[s]}_{x_j}f\right\|_p\right)^r\frac{dt}{t}\right)^{1/r},$$

where $[s]$ denotes the integer part of the real number s and $\triangle_h$ denotes the first order difference operator.

Similarly, we can get a equivalent norm for the homogeneous Besov space.

Theorem 6.7.4.

Assume that $s>0$, and let m and N be integers, such that $m+N>s$ and $0\leqslant N<s$. Then, with $1\leqslant p,\ r\leqslant\infty$,

$$\|f\|_{\dot B^s_{p,r}}\sim\sum_{j=1}^n\left(\int_0^\infty\left(t^{N-s}\omega_p^m\left(t,\frac{\partial^N f}{\partial x_j^N}\right)\right)^r\frac{dt}{t}\right)^{1/r}.$$

In particular, if $s>0$ and $s\notin\mathbb{N}$, then

$$\|f\|_{\dot B^s_{p,r}}\sim\sum_{j=1}^n\left(\int_0^\infty\left(t^{[s]-s}\sup_{|h|\leqslant t}\left\|\triangle_h\,\partial^{[s]}_{x_j}f\right\|_p\right)^r\frac{dt}{t}\right)^{1/r},$$

One of the following result is a straightforward consequence of Theorem 6.7.1 and Theorem 6.7.4, which indicates the relation between homogeneous and non-homogeneous spaces.

Theorem 6.7.5.

Suppose that $f\in\mathscr{S}'$ and $0\notin\operatorname{supp}\widehat{f}$. Then

$$f\in B^s_{p,r}\Leftrightarrow f\in\dot B^s_{p,r},\quad\forall s\in\mathbb{R},\ 1\leqslant p,r\leqslant\infty.$$

Moreover,

$$B^s_{p,r}=L^p\cap\dot B^s_{p,r},\quad\forall s>0,\ 1\leqslant p,r\leqslant\infty,$$

$$B^s_{p,r}=L^p+\dot B^s_{p,r},\quad\forall s<0,\ 1\leqslant p,r\leqslant\infty.$$

Proof. One can see [BL76, Chapter 6]. ■

§ 6.8 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 of the value of functions. Therefore, at least we want to go back 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 6.8.1.

Let $1 \leqslant p, r \leqslant \infty$. Assume

$$s < \frac{n}{p}, \quad \text{or} \quad s = \frac{n}{p} \text{ and } r = 1. \tag{6.8.1}$$

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 $\|\dot{\Delta}_k f\|_\infty \leqslant C2^{kn/p}\|\dot{\Delta}_k f\|_p$. It follows that

$$\begin{aligned}\left\|\sum_{k=-\infty}^{0} \dot{\Delta}_k f\right\|_\infty \leqslant & C \sum_{k=-\infty}^{0} \|\dot{\Delta}_k f\|_\infty \leqslant C \sum_{k=-\infty}^{0} 2^{k(n/p-s)}2^{ks}\|\dot{\Delta}_k f\|_p \\ \leqslant & \begin{cases} C\|f\|_{\dot{B}^s_{p,\infty}} \leqslant C\|f\|_{\dot{B}^s_{p,r}}, & \text{if } s < n/p, \\ C\|f\|_{\dot{B}^{n/p}_{p,1}}, & \text{if } s = n/p \text{ and } r = 1. \end{cases}\end{aligned}$$

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 of the regularity index. For convenience, we first define a subspace of $\mathscr{S}'(\mathbb{R}^n)$ which will play an important role to study PDEs.

Definition 6.8.2.

We denote by $\mathscr{S}'_h(\mathbb{R}^n)$ the space of tempered distributions f such that

$$\lim_{\lambda\to\infty} \|\theta(\lambda D)f\|_\infty = 0, \quad \forall\theta \in \mathscr{D}(\mathbb{R}^n), \tag{6.8.2}$$

where the operator $\theta(D)$ is defined by $\theta(D)f := \mathscr{F}^{-1}(\theta\widehat{f})$, for a measurable function f on $\mathbb{R}^n$ with at most polynomial growth at infinity.

Remark 6.8.3. We have the following facts about $\mathscr{S}'_h(\mathbb{R}^n)$.

1) It holds

$$\mathscr{S}'_h(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) : \lim_{k\to-\infty} \dot{S}_k f = 0 \text{ in } \mathscr{S}'(\mathbb{R}^n) \right\}. \tag{6.8.3}$$

Thus, we obtain

$$\mathscr{S}'_h(\mathbb{R}^n) = \left\{ f \in \mathscr{S}'(\mathbb{R}^n) : f = \sum_{k\in\mathbb{Z}} \dot{\Delta}_k f \text{ in } \mathscr{S}'(\mathbb{R}^n) \right\}. \tag{6.8.4}$$

In fact, since $\psi \in \mathscr{D}$, 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 (6.8.2). It also implies $\lim_{k\to-\infty} \dot{S}_k f = 0$ in $\mathscr{S}'$. Conversely, for 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., $k > \log_2 \frac{C}{\lambda\alpha}$. Due to (6.8.4), it holds for any $g \in \mathscr{S}$,

$$|\langle\theta(\lambda D)f, g\rangle| = |\langle\widehat{f}(\xi), \theta(\lambda\xi)\check{g}\rangle|$$

$$
\begin{aligned}
&= \left| \left\langle \sum_{k\in\mathbb{Z}} \varphi_k(\xi)\widehat{f}, \theta(\lambda\xi)\check{g} \right\rangle \right| \\
&= \left| \left\langle \sum_{k\leqslant[\log_2 \frac{C}{\lambda\alpha}]} \varphi_k(\xi)\widehat{f}, \theta(\lambda\xi)\check{g} \right\rangle \right| \\
&= \left| \left\langle \dot{S}_{[\log_2 \frac{C}{\lambda\alpha}]+1} f, \widehat{\theta(\lambda\cdot)} * g \right\rangle \right| \to 0 \text{ as } \lambda\to\infty,
\end{aligned}
$$

by (6.8.3) and the fact that $\|\widehat{\theta(\lambda\cdot)} * g\|_\infty \leqslant \|\widehat{\theta(\lambda\cdot)}\|_1\|g\|_\infty = \|\widehat{\theta}\|_1\|g\|_\infty$ by Young's inequality, i.e., $\widehat{\theta(\lambda\cdot)}*g$ 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 $\lambda\to\infty$. For (6.8.4), noticing $\dot{\Delta}_k = \dot{S}_{k+1} - \dot{S}_k$ and by Proposition 6.4.4 and (6.8.3), 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 6.4.4 and (6.8.4), it follows that (6.8.3).

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 get

$$
\begin{aligned}
|\langle \dot{S}_k f, g\rangle| = |\langle \psi(2^{-k}\xi)\widehat{f}(\xi), \check{g}(\xi)\rangle| &\leqslant \int_{|\xi|\leqslant 2^k\alpha} |\widehat{f}(\xi)||\check{g}(\xi)| d\xi \\
&\leqslant C \int_{|\xi|\leqslant 2^k\alpha} |\widehat{f}(\xi)| d\xi \to 0, \text{ as } k\to-\infty,
\end{aligned}
$$

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

$$
\begin{aligned}
|\langle \dot{S}_k f, g\rangle| &= \left| \left\langle \theta(2^{l-k}\xi)\widehat{f}(\xi), \frac{\psi(2^{-k}\xi)}{\theta(2^{l-k}\xi)}\check{g}(\xi) \right\rangle \right| \\
&\leqslant (2\pi)^{-n/2}\|\theta(2^{l-k}D)f\|_\infty \left\| \mathscr{F}\left(\frac{\psi(2^{-k}\xi)}{\theta(2^{l-k}\xi)} \right) \right\|_1 \|g\|_1 \\
&= (2\pi)^{-n/2}\|\theta(2^{l-k}D)f\|_\infty \left\| \mathscr{F}\left(\frac{\psi_\ell}{\theta} \right) \right\|_1 \|g\|_1 \qquad (6.8.5) \\
&\leqslant C\|\theta(2^{l-k}D)f\|_\infty \to 0, \text{ as } k\to-\infty,
\end{aligned}
$$

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, as similar as in (6.8.5), we can also get for any $k<\ell$

$$
|\langle \dot{S}_k f, g\rangle| = \left| \left\langle \theta(\xi)\widehat{f}(\xi), \frac{\psi(2^{-k}\xi)}{\theta(\xi)}\check{g}(\xi) \right\rangle \right|
$$

$$
\begin{aligned}
&\leqslant (2\pi)^{-n/2}\|\theta(D)f\|_p \left\| \mathscr{F}\left(\frac{\psi(2^{-k}\xi)}{\theta(\xi)}\right)\right\|_{p'} \|g\|_1 \\
&=(2\pi)^{-n/2}\|\theta(D)f\|_p \left\| 2^{kn}\mathscr{F}\left(\frac{\psi(\cdot)}{\theta(2^k\cdot)}\right)(2^k\cdot)\right\|_{p'} \|g\|_1 \\
&=(2\pi)^{-n/2}2^{kn/p}\|\theta(D)f\|_p \left\| \mathscr{F}\left(\frac{\psi(\cdot)}{\theta(2^k\cdot)}\right)\right\|_{p'} \|g\|_1 \to 0, \text{ as } k\to -\infty,
\end{aligned}
$$

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 belongs 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, $\forall\alpha\in\mathbb{N}_0^n$, $\forall g\in\mathscr{S}$,

$$
\begin{aligned}
\langle\theta(\lambda D)x^\alpha, g(x)\rangle =&\langle\theta(\lambda\xi)\widehat{x^\alpha}(\xi), \check{g}(\xi)\rangle = \langle x^\alpha, \widehat{\theta(\lambda\xi)\check{g}(\xi)}\rangle = \langle 1, x^\alpha\widehat{\theta(\lambda\xi)\check{g}(\xi)}\rangle \\
=&\langle 1, \widehat{(-i\partial_\xi)^\alpha(\theta(\lambda\xi)\check{g}(\xi))}\rangle \\
=&\left\langle (2\pi)^{n/2}\delta_0(\xi), \sum_{\alpha=\beta+\gamma} C_\alpha^\beta(-i\lambda)^\beta(\partial_\xi^\beta\theta)(\lambda\xi)(-i\partial_\xi)^\gamma\check{g}(\xi)\right\rangle \\
=&(2\pi)^{n/2}\sum_{\alpha=\beta+\gamma} C_\alpha^\beta(-i\lambda)^\beta(\partial_\xi^\beta\theta)(0)\widetilde{(x^\gamma g)}(0) \\
=&(2\pi)^{n/2}\widetilde{(x^\alpha g)}(0) = \langle(2\pi)^{n/2}\delta_0, \widetilde{(x^\alpha g)}\rangle = \langle 1, x^\alpha g\rangle = \langle x^\alpha, g(x)\rangle,
\end{aligned}
$$

since $(\partial^\beta\theta)(0)=0$ for any $\beta\neq 0$.

7) A non-zero constant function f does not belong to $\mathscr{S}'_h$ because $\dot{S}_k f=f$, $\forall k\in\mathbb{Z}$, i.e., $\lim_{k\to-\infty}\dot{S}_k f\neq 0$. 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 which 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 which can be used in the context of PDEs.

Definition 6.8.4: 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 6.8.5.

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 semi-norm. 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 6.8.6. 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,

$$\begin{aligned}2^{ks}\|\tilde{\varphi}(2^{-k}D)f\|_p =& 2^{ks}\left\|\sum_{|k-k'|\leqslant N_0} \tilde{\varphi}(2^{-k}D)\dot{\Delta}_{k'}f\right\|_p \\ \leqslant& C2^{N_0|s|}\sum_{k'}\chi_{[-N_0,N_0]}(k-k')2^{k's}\|\dot{\Delta}_{k'}f\|_p,\end{aligned}$$

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 6.8.7.

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 if and only if $\|f_N\|_{\dot{\mathcal{B}}^s_{p,r}}$ 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 get

$$\begin{aligned}\dot{\Delta}_k f_N(x) =& \mathscr{F}^{-1}(\varphi(2^{-k}\xi)\widehat{f(2^N x)}(\xi))(x) \\ =& \mathscr{F}^{-1}(\varphi(2^{-k}\xi)2^{-nN}\widehat{f}(2^{-N}\xi))(x) \\ =& \mathscr{F}^{-1}(\varphi(2^{-(k-N)}\xi)\widehat{f}(\xi))(2^N x) = \dot{\Delta}_{k-N}f(2^N x).\end{aligned}$$

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., 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 6.8.8.

Let $\sigma \in (0,n)$. Then for any $p\in[1,\infty]$, it holds

$$\frac{1}{|x|^\sigma} \in \dot{\mathcal{B}}^{\frac{n}{p}-\sigma}_{p,\infty}(\mathbb{R}^n). \tag{6.8.6}$$

Proof. By Proposition 6.6.1, it is enough to prove that $\rho_\sigma := |\cdot|^{-\sigma} \in \dot{\mathcal{B}}^{n-\sigma}_{1,\infty}$. In order 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

$$\dot{\Delta}_k\rho_\sigma = (2\pi)^{-n/2}\widecheck{\varphi(2^{-k}\cdot)} * \rho_\sigma = (2\pi)^{-n/2}2^{kn}\check{\varphi}(2^k\cdot) * \rho_\sigma$$

$$=(2\pi)^{-n/2}2^{k(n+\sigma)}\check{\varphi}(2^k\cdot)*\rho_\sigma(2^k\cdot)=2^{k\sigma}(\dot{\Delta}_0\rho_\sigma)(2^k\cdot).$$

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 get

$$\|\dot{\Delta}_0\rho_1\|_1\leqslant C_k\sup_{|\alpha|=k}\|\partial^\alpha\dot{\Delta}_0\rho_1\|_1\leqslant C_k\sup_{|\alpha|=k}\|\partial^\alpha\rho_1\|_1.$$

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 such that $k>n-\sigma$. ■

The following lemma provides a useful criterion for determining whether the sum of a series belongs to a homogeneous Besov space.

Lemma 6.8.9.

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 $\sum_{k\in\mathbb{Z}}f_k$ converges in $\mathscr{S}'$ to some $f\in\mathscr{S}'_h$, then $f\in\dot{\mathrm{B}}^s_{p,r}$ and

$$\|f\|_{\dot{\mathrm{B}}^s_{p,r}}\leqslant C_s\left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})}.$$

Proof. It is clear that there exists some positive integer N_0 such that $\dot{\Delta}_jf_k=0$ for $|j-k|\geqslant N_0$. Hence,

$$\|\dot{\Delta}_jf\|_p=\left\|\sum_{|j-k|<N_0}\dot{\Delta}_jf_k\right\|_p\leqslant C\sum_{|j-k|<N_0}\|f_k\|_p.$$

Therefore, we obtain that

$$2^{js}\|\dot{\Delta}_jf\|_p\leqslant C\sum_{|j-k|<N_0}2^{(j-k)s}2^{ks}\|f_k\|_p=C\sum_{k\in\mathbb{Z}}2^{(j-k)s}\chi_{|j-k|<N_0}(k)2^{ks}\|f_k\|_p.$$

Thus, by Young's inequality, we get

$$\|f\|_{\dot{\mathrm{B}}^s_{p,r}}\leqslant C\left(\sum_{j=-N_0+1}^{N_0-1}2^{js}\right)\left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})}\leqslant C_s\left\|\{2^{ks}\|f_k\|_p\}_k\right\|_{\ell^r(\mathbb{Z})}.$$

As $f\in\mathscr{S}'_h$ by assumption, this proves the lemma. ■

Remark 6.8.10. The above convergence assumption concerns $\{f_k\}_{k<0}$. We note that if (s,p,r) satisfies the condition (6.8.1), i.e.,

$$s<\frac{n}{p},\quad\text{or}\quad s=\frac{n}{p}\text{ and }r=1,\tag{6.8.1}$$

then, owing to Theorem 6.8.1, 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 $\dot{S}_kf$ tends to 0 when k goes to $-\infty$. In particular, we have $f\in\mathscr{S}'_h$.

Theorem 6.8.11.

Let $s \in \mathbb{R}$, $p, r \in [1, \infty]$. Then $\dot{\mathcal{B}}^s_{p,r}(\mathbb{R}^n)$ is a Banach space when $s < \frac{n}{p}$. In addition, $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n)$ is also a Banach space.

Proof. By Proposition 6.8.5, both $\dot{\mathcal{B}}^s_{p,r}(\mathbb{R}^n)$ and $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n)$ are normed spaces.

Step 1. To prove the embedding: $\dot{\mathcal{B}}^s_{p,r}(\mathbb{R}^n) \hookrightarrow \mathscr{S}'$ for $s < \frac{n}{p}$, and $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n) \hookrightarrow \mathscr{S}'$.

We know that $\dot{\mathcal{B}}^s_{p,r}(\mathbb{R}^n) \subset \mathscr{S}'$ for $s < \frac{n}{p}$ and $\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}(\mathbb{R}^n) \subset \mathscr{S}'$ by the definition of Besov spaces due to $\mathscr{S}'_h \subset \mathscr{S}'$, but the embedding relation in topological sense needs to prove. From Bernstein's inequality, it follows that

$$\|\dot{\Delta}_k u\|_\infty \leqslant C 2^{k\frac{n}{p}} \|\dot{\Delta}_k u\|_p. \tag{6.8.7}$$

For $u \in \dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}$, we have

$$\|u\|_\infty \leqslant \sum_{k\in\mathbb{Z}} \|\dot{\Delta}_k u\|_\infty \leqslant C \sum_{k\in\mathbb{Z}} 2^{k\frac{n}{p}} \|\dot{\Delta}_k u\|_p = C\|u\|_{\dot{\mathcal{B}}^{\frac{n}{p}}_{p,1}},$$

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 get

$$\begin{aligned}|\langle \dot{\Delta}_k u, f\rangle| \leqslant& \|\dot{\Delta}_k u\|_\infty \|f\|_1 \leqslant 2^{k\frac{n}{p}} \|\dot{\Delta}_k u\|_p \|f\|_1 \\ \leqslant& C 2^{k\left(\frac{n}{p}-s\right)} \|u\|_{\dot{\mathcal{B}}^s_{p,\infty}} \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1}|f(x)|.\end{aligned} \tag{6.8.8}$$

Thus,

$$\left|\left\langle \sum_{k<0} \dot{\Delta}_k u, f\right\rangle\right| \leqslant C\|u\|_{\dot{\mathcal{B}}^s_{p,r}} \sup_{x\in\mathbb{R}^n} (1+|x|)^{n+1}|f(x)|. \tag{6.8.9}$$

For high frequencies $k \geqslant 0$, we can use, as in (6.4.11),

$$\dot{\Delta}_k u = 2^{-kl} \sum_{|\alpha|=l} \partial^\alpha (2^{kn} g_\alpha(2^k \cdot) * \dot{\Delta}_k u), \quad g_\alpha := (2\pi)^{-n/2} a_\alpha \mathscr{F}^{-1}\left[\frac{(-i\xi)^\alpha}{|\xi|^{2l}} \varphi(\xi)\right]. \tag{6.8.10}$$

Then, it holds for $l \in \mathbb{N}_0$ and any $f \in \mathscr{S}$,

$$\begin{aligned}\langle \dot{\Delta}_k u, f\rangle =& 2^{-kl} \sum_{|\alpha|=l} \langle \partial^\alpha (2^{kn} g_\alpha(2^k\cdot) * \dot{\Delta}_k u), f\rangle \\ =& 2^{-kl} \sum_{|\alpha|=l} \langle (\dot{\Delta}_k u), 2^{kn} g_\alpha(-2^k \cdot) * (-\partial)^\alpha f\rangle \\ \leqslant& C\|\dot{\Delta}_k u\|_\infty 2^{-kl} \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|=l}} (1+|x|)^{n+1}|\partial^\alpha f(x)| \\ \leqslant& C 2^{k\left(\frac{n}{p}-s-l\right)} 2^{ks} \|\dot{\Delta}_k u\|_p \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|=l}} (1+|x|)^{n+1}|\partial^\alpha f(x)|.\end{aligned}$$

Thus, for large $l > \frac{n}{p} - s$, it follows that

$$\left|\left\langle \sum_{k\geqslant 0} \dot{\Delta}_k u, f\right\rangle\right| \leqslant C\|u\|_{\dot{\mathcal{B}}^s_{p,r}} \sup_{\substack{x\in\mathbb{R}^n\\|\alpha|=l}} (1+|x|)^{n+1}|\partial^\alpha f(x)|.$$

Therefore, we obtain

$$|\langle u, f\rangle| \leqslant \sum_{k\in\mathbb{Z}} |\langle \dot{\Delta}_k u, f\rangle| \leqslant C\|u\|_{\dot{\mathrm{B}}^s_{p,r}} \sup_{\substack{x\in\mathbb{R}^n\\ |\alpha|\leqslant l}} (1+|x|)^{n+1}|\partial^\alpha f(x)|, \quad \forall f\in\mathscr{S}, \tag{6.8.11}$$

which implies $\dot{\mathrm{B}}^s_{p,r} \hookrightarrow \mathscr{S}'$.

Step 2. To prove the completeness. Let $\{u_\ell\}_{\ell\in\mathbb{N}}$ is a Cauchy sequence in $\dot{\mathrm{B}}^s_{p,r}$, where $s<\frac{n}{p}$ or $s=\frac{n}{p}$ and $r=1$. Replacing u by $u_\ell-u_j$ in (6.8.11), 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 (6.8.8), we have for any $\ell\in\mathbb{N}$ and $j\in\mathbb{Z}$

$$\begin{aligned}|\langle \dot{S}_j u_\ell, f\rangle| \leqslant& \sum_{k\leqslant j-1} |\langle \dot{\Delta}_k u_\ell, f\rangle| \leqslant \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_\ell\|_\infty \|f\|_1\\ \leqslant& C_s 2^{j\left(\frac{n}{p}-s\right)} \sup_\ell \|u_\ell\|_{\dot{\mathrm{B}}^s_{p,r}} \|f\|_1.\end{aligned}$$

From $u_\ell \xrightarrow{\mathscr{S}'} u\in\mathscr{S}'$, it follows that

$$|\langle \dot{S}_j u, f\rangle| \leqslant C_s 2^{j\left(\frac{n}{p}-s\right)} \sup_\ell \|u_\ell\|_{\dot{\mathrm{B}}^s_{p,r}} \|f\|_1, \quad \forall f\in\mathscr{S}.$$

Hence, we get

$$\lim_{j\to-\infty} \dot{S}_j u = 0, \quad \text{i.e., } u\in\mathscr{S}'_h.$$

For the case $s=\frac{n}{p}$ and $r=1$, since $\{u_\ell\}$ is Cauchy in $\dot{\mathrm{B}}^{\frac{n}{p}}_{p,1} \hookrightarrow \dot{\mathrm{B}}^0_{\infty,1}$, we have $\forall\varepsilon>0$, $\exists \ell_0\in\mathbb{N}$, s.t. $\forall j\in\mathbb{Z}$ and $\ell\geqslant\ell_0$

$$\begin{aligned}\sum_{k\leqslant j-1} \|\dot{\Delta}_k u_\ell\|_\infty \leqslant& \sum_{k\leqslant j-1} \|\dot{\Delta}_k (u_\ell-u_{\ell_0})\|_\infty + \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty\\ \leqslant& \|u_\ell-u_{\ell_0}\|_{\dot{\mathrm{B}}^0_{\infty,1}} + \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty\\ \leqslant& \frac{\varepsilon}{2} + \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_{\ell_0}\|_\infty.\end{aligned}$$

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, $\forall j\leqslant j_0$, $\forall \ell\geqslant\ell_0$

$$\|\dot{S}_j u_\ell\|_\infty \leqslant \sum_{k\leqslant j-1} \|\dot{\Delta}_k u_\ell\|_\infty < \varepsilon. \tag{6.8.12}$$

Since $\dot{\mathrm{B}}^{\frac{n}{p}}_{p,1} \hookrightarrow \dot{\mathrm{B}}^0_{\infty,1} \hookrightarrow L^\infty$, $\{u_\ell\}_{\ell\in\mathbb{N}}$ 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 (6.8.12) 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 $u\in\dot{\mathrm{B}}^s_{p,r}$. 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 does $\{2^{ks}\|\dot\Delta_k u\|_p\}$. It follows that $u \in \dot{\mathcal{B}}^s_{p,r}$ from Lemma 6.8.9.

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 get

$$\left(\sum_{|k|\leqslant K}\left(2^{ks}\|\dot\Delta_k(u_\ell - u)\|_p\right)^r\right)^{\frac1r} = \lim_{m\to\infty}\left(\sum_{|k|\leqslant K}\left(2^{ks}\|\dot\Delta_k(u_\ell - u_m)\|_p\right)^r\right)^{\frac1r}.$$

Noticing that $\{u_\ell\}_{\ell\in\mathbb{N}}$ is Cauchy in $\dot{\mathcal{B}}^s_{p,r}$, thus, for any $\varepsilon > 0$, there exists an $\ell_0 \in \mathbb{N}$ 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 6.8.12. 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 6.8.13. 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

$$\begin{aligned}\|f_k - f_\ell\|_{\dot{\mathcal{B}}^{1/2}_{2,\infty}} &= \sup_{j\in\mathbb{Z}} 2^{j/2}\|\dot\Delta_j(f_k - f_\ell)\|_2 = \sup_{j\in\mathbb{Z}} 2^{j/2}\left(\int_{2^{-k}<|\xi|<2^{-\ell}}\left|\frac{\varphi_j(\xi)}{\xi\ln|\xi|}\right|^2 d\xi\right)^{1/2}\\ &\leqslant \sup_{j\in\mathbb{Z}} 2^{j/2}\frac{1}{\ell\ln 2}\left(\int_{\mathbb{R}}\left|\frac{\varphi(2^{-j}\xi)}{\xi}\right|^2 d\xi\right)^{1/2}\\ &= \frac{1}{\ell\ln 2}\left(\int_{\mathbb{R}}\left|\frac{\varphi(\xi)}{\xi}\right|^2 d\xi\right)^{1/2} \to 0, \text{ as } k, \ell \to \infty,\end{aligned}$$

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 $\phi \in \mathscr{S}$,

$$\langle u, \phi\rangle = \sum_{|k-j|\leqslant 1} \langle \dot{\Delta}_k u, \dot{\Delta}_j \phi\rangle = \sum_{|k-j|\leqslant 1} \int_{\mathbb{R}^n} \dot{\Delta}_k u(x) \dot{\Delta}_j \phi(x) dx.$$

As for the L^p space, we can estimate the norm in $\dot{\mathcal{B}}^s_{p,r}$ by duality.

Proposition 6.8.14.

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

$$|\langle u, \phi\rangle| \leqslant C_s \|u\|_{\dot{\mathcal{B}}^s_{p,r}} \|\phi\|_{\dot{\mathcal{B}}^{-s}_{p',r'}}.$$

In order 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 get

$$\begin{aligned} \|u\|_{\dot{\mathcal{B}}^s_{p,r}} &= \sup_{N\in\mathbb{N}} \left\| \left(\chi_{|k|\leqslant N}(k) 2^{ks} \|\dot{\Delta}_k u\|_p \right)_k \right\|_{\ell^r} \\ &= \sup_{N\in\mathbb{N}} \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \|\dot{\Delta}_k u\|_p 2^{ks} \alpha_k \quad \text{(by duality of } \ell^r\text{)} \\ &= \sup_{N\in\mathbb{N}} \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} 2^{ks} \alpha_k \sup_{\substack{\tilde{\phi}\in\mathscr{S} \\ \|\tilde{\phi}\|_{p'}\leqslant 1}} \langle \dot{\Delta}_k u, \tilde{\phi}\rangle, \quad \text{(by duality of } L^p\text{)}. \end{aligned}$$

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

$$\Phi_N := \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \alpha_k 2^{ks} \dot{\Delta}_k \phi_k.$$

Then, for $r' \in [1, \infty)$

$$
\begin{aligned}
\|\Phi_N\|_{\dot{\mathcal{B}}^{-s}_{p',r'}} =& \left(\sum_{j\in\mathbb{Z}} 2^{-jsr'} \left\| \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \alpha_k 2^{ks} \dot{\Delta}_j \dot{\Delta}_k \phi_k \right\|_{p'}^{r'}\right)^{1/r'} \\
=& \left(\sum_{j\in\mathbb{Z}} 2^{-jsr'} \left\| \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} \chi_{[j-1,j+1]}(k) \alpha_k 2^{ks} \dot{\Delta}_j \dot{\Delta}_k \phi_k \right\|_{p'}^{r'}\right)^{1/r'} \\
\leqslant& C \left(3^{r'-1} \sum_{j\in\mathbb{Z}} \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} |\alpha_k|^{r'} \chi_{[j-1,j+1]}(k) 2^{(k-j)sr'} \|\phi_k\|_{p'}^{r'} \right)^{1/r'} \\
\leqslant& C \left(3^{r'-1} \sum_{j\in\mathbb{Z}} \sup_{(\alpha_k)\in Q^{r'}_N} \left(\sum_{|k|\leqslant N} |\alpha_k|^{r'}\right) \sup_{|k|\leqslant N} \chi_{[j-1,j+1]}(k) 2^{(k-j)sr'} \right)^{1/r'} \\
\leqslant& C 2^{|s|} \left(3 \cdot 3^{r'-1}\right)^{1/r'} \\
\leqslant& 3C2^{|s|},
\end{aligned}
$$

which is independent of N.

Thus, for any N,

$$
\begin{aligned}
\left\| \left(\chi_{|k|\leqslant N}(k) 2^{ks} \|\dot{\Delta}_k u\|_p\right)_k \right\|_{\ell^r} <& \langle u, \Phi_N\rangle + \sup_{(\alpha_k)\in Q^{r'}_N} \sum_{|k|\leqslant N} 2^{ks} |\alpha_k| \frac{\varepsilon 2^{-ks}}{(1+|\alpha_k|)(1+|k|^2)} \\
\leqslant& \langle u, \Phi_N\rangle + \varepsilon.
\end{aligned}
$$

Therefore, we complete the proof. ■

§ 6.9 Hölder spaces

Definition 6.9.1.

Let $0 < \alpha < 1$. Define the ***Hölder (or Lipschitz) space*** $\mathcal{C}^\alpha$ as

$$\mathcal{C}^\alpha = \{f \in L^\infty(\mathbb{R}^n) : \|f(x-t) - f(x)\|_\infty \leqslant A|t|^\alpha\}.$$

The $\mathcal{C}^\alpha$ norm is then given by

$$\|f\|_{\mathcal{C}^\alpha} = \|f\|_\infty + \sup_{|t|>0} \frac{\|f(x-t) - f(x)\|_\infty}{|t|^\alpha}. \tag{6.9.1}$$

The first thing to observe is that the functions in $\mathcal{C}^\alpha$ may be taken to be continuous, and so the relation $|f(x-t) - f(x)| \leqslant A|t|^\alpha$ holds for every x. More precisely,

Proposition 6.9.2.

Every $f \in \mathcal{C}^\alpha$ may be modified on a set of measure zero such that it becomes continuous.

Proof. The proof can be carried out by using the device of regularization. Any smooth regularization will do, and we shall use here that of the Poisson integral.

Thus, we consider

$$u(x,y)=\int_{\mathbb{R}^n}P_y(t)f(x-t)dt,\quad P_y(t)=\frac{c_n y}{(|t|^2+y^2)^{(n+1)/2}},\quad y>0.$$

Then, since $\int_{\mathbb{R}^n}P_y(t)dt=1$, we have

$$u(x,y)-f(x)=\int_{\mathbb{R}^n}P_y(t)[f(x-t)-f(x)]dt,$$

and for $0<\alpha<1$,

$$\begin{aligned}\|u(x,y)-f(x)\|_\infty\leqslant&\int_{\mathbb{R}^n}P_y(t)\|f(x-t)-f(x)\|_\infty dt\\ \leqslant&Ac_n y\int_{\mathbb{R}^n}\frac{|t|^\alpha}{(|t|^2+y^2)^{(n+1)/2}}dt\\ \overset{t=ys}{=\!=}&Ac_n y^\alpha\int_{\mathbb{R}^n}\frac{|s|^\alpha}{(|s|^2+1)^{(n+1)/2}}ds=A'y^\alpha.\end{aligned}$$

In particular, $\|u(x,y_1)-u(x,y_2)\|_\infty\to 0$, as y_1 and y_2 $\to 0$, and since $u(x,y)$ is continuous in x, then $u(x,y)$ converges uniformly to $f(x)$ as $y\to 0$. Therefore, $f(x)$ may be taken to be continuous. ■

We begin by giving a characterization of $f\in\mathcal{C}^\alpha$ in terms of their Poisson integrals $u(x,y)$.

Proposition 6.9.3.

Suppose $f\in L^\infty(\mathbb{R}^n)$ and $0<\alpha<1$. Then $f\in\mathcal{C}^\alpha(\mathbb{R}^n)$ if and only if

$$\left\|\frac{\partial u(x,y)}{\partial y}\right\|_\infty\leqslant Ay^{-1+\alpha}. \tag{6.9.2}$$

Remark 6.9.4. If A_1 is the smallest constant A for which (6.9.2) holds, then $\|f\|_\infty+A_1$ and $\|f\|_{\mathcal{C}^\alpha}$ give equivalent norms.

Proof. For Poisson kernel, we have

$$\begin{aligned}\frac{\partial P_y(x)}{\partial y}=&c_n\frac{(|x|^2+y^2)^{(n+1)/2}-y\frac{n+1}{2}(|x|^2+y^2)^{(n-1)/2}\cdot 2y}{(|x|^2+y^2)^{n+1}}\\ =&c_n\frac{|x|^2+y^2-(n+1)y^2}{(|x|^2+y^2)^{(n+1)/2+1}}=c_n\frac{|x|^2-ny^2}{(|x|^2+y^2)^{(n+1)/2+1}},\end{aligned}\tag{6.9.3}$$

and then

$$\left|\frac{\partial P_y(x)}{\partial y}\right|\leqslant\frac{c}{(|x|^2+y^2)^{(n+1)/2}},\quad y>0. \tag{6.9.4}$$

Differentiating $\int_{\mathbb{R}^n}P_y(x)dx=1$ w.r.t. y, we obtain

$$\int_{\mathbb{R}^n}\frac{\partial P_y(x)}{\partial y}dx=0,\quad y>0. \tag{6.9.5}$$

Thus, it follows

$$\frac{\partial u}{\partial y}(x,y)=\int_{\mathbb{R}^n}\frac{\partial P_y(t)}{\partial y}f(x-t)dt=\int_{\mathbb{R}^n}\frac{\partial P_y(t)}{\partial y}[f(x-t)-f(x)]dt.$$

Hence, by changing variables, we have

$$\begin{aligned}\left\|\frac{\partial u(\cdot,y)}{\partial y}\right\|_\infty\leqslant&\|f\|_{\mathcal{C}^\alpha}\int_{\mathbb{R}^n}\left|\frac{\partial P_y(t)}{\partial y}\right||t|^\alpha dt\leqslant c\|f\|_{\mathcal{C}^\alpha}\int_{\mathbb{R}^n}\frac{|t|^\alpha}{(|t|^2+y^2)^{(n+1)/2}}dt\\ \leqslant&c\|f\|_{\mathcal{C}^\alpha}\int_{\mathbb{R}^n}\frac{1}{(|t|^2+y^2)^{(n+1-\alpha)/2}}dt\end{aligned}$$

$$\overset{t=ys}{=\!=\!=} c\|f\|_{\mathfrak{C}^\alpha} y^{-1+\alpha} \int_{\mathbb{R}^n} \frac{1}{(|s|^2+1)^{(n+1-\alpha)/2}} ds \leqslant C\|f\|_{\mathfrak{C}^\alpha} y^{-1+\alpha}.$$

This proves the necessariness part.

For the sufficiency part, it is far more enlightening, as it reveals an essential feature of the spaces in question, although it is not much more difficult. This insight is contained in the lemma below and the comments that follow. Then we shall return to the proof of the second part (to be continued). ■

Lemma 6.9.5.

Suppose $f \in L^\infty(\mathbb{R}^n)$ and $0 < \alpha < 1$. Then the single condition (6.9.2) is equivalent with the n conditions

$$\left\| \frac{\partial u(x,y)}{\partial x_j} \right\|_\infty \leqslant A' y^{-1+\alpha}, \quad j = 1, ..., n. \tag{6.9.6}$$

Remark 6.9.6. The smallest A in (6.9.2) is comparable to the smallest A' in (6.9.6).

Proof. From the Poisson kernel, we can derive

$$\frac{\partial P_y(x)}{\partial x_j} = -\frac{(n+1)c_n y x_j}{(|x|^2+y^2)^{(n+1)/2+1}}, \quad \left|\frac{\partial P_y(x)}{\partial x_j}\right| \leqslant \frac{C}{(|x|^2+y^2)^{(n+1)/2}}, \quad y>0. \tag{6.9.7}$$

For $y = y_1 + y_2$, it follows from Corollary 2.1.24 that $P_y = P_{y_1} * P_{y_2}$, with $y_1, y_2 > 0$. Thus,

$$u(x,y) = P_y * f = P_{y_1} * P_{y_2} * f = P_{y_1} * u(x, y_2),$$

and therefore, with $y_1 = y_2 = y/2$, we get

$$\frac{\partial^2 u}{\partial y \partial x_j} = \frac{\partial P_{y/2}}{\partial x_j} * \frac{\partial u(x, y/2)}{2\partial y_2}.$$

By Young's inequality, (6.9.7) and (6.9.2), we get

$$\begin{aligned}\left\|\frac{\partial^2 u}{\partial y \partial x_j}\right\|_\infty \leqslant & \left\|\frac{\partial P_{y/2}}{\partial x_j}\right\|_1 \left\|\frac{\partial u(x,y/2)}{2\partial y_2}\right\|_\infty \\ \leqslant & C\int_{\mathbb{R}^n} \frac{dx}{(|x|^2+y^2/4)^{(n+1)/2}} \cdot 2^{-\alpha} A y^{-1+\alpha} \\ \overset{x=yt/2}{=\!=\!=} & CAy^{-2+\alpha} \int_{\mathbb{R}^n} \frac{dt}{(|t|^2+1)^{(n+1)/2}} = A_1 y^{-2+\alpha}.\end{aligned} \tag{6.9.8}$$

However, by Young's inequality and (6.9.7),

$$\left\|\frac{\partial}{\partial x_j} u(x,y)\right\|_\infty = \left\|\frac{\partial P_y}{\partial x_j} * f\right\|_\infty \leqslant \left\|\frac{\partial P_y}{\partial x_j}\right\|_1 \|f\|_\infty \leqslant \frac{c\|f\|_\infty}{y}.$$

So

$$\frac{\partial}{\partial x_j} u(x,y) \to 0, \quad \text{as } y \to \infty,$$

and therefore,

$$\frac{\partial}{\partial x_j} u(x,y) = -\int_y^\infty \frac{\partial^2 u(x,y')}{\partial y' \partial x_j} dy'.$$

Then, for $\alpha < 1$, (6.9.8) gives that

$$\left\|\frac{\partial u}{\partial x_j}\right\|_\infty \leqslant A_1 \int_y^\infty y'^{-2+\alpha} dy' \leqslant A_2 y^{-1+\alpha}.$$

Conversely, suppose that (6.9.6) is satisfied. Reasoning as before, we get that $\|\frac{\partial^2 u}{\partial x_j^2}\|_\infty \leqslant A_3 y^{-2+\alpha}$, $j = 1, ..., n$. However, since u is harmonic, that is because $\frac{\partial^2 u}{\partial y^2} = -\sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2}$, we have $\|\frac{\partial^2 u}{\partial y^2}\|_\infty \leqslant A_4 y^{-2+\alpha}$. Then, a similar integration argument shows that $\|\frac{\partial u}{\partial y}\|_\infty \leqslant A_5 y^{-1+\alpha}$. ■

We can now prove the converse part of Proposition 6.9.3.

Proof of Proposition 6.9.3 (continue). Suppose $\|\frac{\partial}{\partial y} u(x, y)\|_\infty \leqslant A y^{-1+\alpha}$. Then Lemma 6.9.5 also shows that $\|\frac{\partial}{\partial x_j} u(x, y)\|_\infty \leqslant A' y^{-1+\alpha}$. We write

$$\begin{aligned} f(x+t) - f(x) =& [u(x+t, y) - u(x, y)] + [f(x+t) - u(x+t, y)] \\ & - [f(x) - u(x, y)]. \end{aligned}$$

Here y does not necessarily depend on t but it is best to choose $y = |t|$. Now, we have $|u(x+t, y) - u(x, y)| \leqslant \int_L |\nabla_x u(x+s, y)| ds$ where L is the line segment (of length $|t|$) joining x with $x+t$. Thus, it follows

$$|u(x+t, y) - u(x, y)| \leqslant |t| \sum_{j=1}^n \|u_{x_j}(x, y)\|_\infty \leqslant C|t||t|^{-1+\alpha} = C|t|^\alpha.$$

Since

$$f(x+t) - u(x+t, y) = -\int_0^y \frac{\partial}{\partial y'} u(x+t, y') dy',$$

we get

$$|f(x+t) - u(x+t, y)| \leqslant \int_0^y \left\| \frac{\partial u}{\partial y'} \right\|_\infty dy' \leqslant C y^\alpha = C|t|^\alpha.$$

With a similar estimate for $f(x) - u(x, y)$, the proof of the proposition is concluded. ■

Similar to Lemma 6.9.5, we can prove the following lemma, and remaind the proof to interested readers.

Lemma 6.9.7.

Suppose $f \in L^\infty(\mathbb{R}^n)$, and $\alpha > 0$. Let k and l be two integers, both greater than α. Then the two conditions

$$\left\| \frac{\partial^k u(x, y)}{\partial y^k} \right\|_\infty \leqslant A_k y^{-k+\alpha}, \quad \text{and} \quad \left\| \frac{\partial^l u(x, y)}{\partial y^l} \right\|_\infty \leqslant A_l y^{-l+\alpha}$$

are equivalent. Moreover, the smallest A_k and A_l holding in the above inequalities are comparable.

The utility of this lemma will be apparent soon.

We now can define the space $\mathcal{C}^\alpha(\mathbb{R}^n)$ for any $\alpha > 0$. Suppose that $k = \lceil \alpha \rceil$ is the smallest integer greater than α, i.e., the ceiling function of α. We set

$$\mathcal{C}^\alpha = \left\{ f \in L^\infty(\mathbb{R}^n) : \left\| \frac{\partial^k}{\partial y^k} u(x, y) \right\|_\infty \leqslant A y^{-k+\alpha} \right\}. \tag{6.9.9}$$

If A_k denotes the smallest A appearing in the inequality in (6.9.9), then we can define the $\mathcal{C}^\alpha$ norm by

$$\|f\|_{\mathcal{C}^\alpha} = \|f\|_\infty + A_k. \tag{6.9.10}$$

According to Proposition 6.9.3, when $0 < \alpha < 1$, this definition is equivalent with the previous one and the resulting norms are also equivalent. Lemma 6.9.7

also shows us that we could have replaced the $\frac{\partial^k u(x,y)}{\partial y^k}$ by the corresponding estimate for $\frac{\partial^l u(x,y)}{\partial y^l}$ where l is any integer greater than α.

A remark about the condition in (6.9.9) is in order. The estimate

$$\left\| \frac{\partial^k}{\partial y^k} u(x,y) \right\|_\infty \leqslant Ay^{-k+\alpha}$$

is of interest only for y near zero, since the inequality $\left\| \frac{\partial^k}{\partial y^k} u(x,y) \right\|_\infty \leqslant Ay^{-k}$ (which is stronger away from zero) follows already from the fact that $f \in L^\infty$, (as the argument of Lemma 6.9.5 shows). This observation allows us to assert the inclusion $\mathcal{C}^\alpha \hookrightarrow \mathcal{C}^{\alpha'}$, if $\alpha > \alpha'$.

In the case of $0 < \alpha < 1$, we considered the first order difference, next, we will consider the case $0 < \alpha < 2$, it would be better to use the second order differences. We recall the ***m-th order difference operator*** $\triangle_t^m$

$$\triangle_t^m f(x) = \sum_{k=0}^m C_m^k (-1)^k f(x+kt).$$

Thus, $\triangle_t^2 f(x) = f(x) - 2f(x+t) + f(x+2t)$. But for simplicity, we denote

$$\triangle_t^2 f(x) = f(x-t) - 2f(x) + f(x+t)$$

in this section.

Proposition 6.9.8.

Suppose $0 < \alpha < 2$. Then $f \in \mathcal{C}^\alpha$ if and only if $f \in L^\infty(\mathbb{R}^n)$ and $\|f(x-t) - 2f(x) + f(x+t)\|_\infty \leqslant A|t|^\alpha$. The expression

$$\|f\|_\infty + \sup_{|t|>0} \frac{\|f(x-t) - 2f(x) + f(x+t)\|_\infty}{|t|^\alpha}$$

is equivalent with the $\mathcal{C}^\alpha$ norm.

Proof. Differentiating $\int_{\mathbb{R}^n} P_y(t)dt = 1$ twice w.r.t. y, we obtain

$$\int_{\mathbb{R}^n} \frac{\partial^2 P_y(t)}{\partial y^2} dt = 0, \quad y > 0. \tag{6.9.11}$$

From (6.9.3), we have

$$\frac{\partial^2 P_y(t)}{\partial y^2} = -\frac{c_n(n+1)(3|t|^2 - ny^2)y}{(|t|^2+y^2)^{(n+5)/2}},$$

and then

$$\frac{\partial^2 P_y(t)}{\partial y^2} = \frac{\partial^2 P_y(-t)}{\partial y^2}, \quad \left| \frac{\partial^2 P_y(t)}{\partial y^2} \right| \leqslant \frac{c}{(|t|^2+y^2)^{(n+2)/2}}. \tag{6.9.12}$$

Thus, we get

$$\frac{\partial^2}{\partial y^2} u(x,y) = \frac{1}{2} \int_{\mathbb{R}^n} \frac{\partial^2}{\partial y^2} P_y(t)[f(x-t) - 2f(x) + f(x+t)]dt,$$

and so, for $\alpha < 2$,

$$\begin{aligned} \left\| \frac{\partial^2}{\partial y^2} u(x,y) \right\|_\infty \leqslant & \frac{Ac}{2} \left[y^{-n-2} \int_{|t|\leqslant y} |t|^\alpha dt + \int_{|t|>y} |t|^{-n-2+\alpha} dt \right] \\ \leqslant & C \left[y^{-n-2} \int_0^y r^{\alpha+n-1} dr + \int_y^\infty r^{-3+\alpha} dr \right] \\ \leqslant & Cy^{-2+\alpha}. \end{aligned}$$

To prove the converse, we observe that if F has two orders continuous derivatives, then we have

$$\triangle_t^2 F(x) = \int_0^{|t|}\int_{-s}^{s} \frac{d^2}{d\tau^2}F(x+t'\tau)d\tau ds, \quad \text{where } t' = t/|t|.$$

It follows immediately that

$$\| \triangle_t^2 F(x)\|_\infty \leqslant |t|^2 \sum_{i,j} \left\| \frac{\partial^2 F}{\partial x_i \partial x_j}\right\|_\infty . \tag{6.9.13}$$

By the definition (6.9.9), it is clear that $f \in \mathcal{C}^\alpha \Rightarrow f \in \mathcal{C}^{\alpha'}$ where $\alpha' < \alpha$. If we choose an $\alpha' < 1$, then by the results in Propositions 6.9.2 and 6.9.3, we get

$$\|u(x,y) - f(x)\|_\infty \to 0, \quad \text{and } y\|u_y(x,y)\|_\infty \to 0, \text{ as } y \to 0. \tag{6.9.14}$$

Thus, the identity

$$f(x) = u(x,0) = \int_0^y y' \frac{\partial^2}{\partial y'^2}u(x,y')dy' - y\frac{\partial u}{\partial y}(x,y) + u(x,y) \tag{6.9.15}$$

is obtained by noticing that the derivative w.r.t. y of the extreme r.h.s. vanishes, and by the use of the end-point conditions (6.9.14). However, the arguments of Lemma 6.9.5 and 6.9.7 show that the inequality $\|\frac{\partial^2 u(x,y)}{\partial y^2}\|_\infty \leqslant Ay^{-2+\alpha}$ implies the estimates

$$\left\|\frac{\partial^2 u(x,y)}{\partial x_i \partial x_j}\right\|_\infty \leqslant A' y^{-2+\alpha}, \quad \left\|\frac{\partial^3 u(x,y)}{\partial y \partial x_i \partial x_j}\right\|_\infty \leqslant A' y^{-3+\alpha}.$$

Thus, by using(6.9.13) to the last two terms of the r.h.s. of (6.9.15),

$$\begin{aligned}
&\| \triangle_t^2 f\|_\infty \\
\leqslant& \| \triangle_t^2 \int_0^y y' \frac{\partial^2}{\partial y'^2}u(x,y')dy'\|_\infty + y\| \triangle_t^2 \frac{\partial u}{\partial y}(x,y)\|_\infty + \| \triangle_t^2 u(x,y)\|_\infty \\
\leqslant& 4\int_0^y y'\|\frac{\partial^2}{\partial y'^2}u(x,y')\|_\infty dy' + |t|^2 \sum_{i,j}\left[y\left\|\frac{\partial^3 u}{\partial y\partial x_i \partial x_j}\right\|_\infty + \left\|\frac{\partial^2 u}{\partial x_i\partial x_j}\right\|_\infty\right] \\
\leqslant& C\int_0^y y' y'^{-2+\alpha}dy' + C|t|^2[yy^{-3+\alpha} + y^{-2+\alpha}] \\
\leqslant& Cy^\alpha + C|t|^2 y^{-2+\alpha}.
\end{aligned}$$

Taking $y = |t|$ gives

$$\|\triangle_t^2 f\|_\infty \leqslant C|t|^\alpha, \quad \text{if } \alpha > 0,$$

which is the desired result. ■

Proposition 6.9.9.

Suppose $\alpha > 1$. Then $f \in \mathcal{C}^\alpha$ if and only if $f \in L^\infty$ and $\frac{\partial f}{\partial x_j} \in \mathcal{C}^{\alpha-1}$, $j = 1, ..., n$. The norms $\|f\|_{\mathcal{C}^\alpha}$ and $\|f\|_\infty + \sum_{j=1}^n \|\frac{\partial f}{\partial x_j}\|_{\mathcal{C}^{\alpha-1}}$ are equivalent.

Proof. Let us suppose for simplicity that $1 < \alpha \leqslant 2$, the other cases can be argued similarly.

We first prove that $\frac{\partial f}{\partial x_j} \in L^\infty$. We have $\|\frac{\partial^3 u}{\partial y^3}\|_\infty \leqslant Ay^{-3+\alpha}$ since $f \in \mathcal{C}^\alpha$, which implies, as we know, $\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\|_\infty \leqslant Ay^{-3+\alpha}$. Equivalently, we see that $\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\|_\infty \leqslant Ay^{-1-\beta}$, where $0 \leqslant \beta < 1$ since $1 < \alpha \leqslant 2$. We restrict to $0 < y \leqslant 1$, then an integration in y gives

$$\int_y^1 \frac{\partial^3 u}{\partial y^2 \partial x_j} dy = \left[\frac{\partial^2 u}{\partial y \partial x_j}\right]_{y=1} - \frac{\partial^2 u}{\partial y \partial x_j},$$

and then

$$\left\|\frac{\partial^2 u}{\partial y \partial x_j}\right\|_\infty \leqslant \int_y^1 \left\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\right\|_\infty dy + \left\|\left[\frac{\partial^2 u}{\partial y \partial x_j}\right]_{y=1}\right\|_\infty \leqslant Cy^{-\beta} + C.$$

Another integration,

$$\int_{y_1}^{y_2} \frac{\partial^2 u}{\partial y \partial x_j} dy = \frac{\partial}{\partial x_j} u(x, y_2) - \frac{\partial}{\partial x_j} u(x, y_1),$$

then shows that

$$\left\|\frac{\partial}{\partial x_j} u(x, y_2) - \frac{\partial}{\partial x_j} u(x, y_1)\right\|_\infty$$
$$\leqslant \int_{y_1}^{y_2} [Cy^{-\beta} + C] dy = C(y_2^{1-\beta} - y_1^{1-\beta}) + C(y_2 - y_1).$$

Thus, $\{\frac{\partial}{\partial x_j} u(x, y)\}$ is Cauchy in the L^∞ norm as $y \to 0$, and so its limit can be taken to be $\frac{\partial f}{\partial x_j}$. The argument also gives the bound

$$\left\|\frac{\partial f}{\partial x_j}\right\|_\infty \leqslant \left\|\frac{\partial u}{\partial x_j}\right\|_\infty \leqslant CAy^{-1+\alpha} \leqslant CA \leqslant C\|f\|_{\mathcal{C}^\alpha},$$

since $0 < y \leqslant 1$ and $\alpha > 1$.

Since the (weak) derivative of f is $\frac{\partial f}{\partial x_j}$, the Poisson integral of the latter is $\frac{\partial u}{\partial x_j}$. But $\|\frac{\partial^3 u}{\partial y^2 \partial x_j}\|_\infty \leqslant Ay^{-3+\alpha}$. Therefore, $\frac{\partial f}{\partial x_j} \in \mathcal{C}^{\alpha-1}$. The converse implication is proved in the same way. ■

The last proposition reduces the study of the spaces $\mathcal{C}^\alpha$ to those α such that $0 < \alpha \leqslant 1$. Concerning the space $\mathcal{C}^\alpha$, $0 < \alpha \leqslant 1$, the following additional remark is in order.

Remark 6.9.10. When $0 < \alpha < 1$, Proposition 6.9.8 shows that if $f \in L^\infty$, the two conditions $\|f(x+t) - f(x)\|_\infty \leqslant A|t|^\alpha$ and $\|f(x-t) - 2f(x) + f(x+t)\|_\infty \leqslant A'|t|^\alpha$ are equivalent. However, this is not the case when $\alpha = 1$.

Example 6.9.11. There exists $f \in L^\infty(\mathbb{R}^n)$ such that

$$\|f(x-t) - 2f(x) + f(x+t)\|_\infty \leqslant A|t|, \quad |t| > 0,$$

but $\|f(x+t) - f(x)\|_\infty \leqslant A'|t|$ fails for all A'.

Solution. One can construct such f by lacunary series, and more particularly as ***Hardy-Weierstrass non-differentiable functions***. To do this, we consider the function of one variable x, given by $f(x) = \sum_{k=1}^\infty a^{-k} e^{2\pi i a^k x}$. Here $a > 1$, for simplicity, we take a to be an integer and this makes f periodic. Now

$$f(x-t) - 2f(x) + f(x+t) = \sum_{k=1}^\infty a^{-k} [e^{2\pi i a^k (x-t)} - 2e^{2\pi i a^k x} + e^{2\pi i a^k (x+t)}]$$
$$= \sum_{k=1}^\infty a^{-k} [e^{-2\pi i a^k t} - 2 + e^{2\pi i a^k t}] e^{2\pi i a^k x} = 2\sum_{k=1}^\infty a^{-k} [\cos 2\pi a^k t - 1] e^{2\pi i a^k x}.$$

Therefore, (assume $|t| < 1$ without loss of generalities)

$$\|f(x-t) - 2f(x) + f(x+t)\|_\infty$$

$$\begin{aligned}
&\leqslant 2\sum_{a^k|t|\leqslant 1} a^{-k}B(a^kt)^2 + 4\sum_{a^k|t|>1} a^{-k}\\
&\leqslant 2B|t|^2 \sum_{k\leqslant[\log_a |t|^{-1}]} a^k + 4\sum_{k>[\log_a |t|^{-1}]} a^{-k}\\
&\leqslant 2B|t|^2 \frac{a^{[\log_a |t|^{-1}]+1}-1}{a-1} + 4\frac{a^{-[\log_a |t|^{-1}]-1}}{1-a^{-1}}\\
&\leqslant A|t|.
\end{aligned}$$

We have used merely the fact that $|\cos 2\pi a^k t - 1| \leqslant \min(B(a^kt)^2, 2)$ with $B = 2\pi^2$ since $\sin x \leqslant x$ for any $x \geqslant 0$.

However, if we had $\|f(x+t) - f(x)\|_\infty \leqslant A'|t|$, then by Bessel's inequality for L^2 periodic functions we would get

$$\begin{aligned}
(A'|t|)^2 \geqslant &\int_0^1 |f(x+t) - f(x)|^2 dx = \int_0^1 \left|\sum_{k=1}^\infty a^{-k}[e^{2\pi i a^k t} - 1]e^{2\pi i a^k x}\right|^2 dx\\
\geqslant &\sum_{k=1}^\infty a^{-2k}|e^{2\pi i a^k t} - 1|^2 \geqslant \sum_{a^k|t|<1/2} a^{-2k}|e^{2\pi i a^k t} - 1|^2.
\end{aligned}$$

In the range $a^k|t| < 1/2$, we have $|e^{2\pi i a^k t} - 1|^2 > 16(a^kt)^2$ due to the inequality $\frac{\sin x}{x} > \frac{2}{\pi}$ for any $x \in (-\pi/2, \pi/2)$, and so we would arrive at the contradiction

$$(A'|t|)^2 \geqslant 16|t|^2 \sum_{a^k|t|<1/2} 1 \geqslant 16|t|^2 \sum_{1\leqslant k\leqslant[\log_a |t|^{-1}/2]-1} 1,$$

which implies that $A'^2 \geqslant 16\left[\log_a |t|^{-1}/2\right] - 16$ and so $A' \to \infty$ as $|t| \to 0$. ■

Now, we give a relation between Hölder spaces and Besov spaces.

Corollary 6.9.12.

Let $s > 0$. Then we have $B^s_{\infty,\infty} = \mathcal{C}^s$.

Proof. By Theorem 6.7.1 with $p = r = \infty$ and $m = 2$, and Proposition 6.9.8, for $0 < s \leqslant 1$, we can take $N = 0$ and then

$$\|f\|_{B^s_{\infty,\infty}} \sim \|f\|_\infty + \sup_{t>0} t^{-s}\omega^2_\infty(t, f) = \|f\|_{\mathcal{C}^s}.$$

By Proposition 6.9.9, we can extend to any $s > 1$. This completes the proof. ■

References

[BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343 of *GMW*. Springer-Verlag, Berlin Heidelberg, 2011.

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[Gra14] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[HMOW11] Hichem Hajaiej, Luc Molinet, Tohru Ozawa, and Baoxiang Wang. Necessary and sufficient conditions for the fractional Gagliardo-Nirenberg inequalities and applications to Navier-Stokes and generalized boson equations. In *Harmonic analysis and nonlinear partial differential equations*, RIMS Kôkyûroku Bessatsu, B26, pages 159–175. Res. Inst. Math. Sci. (RIMS), Kyoto, 2011.

[Jaw77] Björn Jawerth. Some observations on Besov and Lizorkin-Triebel spaces. *Math. Scand.*, 40(1):94–104, 1977.

[MWZ12] Changxing Miao, Jiahong Wu, and Zhifei Zhang. *Littlewood-Paley theory and its application in hydrodynamic equations*, volume 126 of *Fundamental Series in Mordon Mathematics*. Science Press, Beijing, 2012.

[Saw18] Yoshihiro Sawano. *Theory of Besov spaces*, volume 56 of *Developments in Mathematics*. Springer, Singapore, 2018.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[Tri83] Hans Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983.

[WHHG11] Baoxiang Wang, Zhaohui Huo, Chengchun Hao, and Zihua Guo. *Harmonic analysis method for nonlinear evolution equations*, volume I. World Scientific Publishing Co. Pte. Ltd., 2011.

VII

BMO and Carleson Measures

§ 7.1 The sharp maximal functions and BMO spaces

Functions of bounded mean oscillation were introduced by F. John and L. Nirenberg [JN61], in connection with differential equations.

Definition 7.1.1.

The ***mean oscillation*** of a locally integrable function f (i.e. a function belonging to $L^1_{loc}(\mathbb{R}^n)$) over a cube Q in $\mathbb{R}^n$ is defined as the following integral:

$$\tilde{f}_Q = \frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f|dx,$$

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 7.1.2.

A ***BMO function*** is any function f belonging to $L^1_{loc}(\mathbb{R}^n)$ whose mean oscillation has a finite supremum over the set of all cubes[a] Q contained in $\mathbb{R}^n$. For $f \in L^1_{loc}(\mathbb{R}^n)$, we define the ***maximal BMO function*** or the ***sharp maximal function***

$$M^\# f(x) = \sup_{Q\ni x}\frac{1}{|Q|}\int_Q |f(t) - \operatorname*{Avg}_Q f|dt,$$

where the supremum is take over all cubes Q in $\mathbb{R}^n$ that contains the given point x, and $M^\#$ is called the ***sharp maximal operator***. Then we denote the norm of f in this space by $\|f\|_{\text{BMO}} = \|M^\# f\|_\infty$. The set

$$\text{BMO}(\mathbb{R}^n) = \{f \in L^1_{\text{loc}}(\mathbb{R}^n) : \|f\|_{\text{BMO}} < \infty\}$$

is called the ***function space of bounded mean oscillation*** or the ***BMO space***.

[a]The use of cubes Q in $\mathbb{R}^n$ as the integration domains on which the mean oscillation is calculated, is not mandatory: Wiegerinck (2001) uses balls instead and, as remarked by Stein ([Ste93], p. 140), in doing so a perfectly equivalent of definition of functions of bounded mean

oscillation arises.

Several remarks are in order. First, it is a simple fact that $\mathrm{BMO}(\mathbb{R}^n)$ 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

$$\begin{aligned}\|f+g\|_{\mathrm{BMO}} \leqslant & \|f\|_{\mathrm{BMO}} + \|g\|_{\mathrm{BMO}},\\ \|\lambda f\|_{\mathrm{BMO}} = & |\lambda| \|f\|_{\mathrm{BMO}}.\end{aligned}$$

But $\|\cdot\|_{\mathrm{BMO}}$ 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 semi-norm, 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 7.1.3.

The following properties of the space $\mathrm{BMO}(\mathbb{R}^n)$ are valid:

(1) If $\|f\|_{\mathrm{BMO}} = 0$, then f is a.e. equal to a constant.

(2) $L^\infty(\mathbb{R}^n) \hookrightarrow \mathrm{BMO}(\mathbb{R}^n)$ and $\|f\|_{\mathrm{BMO}} \leqslant 2\|f\|_\infty$.

(3) 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

$$\sup_Q \frac{1}{|Q|}\int_Q |f(x) - c_Q| dx \leqslant A. \tag{7.1.1}$$

Then $f \in \mathrm{BMO}(\mathbb{R}^n)$ and $\|f\|_{\mathrm{BMO}} \leqslant 2A$.

(4) For all $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, we have

$$\frac{1}{2}\|f\|_{\mathrm{BMO}} \leqslant \sup_Q \frac{1}{|Q|} \inf_{c_Q} \int_Q |f(x) - c_Q| dx \leqslant \|f\|_{\mathrm{BMO}}.$$

(5) 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

$$\|\tau^h(f)\|_{\mathrm{BMO}} = \|f\|_{\mathrm{BMO}}.$$

(6) 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

$$\|\delta^\lambda(f)\|_{\mathrm{BMO}} = \|f\|_{\mathrm{BMO}}.$$

(7) If $f \in \mathrm{BMO}(\mathbb{R}^n)$, then so is $|f|$. Similarly, if f, g are real-valued BMO functions, then so are $\max(f,g)$ and $\min(f,g)$. In other words, BMO is a lattice. Moreover,

$$\begin{aligned}\||f|\|_{\mathrm{BMO}} \leqslant & 2\|f\|_{\mathrm{BMO}},\\ \|\max(f,g)\|_{\mathrm{BMO}} \leqslant & \frac{3}{2}\left(\|f\|_{\mathrm{BMO}} + \|g\|_{\mathrm{BMO}}\right),\end{aligned}$$

$$\|\min(f,g)\|_{\mathrm{BMO}} \leqslant \frac{3}{2}\left(\|f\|_{\mathrm{BMO}} + \|g\|_{\mathrm{BMO}}\right).$$

(8) For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, define

$$\|f\|_{\mathrm{BMO}_{\mathrm{balls}}} = \sup_B \frac{1}{|B|}\int_B |f(x) - \operatorname*{Avg}_B f|dx, \tag{7.1.2}$$

where the supremum is taken over all balls B in $\mathbb{R}^n$. Then there are positive constants c_n, C_n such that

$$c_n\|f\|_{\mathrm{BMO}} \leqslant \|f\|_{\mathrm{BMO}_{\mathrm{balls}}} \leqslant C_n\|f\|_{\mathrm{BMO}}.$$

Proof. To prove (1), 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 (2), 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 (3), 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 (7.1.1), we obtain that $\|f\|_{\mathrm{BMO}} \leqslant 2A$.

The lower inequality in (4) follows from (3) while the upper one is trivial. (5) is immediate.

For (6), 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 (7) is a consequence of the fact that

$$\begin{aligned}\left||f(x)| - \operatorname*{Avg}_Q |f|\right| =& \left||f(x)| - \frac{1}{|Q|}\int_Q |f(t)|dt\right| = \left|\frac{1}{|Q|}\int_Q (|f(x)| - |f(t)|)dt\right| \\ \leqslant& \left|\frac{1}{|Q|}\int_Q (|f(x) - f(t)|)dt\right| \\ \leqslant& \left|\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f| + \frac{1}{|Q|}\int_Q |\operatorname*{Avg}_Q f - f(t)|dt\right| \\ \leqslant& |f - \operatorname*{Avg}_Q f| + \operatorname*{Avg}_Q |f - \operatorname*{Avg}_Q f|.\end{aligned}$$

The second and the third inequalities in (7) follow from the first one in (7) 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 (8). 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

$$\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_B f|dx \leqslant \frac{|B|}{|Q|}\frac{1}{|B|}\int_B |f(x) - \operatorname*{Avg}_B f|dx \leqslant \frac{V_n\sqrt{n^n}}{2^n}\|f\|_{\mathrm{BMO}_{\mathrm{balls}}}.$$

It follows from (3) that

$$\|f\|_{\mathrm{BMO}} \leqslant 2^{1-n}V_n\sqrt{n^n}\|f\|_{\mathrm{BMO}_{\mathrm{balls}}}.$$

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 (3) for the space $\mathrm{BMO}_{\text{balls}}$. ■

Example 7.1.4. It is trivial that any bounded function is in BMO, i.e., $L^\infty \hookrightarrow \mathrm{BMO}$. 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 function $f(x) = \ln|x|$. To check that this function is in BMO, 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 $\overline{B(0,R)} = \{x \in \mathbb{R}^n : |x - x_0| \leqslant R\}$ 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

$$\begin{aligned}&\frac{1}{V_n R^n}\int_{|x-x_0|\leqslant R} |\ln|x| - C_{x_0,R}|\,dx\\ =&\frac{1}{V_n R^n}\int_{|x-x_0|\leqslant R}\left|\ln\frac{|x|}{|x_0|}\right|dx\\ \leqslant&\max\left(\ln\frac{3}{2}, \left|\ln\frac{1}{2}\right|\right)\\ =&\ln 2,\end{aligned}$$

since $\frac{1}{2}|x_0| \leqslant |x| \leqslant \frac{3}{2}|x_0|$ when $|x - x_0| \leqslant R$ and $|x_0| > 2R$. Also, if $|x_0| \leqslant 2R$, then

$$\begin{aligned}&\frac{1}{V_n R^n}\int_{|x-x_0|\leqslant R} |\ln|x| - C_{x_0,R}|\,dx\\ =&\frac{1}{V_n R^n}\int_{|x-x_0|\leqslant R}\left|\ln\frac{|x|}{R}\right|dx \leqslant \frac{1}{V_n R^n}\int_{|x|\leqslant 3R}\left|\ln\frac{|x|}{R}\right|dx\\ =&\frac{1}{V_n}\int_{|x|\leqslant 3}|\ln|x||\,dx = \frac{\omega_{n-1}}{V_n}\int_0^3 r^{n-1}|\ln r|dr\\ =&n\int_0^1(-1)r^n\ln r\frac{dr}{r} + n\int_1^3 r^{n-1}\ln r dr\\ \leqslant&n\int_0^\infty te^{-nt}dt + n\ln 3\int_1^3 r^{n-1}dr\\ =&\frac{1}{n} + 3^n\ln 3.\end{aligned}$$

Thus, $\ln|x|$ is in BMO.

It is interesting to observe that an abrupt cutoff of a BMO function may not give a function in the same space.

Example 7.1.5. The function $h(x) = \chi_{x>0}\ln\frac{1}{x}$ is not in $\mathrm{BMO}(\mathbb{R})$. Indeed, the problem is at the origin. Consider the intervals $(-\varepsilon, \varepsilon)$, where $\varepsilon \in (0, 1/2)$. We have that

$$\operatorname*{Avg}_{(-\varepsilon,\varepsilon)} h = \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} h(x)dx = \frac{1}{2\varepsilon}\int_0^{\varepsilon}\ln\frac{1}{x}dx = \frac{1}{2}\int_0^1\left(\ln\frac{1}{\varepsilon} + \ln\frac{1}{y}\right)dy = \frac{1+\ln\frac{1}{\varepsilon}}{2}.$$

But then

$$\frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon}|h(x) - \operatorname*{Avg}_{(-\varepsilon,\varepsilon)} h|dx \geqslant \frac{1}{2\varepsilon}\int_{-\varepsilon}^{0}|\operatorname*{Avg}_{(-\varepsilon,\varepsilon)} h|dx = \frac{1+\ln\frac{1}{\varepsilon}}{4},$$

and the latter is clearly unbounded as $\varepsilon \to 0$.

A useful related fact is the following, which describes the behavior of BMO functions at infinity.

Theorem 7.1.6.

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 $Q_0 = Q(0,1)$.

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 $x \in S_k$, we have $|x| > 2^{k-2}$ and then

$$1 + |x|^{n+1} > 1 + 2^{(k-2)(n+1)} > 4^{-(n+1)} 2^{k(n+1)}.$$

Hence,

$$\begin{aligned}
I_k \leqslant & 4^{n+1} 2^{-k(n+1)} \int_{Q_k} |f(x) - \operatorname*{Avg}_{Q_0} f| dx \\
\leqslant & 4^{n+1} 2^{-k(n+1)} \int_{Q_k} [|f(x) - \operatorname*{Avg}_{Q_k} f| + |\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f|] dx \\
\leqslant & 4^{n+1} 2^{-k(n+1)} |Q_k| \left(\|f\|_{\mathrm{BMO}} + |\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f|\right) \\
= & 4^{n+1} 2^{-k(n+1)} 2^{kn} (\|f\|_{\mathrm{BMO}} + |\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f|).
\end{aligned}$$

The second term can be controlled as follows:

$$\begin{aligned}
|\operatorname*{Avg}_{Q_k} f - \operatorname*{Avg}_{Q_0} f| \leqslant & \sum_{i=1}^{k} |\operatorname*{Avg}_{Q_i} f - \operatorname*{Avg}_{Q_{i-1}} f| \\
\leqslant & \sum_{i=1}^{k} \frac{1}{|Q_{i-1}|} \int_{Q_{i-1}} |f(x) - \operatorname*{Avg}_{Q_i} f| dx \\
\leqslant & \sum_{i=1}^{k} \frac{2^n}{|Q_i|} \int_{Q_i} |f(x) - \operatorname*{Avg}_{Q_i} f| dx \\
\leqslant & k \cdot 2^n \|f\|_{\mathrm{BMO}}.
\end{aligned} \tag{7.1.3}$$

Therefore,

$$I_k \leqslant 4^{n+1} 2^{-k(n+1)} 2^{nk} (1 + k2^n) \|f\|_{\mathrm{BMO}},$$

where $C_k = Ck2^{-k}$ and $\sum_k C_k < \infty$. This completes the proof. ■

Let us now look at some basic properties of BMO functions. As the same as in (7.1.3), we observe that if a cube Q_1 is contained in a cube Q_2, then

$$\begin{aligned}|\operatorname*{Avg}_{Q_1} f - \operatorname*{Avg}_{Q_2} f| &= \left|\frac{1}{|Q_1|}\int_{Q_1} f dx - \operatorname*{Avg}_{Q_2} f\right| \leqslant \frac{1}{|Q_1|}\int_{Q_1} |f - \operatorname*{Avg}_{Q_2} f| dx \\ &\leqslant \frac{1}{|Q_1|}\int_{Q_2} |f - \operatorname*{Avg}_{Q_2} f| dx \\ &\leqslant \frac{|Q_2|}{|Q_1|}\|f\|_{\mathrm{BMO}}. \end{aligned} \tag{7.1.4}$$

The same estimate holds if the 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 $\|f\|_{\mathrm{BMO}}$ the expression given by $\|F\|_{\mathrm{BMO}_{\text{balls}}}$ in (7.1.2), 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 7.1.7.

(i) Let $f \in \mathrm{BMO}(\mathbb{R}^n)$. Given a ball B and a positive integer m, we have

$$|\operatorname*{Avg}_{B} f - \operatorname*{Avg}_{2^m B} f| \leqslant 2^n m \|f\|_{\mathrm{BMO}}. \tag{7.1.5}$$

(ii) For any $\delta > 0$, there is a constant $C_{n,\delta}$ such that for any ball $B(x_0, R)$ we have

$$R^\delta \int_{\mathbb{R}^n} \frac{|f(x) - \operatorname*{Avg}_{B(x_0,R)} f|}{(R + |x - x_0|)^{n+\delta}} dx \leqslant C_{n,\delta}\|f\|_{\mathrm{BMO}}. \tag{7.1.6}$$

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

$$\sup_{y\in\mathbb{R}^n} \sup_{t>0} \int_{\mathbb{R}^n} |f(x) - (P_t * f)(y)| P_t(x - y) dx \leqslant C_n \|f\|_{\mathrm{BMO}}. \tag{7.1.7}$$

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

$$C_n'\|f\|_{\mathrm{BMO}} \leqslant \sup_{y\in\mathbb{R}^n} \sup_{t>0} \int_{\mathbb{R}^n} |f(x) - (P_t * f)(y)| P_t(x - y) dx. \tag{7.1.8}$$

Proof. (i) We have the desired result as the same as (7.1.3).

(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

$$\begin{aligned}
&\int_{\mathbb{R}^n}\frac{|f(x)-\operatorname{Avg}_B f|}{(1+|x|)^{n+\delta}}dx\\
\leqslant&\int_B\frac{|f(x)-\operatorname{Avg}_B f|}{(1+|x|)^{n+\delta}}dx+\sum_{k=0}^\infty\int_{2^{k+1}B\setminus 2^kB}\frac{|f(x)-\operatorname{Avg}_{2^{k+1}B} f|+|\operatorname{Avg}_{2^{k+1}B} f-\operatorname{Avg}_B f|}{(1+|x|)^{n+\delta}}dx\\
\leqslant&\int_B|f(x)-\operatorname{Avg}_B f|dx+\sum_{k=0}^\infty 2^{-k(n+\delta)}\int_{2^{k+1}B}(|f(x)-\operatorname{Avg}_{2^{k+1}B} f|+|\operatorname{Avg}_{2^{k+1}B} f-\operatorname{Avg}_B f|)dx\\
\leqslant&V_n\|f\|_{\mathrm{BMO}}+\sum_{k=0}^\infty 2^{-k(n+\delta)}(1+2^n(k+1))(2^{k+1})^nV_n\|f\|_{\mathrm{BMO}}\\
=&C'_{n,\delta}\|f\|_{\mathrm{BMO}}.
\end{aligned}$$

(iii) The proof of (7.1.7) is a reprise of the argument given in (ii). Set $B_t=B(y,t)$. We first prove a version of (7.1.7) in which the expression $(P_t*f)(y)$ is replaced by $\operatorname{Avg}_{B_t} f$. For fixed y,t we have by (ii)

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}}\int_{\mathbb{R}^n}\frac{t|f(x)-\operatorname{Avg}_{B_t} f|}{(t^2+|x-y|^2)^{\frac{n+1}{2}}}dx\leqslant C''_n\|f\|_{\mathrm{BMO}}. \tag{7.1.9}$$

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 (7.1.9) yields (7.1.7) with constant $C_n=2C''_n$.

(iv) Conversely, let A be the expression on the right in (7.1.8). For $|x-y|\leqslant t$, we have $P_t(x-y)\geqslant c_nt(2t^2)^{-(n+1)/2}=c'_nt^{-n}$, 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 7.1.3 (3) now implies that

$$\|f\|_{\mathrm{BMO}}\leqslant 2A/(V_nc'_n).$$

This concludes the proof of the proposition. ■

§ 7.2 John-Nirenberg theorem

Having set down some basic facts about BMO, we now turn to a deeper property of BMO functions: their exponential integrability. We begin with a preliminary example.

Example 7.2.1. Let $f(x)=\ln|x|$, $I=(0,b)$, and

$$E_\alpha=\{x\in I:|\ln x-\operatorname{Avg}_I f|>\alpha\},$$

then we have

$$E_\alpha=\{x\in I:\ln x-\operatorname{Avg}_I f>\alpha\}\cup\{x\in I:\ln x-\operatorname{Avg}_I f<-\alpha\}$$

$$=\{x\in I : x > e^{\alpha+\operatorname*{Avg}_I f}\}\cup\{x\in I : x < e^{-\alpha+\operatorname*{Avg}_I f}\}.$$

When α is large enough, the first set is an empty set and the second one is $(0, e^{-\alpha+\operatorname*{Avg}_I f})$. Thus

$$|E_\alpha| = e^{-\alpha+\operatorname*{Avg}_I f}.$$

By Jensen's inequality, we get

$$e^{\operatorname*{Avg}_I f} \leqslant \frac{1}{|I|}\int_I e^{\ln t}dt = \frac{|I|}{2}.$$

Therefore,

$$|E_\alpha| \leqslant \frac{1}{2}|I|e^{-\alpha}.$$

Although the above relation is obtained from the function $\ln|x|$ over $(0, b)$, it indeed reflects an essential property for any BMO function in the BMO space.

Theorem 7.2.2: The John-Nirenberg theorem

For all $f\in \mathrm{BMO}(\mathbb{R}^n)$, for all cubes Q, and all $\alpha>0$, we have

$$\left|\{x\in Q : |f(x)-\operatorname*{Avg}_Q f|>\alpha\}\right| \leqslant e|Q|e^{-A\alpha/\|f\|_{\mathrm{BMO}}} \tag{7.2.1}$$

with $A=(2^n e)^{-1}$.

Proof. Since inequality (7.2.1) is not altered when we multiply both f and α by the same constant, it suffices to assume that $\|f\|_{\mathrm{BMO}}=1$. Let us now fix a closed cube Q and a constant $b>1$ to be chosen later.

We apply the Calderón-Zygmund decomposition to the function $f-\operatorname*{Avg}_Q f$ inside the cube Q. We introduce the following selection criterion for a cube R:

$$\frac{1}{|R|}\int_R |f(x)-\operatorname*{Avg}_Q f|dx > b. \tag{7.2.2}$$

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 (7.2.2). 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 (7.2.2). 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 (7.2.2). Continue this process indefinitely. We obtain a countable collection of cubes $\{Q_j^{(1)}\}_j$ satisfying the following properties:

(A-1) The interior of every $Q_j^{(1)}$ is contained in $Q^{(0)}$.

(B-1) $b < \left|Q_j^{(1)}\right|^{-1}\int_{Q_j^{(1)}}|f(x)-\operatorname*{Avg}_{Q^{(0)}} f|dx \leqslant 2^n b$.

(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 first generation. Note that the second inequality in (D-1) requires (B-1) and the fact that $Q^{(0)}$ does not satisfy (7.2.2). We now fix a selected first-generation cube $Q_j^{(1)}$ and we introduce the following selection criterion for a cube R:

$$\frac{1}{|R|}\int_R |f(x) - \operatorname*{Avg}_{Q_j^{(1)}} f|dx > b. \tag{7.2.3}$$

Observe that $Q_j^{(1)}$ does not satisfy the selection criterion (7.2.3). We apply a similar Calderón-Zygmund decomposition to the function

$$f - \operatorname*{Avg}_{Q_j^{(1)}} f$$

inside the cube $Q_j^{(1)}$. Subdivide $Q_j^{(1)}$ into 2n equal closed subcubes of side length equal to half of the side length of $Q_j^{(1)}$ by bisecting the sides, and select such a subcube R if it satisfies the selection criterion (7.2.3). Continue this process indefinitely. Also repeat this process for any other cube $Q_j^{(1)}$ of the first generation. We obtain a collection of cubes $\{Q_l^{(2)}\}_l$ of second generation each contained in some $Q_j^{(1)}$ such that versions of (A-1)-(E-1) are satisfied, with the superscript (2) replacing (1) and the superscript (1) replacing (0). We use the superscript (k) to denote the generation of the selected cubes.

For a fixed selected cube $Q_l^{(2)}$ of second generation, introduce the selection criterion

$$\frac{1}{|R|}\int_R |f(x) - \operatorname*{Avg}_{Q_l^{(2)}} f|dx > b. \tag{7.2.4}$$

and repeat the previous process to obtain a collection of cubes of third generation inside $Q_l^{(2)}$. Repeat this procedure for any other cube $Q_l^{(2)}$ of the second generation. Denote by $\{Q_s^{(3)}\}_s$ the thus obtained collection of all cubes of the third generation.

We iterate this procedure indefinitely to obtain a doubly indexed family of cubes $Q_j^{(k)}$ satisfying the following properties:

(A-k) The interior of every $Q_j^{(k)}$ is contained in $Q_{j'}^{(k-1)}$.

(B-k) $b < \left|Q_j^{(k)}\right|^{-1} \int_{Q_j^{(k)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f|dx \leqslant 2^n b$.

(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 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 prove (D-k). We have

$$\begin{aligned}
\sum_j \left|Q_j^{(k)}\right| <& \frac{1}{b}\sum_j \int_{Q_j^{(k)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f| dx \\
=& \frac{1}{b}\sum_{j'} \sum_{j \text{ corresp. to } j'} \int_{Q_j^{(k)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f| dx \\
\leqslant& \frac{1}{b}\sum_{j'} \int_{Q_{j'}^{(k-1)}} |f(x) - \operatorname*{Avg}_{Q_{j'}^{(k-1)}} f| dx \\
\leqslant& \frac{1}{b}\sum_{j'} \left|Q_{j'}^{(k-1)}\right| \|f\|_{\mathrm{BMO}} \\
=& \frac{1}{b}\sum_{j'} \left|Q_{j'}^{(k-1)}\right|.
\end{aligned}$$

Having established (A-k)-(E-k) we turn to some consequences. Applying (D-k) successively $k-1$ times, we obtain

$$\sum_j \left|Q_j^{(k)}\right| \leqslant b^{-k}\left|Q^{(0)}\right|. \tag{7.2.5}$$

For any fixed j we have that $|\operatorname*{Avg}_{Q_j^{(1)}} f - \operatorname*{Avg}_{Q^{(0)}} f| \leqslant 2^n b$ and $|f - \operatorname*{Avg}_{Q_j^{(1)}} f| \leqslant b$ a.e. on $Q_j^{(1)} \setminus \cup_l Q_l^{(2)}$. This gives

$$|f - \operatorname*{Avg}_{Q^{(0)}} f| \leqslant 2^n b + b \quad \text{a.e. on } Q_j^{(1)} \setminus \cup_l Q_l^{(2)},$$

which, combined with (E-1), yields

$$|f - \operatorname*{Avg}_{Q^{(0)}} f| \leqslant 2^n 2b \quad \text{a.e. on } Q^{(0)} \setminus \cup_l Q_l^{(2)}. \tag{7.2.6}$$

For every fixed l, we also have that $|f - \operatorname*{Avg}_{Q_l^{(2)}} f| \leqslant b$ a.e. on $Q_l^{(2)} \setminus \cup_s Q_s^{(3)}$, which combined with $|\operatorname*{Avg}_{Q_l^{(2)}} f - \operatorname*{Avg}_{Q_{l'}^{(1)}} f| \leqslant 2^n b$ and $|\operatorname*{Avg}_{Q_{l'}^{(1)}} f - \operatorname*{Avg}_{Q^{(0)}} f| \leqslant 2^n b$ 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 (7.2.6), 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

$$|f - \operatorname*{Avg}_{Q^{(0)}} f| \leqslant 2^n kb \quad \text{a.e. on } Q^{(0)} \setminus \cup_s Q_s^{(k)}. \tag{7.2.7}$$

This proves the almost everywhere inclusion

$$\left\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > 2^n kb\right\} \subset \cup_j Q_j^{(k)}$$

for all $k = 1, 2, 3, \cdots$. (This also holds when $k = 0$.) We now use (7.2.5) and (7.2.7) to prove (7.2.1). We fix an $\alpha > 0$. If

$$2^n kb < \alpha \leqslant 2^n (k+1)b$$

for some $k \geqslant 0$, then

$$\begin{aligned}\left|\left\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > \alpha\right\}\right| \leqslant &\left|\left\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > 2^n kb\right\}\right| \\ \leqslant &\sum_j \left|Q_j^{(k)}\right| \leqslant \frac{1}{b^k}\left|Q^{(0)}\right| \\ = &|Q|e^{-k\ln b} \\ \leqslant &|Q|be^{-\alpha \ln b/(2^n b)},\end{aligned}$$

since $-k \leqslant 1 - \frac{\alpha}{2^n b}$. Choosing $b = e > 1$ yields (7.2.1). ■

Having proved the important distribution inequality (7.2.1), we are now in a position to deduce from it a few corollaries.

Corollary 7.2.3.

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

$$\frac{1}{|Q|}\int_Q e^{\gamma|f(x) - \operatorname*{Avg}_Q f|/\|f\|_{\mathrm{BMO}}} dx \leqslant 1 + \frac{2^n e^2 \gamma}{1 - 2^n e\gamma}.$$

Proof. Using identity (1.1.2) with $\varphi(t) = e^t - 1$, we write

$$\begin{aligned}\frac{1}{|Q|}\int_Q e^h dx = &1 + \frac{1}{|Q|}\int_Q (e^h - 1)dx \\ = &1 + \frac{1}{|Q|}\int_0^\infty e^\alpha \left|\{x \in Q : |h(x)| > \alpha\}\right| d\alpha\end{aligned}$$

for a measurable function h. Then we take $h = \gamma|f(x) - \operatorname*{Avg}_Q f|/\|f\|_{\mathrm{BMO}}$ and we use inequality (7.2.1) with $\gamma < A = (2^n e)^{-1}$ to obtain

$$\begin{aligned}&\frac{1}{|Q|}\int_Q e^{\gamma|f(x) - \operatorname*{Avg}_Q f|/\|f\|_{\mathrm{BMO}}} dx \\ \leqslant &1 + \int_0^\infty e^\alpha e e^{-A(\frac{\alpha}{\gamma}\|f\|_{\mathrm{BMO}})/\|f\|_{\mathrm{BMO}}} d\alpha \\ = &1 + e\int_0^\infty e^{\alpha(1 - 1/(2^n e\gamma))} d\alpha = 1 + \frac{2^n e^2 \gamma}{1 - 2^n e\gamma},\end{aligned}$$

thus, we complete the proof. ■

Another important corollary is the following.

Corollary 7.2.4.

For all $1 < p < \infty$, there exits a finite constant $B_{p,n}$ such that

$$\sup_Q \left(\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f|^p dx\right)^{1/p} \leqslant B_{p,n}\|f\|_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{7.2.8}$$

Proof. This result can be obtained from the one in the preceding corollary or directly in the following way:

$$\frac{1}{|Q|}\int_Q |f(x) - \operatorname*{Avg}_Q f|^p dx = \frac{p}{|Q|}\int_0^\infty \alpha^{p-1}\left|\{x \in Q : |f(x) - \operatorname*{Avg}_Q f| > \alpha\}\right| d\alpha$$

$$\leqslant \frac{p}{|Q|} e|Q| \int_0^\infty \alpha^{p-1} e^{-A\alpha/\|f\|_{\mathrm{BMO}}} d\alpha$$
$$= p\Gamma(p) \frac{e}{A^p} \|f\|_{\mathrm{BMO}}^p,$$

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 7.2.4 can be reversed via Hölder's inequality, we obtain the following important L^p characterization of BMO norms.

Corollary 7.2.5.

For all $1 < p < \infty$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, we have

$$\sup_Q \left(\frac{1}{|Q|} \int_Q |f(x) - \operatorname{Avg}_Q f|^p dx \right)^{1/p} \approx \|f\|_{\mathrm{BMO}}. \tag{7.2.9}$$

Proof. One direction follows from Corollary 7.2.4. Conversely, the supremum in (7.2.9) is bigger than or equal to the corresponding supremum with $p = 1$, which is equal to the BMO norm of f, by definition. ■

§7.3 Non-tangential maximal functions and Carleson measures

We recall the definition of a cone over a point given in Definition 5.1.9.

Definition 7.3.1: Cone

Let $x \in \mathbb{R}^n$. We define the cone over x:

$$\Gamma(x) = \{(y,t) \in \mathbb{R}^{n+1}_+ : |x - y| < t\}.$$

Definition 7.3.2: Non-tangential maximal function

Let $F : \mathbb{R}^{n+1}_+ \to \mathbb{C}$ and define the non-tangential maximal function of F:

$$M^*F(x) = \sup_{(y,t)\in\Gamma(x)} |F(y,t)| \in [0,\infty].$$

Remark 7.3.3. (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 measure, a contradiction.

(ii) Given a Borel measure μ on $\mathbb{R}^{n+1}_+$, we can define the non-tangential maximal function M^*_μ w.r.t. μ by replacing sup with ess sup. Note then that M^*_μ is defined μ-a.e.

Proposition 7.3.4.

M^*F is lower semi-continuous and hence Borel.

Proof. Let $\alpha \geqslant 0$ and $x \in \mathbb{R}^n$ such that $M^*F(x) > \alpha$. Now, there exists a $(y,t) \in \Gamma(x)$ such that $|F(y,t)| > \alpha$. Therefore, for all $z \in B(y,t)$, we have $(y,t) \in \Gamma(z)$ and hence $M^*F(z) \geqslant |F(y,t)| > \alpha$. That is, $x \in B(y,t) \subset \{x \in \mathbb{R}^n : M^*F(x) > \alpha\}$. ■

Definition 7.3.5: 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 7.3.6: 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 7.3.7. 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.

The following measures are not Carleson measures.

Example 7.3.8. (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

$$\mu(B \times (0,r]) = |B| \int_0^r \frac{dt}{t^\alpha} = \begin{cases} |B|\dfrac{r^{1-\alpha}}{1-\alpha}, & 1-\alpha > 0, \\ \infty, & \text{otherwise.} \end{cases}$$

So we only need to consider the case $\alpha < 1$ but in this case, we cannot get uniform control via a constant C.

The following are examples of Carleson measures.

Example 7.3.9. (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 $[x_0 - r, x_0 + r] \times (0, r]$ is at most equal to the diagonal of the box, i.e., $\sqrt{5}r$.

Definition 7.3.10: 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 $\|\mathscr{C}(\mu)\|_\infty = \|\mu\|_{\mathscr{C}}$.

Theorem 7.3.11: Carleson's 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

$$\mu(\{(x,t) \in \mathbb{R}^{n+1}_+ : |F(x,t)| > \alpha\}) \leqslant C_n \int_{\{M^*F>\alpha\}} \mathscr{C}(\mu)(x)dx. \tag{7.3.1}$$

In particular, if μ is a Carleson measure, then

$$\mu(\{|F| > \alpha\}) \leqslant C_n\|\mu\|_{\mathscr{C}}|\{M^*F > \alpha\}|. \tag{7.3.2}$$

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 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 $|F(y_0,t_0)| > \alpha$. If d_0 is the distance from (y_0,t_0) to the sphere formed by the intersection of the hyperplane $t_0 + \mathbb{R}^n$ 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 $|z - y_0| \leqslant |z - x_0| + |x_0 - y_0| < d_0 + t_0 - d_0 = t_0$, hence $M^*F(z) \geqslant |F(y_0,t_0)| > \alpha$.

Let $\{Q_k\}$ be the Whitney decomposition (i.e., Theorem 3.1.2) 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

$$\delta_\alpha(z) \leqslant \sqrt{n}\ell(Q_k) + \operatorname{dist}(Q_k,\Omega_\alpha^c) \leqslant 5\sqrt{n}\ell(Q_k) \tag{7.3.3}$$

in view of Theorem 3.1.2 (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 (7.3.3) to obtain that for any $z \in Q_k$ and $y \in B(z,\delta_\alpha(z))$

$$|y - z_0| \leqslant |y - z| + |z - z_0| \leqslant \delta_\alpha(z) + \sqrt{n}\ell(Q_k)/2 \leqslant \frac{11}{2}\sqrt{n}\ell(Q_k) \leqslant 11\text{rad}(B_k),$$

namely,

$$z \in Q_k \Longrightarrow B(z,\delta_\alpha(z)) \subset 11B_k.$$

This implies that

$$\bigcup_{z\in\Omega_\alpha} T(B(z,\delta_\alpha(z))) \subset \bigcup_k T(11B_k). \tag{7.3.4}$$

Next, we claim that

$$\{|F| > \alpha\} \subset \bigcup_{z\in\Omega_\alpha} T(B(z,\delta_\alpha(z))). \tag{7.3.5}$$

Indeed, let $(x,t) \in \mathbb{R}^{n+1}_+$ 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 (7.3.5).

Combining (7.3.4) and (7.3.5), 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 (7.3.1). It follows (7.3.2) in view of $\|\mathscr{C}(\mu)\|_\infty = \|\mu\|_{\mathscr{C}}$. ∎

Corollary 7.3.12.

For any Carleson measure μ and every μ-measurable function F on $\mathbb{R}^{n+1}_+$, we have

$$\int_{\mathbb{R}^{n+1}_+} |F(x,t)|^p d\mu(x,t) \leqslant C_n\|\mu\|_{\mathscr{C}} \int_{\mathbb{R}^n} (M^*F(x))^p dx \tag{7.3.6}$$

for all $p\in[1,\infty)$.

Proof. From (7.3.2), applying Theorem 1.1.4 twice, we get

$$\begin{aligned}\int_{\mathbb{R}^{n+1}_+} |F(x,t)|^p d\mu(x,t) =&p\int_0^\infty \alpha^{p-1}\mu(\{|F|>\alpha\})d\alpha\\ \leqslant &C_n\|\mu\|_{\mathscr{C}} p\int_0^\infty \alpha^{p-1}|\{M^*F>\alpha\}|d\alpha\\ =&C_n\|\mu\|_{\mathscr{C}}\int_{\mathbb{R}^n}(M^*F(x))^p dx.\end{aligned}$$

∎

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 7.3.13.

Let Φ be a function on $\mathbb{R}^n$ that satisfies for some $0<C,\delta<\infty$,

$$|\Phi(x)| \leqslant \frac{C}{(1+|x|)^{n+\delta}}. \tag{7.3.7}$$

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

$$\int_{\mathbb{R}^{n+1}_+} |(\Phi_t*f)(x)|^p d\mu(x,t) \leqslant C_{p,n}(\mu)\int_{\mathbb{R}^n}|f(x)|^p dx, \tag{7.3.8}$$

where $C_{p,n}(\mu)\leqslant C(p,n)\|\mu\|_{\mathscr{C}}$.

Conversely, suppose that Φ is a non-negative function that satisfies (7.3.7) 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 (7.3.8) 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 (7.3.8) as a sequence of Corollary 7.3.12. Indeed, for $F(x,t) = (\Phi_t * f)(x)$, we have

$$\begin{aligned}
M^*F(x) =& \sup_{t>0} \sup_{\substack{y\in\mathbb{R}^n \\ |y-x|<t}} |(\Phi_t * f)(y)| \\
\leqslant& \sup_{t>0} \sup_{\substack{y\in\mathbb{R}^n \\ |y-x|<t}} \int_{\mathbb{R}^n} |\Phi_t(y-z)||f(z)|dz \\
=& \sup_{t>0} \sup_{\substack{y\in\mathbb{R}^n \\ |y-x|<t}} \int_{\mathbb{R}^n} |\Phi_t(y-x+x-z)||f(z)|dz \\
\leqslant& \sup_{t>0} \left(\sup_{\substack{y\in\mathbb{R}^n \\ |y-x|<t}} |\Phi_t(y-x+\cdot)| * |f| \right)(x) \\
=& \sup_{t>0} (\Psi_t * |f|)(x),
\end{aligned}$$

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 the condition (7.3.7). Thus, it is clear that $\|\Psi\|_{L^1(\mathbb{R}^n)} \leqslant C(V_n + \omega_{n-1}/\delta)$. It follows from Theorem 3.2.12 that

$$M^*F(x) \leqslant C(n,\delta) M(|f|)(x).$$

Then, by Theorem 3.2.7, 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 7.3.12, (7.3.8) follows.

Conversely, if (7.3.8) 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

$$\begin{aligned}
\mu(T(B)) =& \int_{T(B)} d\mu(x,t) \leqslant \frac{1}{c_n^p} \int_{T(B)} c_n^p d\mu(x,t) \\
\leqslant& \frac{1}{c_n^p} \int_{T(B)} |(\Phi_t * \chi_{2B})(x)|^p d\mu(x,t) \\
\leqslant& \frac{1}{c_n^p} \int_{\mathbb{R}^{n+1}_+} |(\Phi_t * \chi_{2B})(x)|^p d\mu(x,t) \\
\leqslant& \frac{C_{p,n}(\mu)}{c_n^p} \int_{\mathbb{R}^n} |\chi_{2B}(x)|^p d\mu(x,t)
\end{aligned}$$

$$=\frac{2^n C_{p,n}(\mu)}{c_n^p}|B|.$$

This proves that μ is a Carleson measure with $\|\mu\|_{\mathscr{C}}$. ■

§7.4 BMO functions and Carleson measures

We now turn to an interesting connection between BMO functions and Carleson measures. We have the following.

Theorem 7.4.1.

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

$$|\Psi(x)| \leqslant A(1+|x|)^{-n-\delta} \tag{7.4.1}$$

for some $0 < A, \delta < \infty$. Consider the dilation $\Psi_t = t^{-n}\Psi(t^{-1}x)$ and define the Littlewood-Paley operators $\dot{\Delta}_j f = f * \Psi_{2^{-j}}$.

(i) Suppose that

$$\sup_{\xi\in\mathbb{R}^n} \sum_{j\in\mathbb{Z}} |\widehat{\Psi}(2^{-j}\xi)|^2 \leqslant B^2 < \infty \tag{7.4.2}$$

and let $\delta_{2^{-j}}(t)$ be 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 $C_{n,\delta}(A+B)^2\|b\|^2_{\mathrm{BMO}}$.

(ii) Suppose that

$$\sup_{\xi\in\mathbb{R}^n} \int_0^\infty |\widehat{\Psi}(t\xi)|^2 \frac{dt}{t} \leqslant B^2 < \infty. \tag{7.4.3}$$

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

$$|K_t(x,y)| \leqslant \frac{At^\delta}{(t+|x-y|)^{n+\delta}} \tag{7.4.4}$$

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

$$\int_0^\infty \int_{\mathbb{R}^n} |R_t(f)(x)|^2 \frac{dxdt}{t} \leqslant B^2 \|f\|^2_{L^2(\mathbb{R}^n)} \tag{7.4.5}$$

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

$$\int_{\mathbb{R}^{n+1}_+} F(x,t)d\mu(x,t) = \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} |(\Psi_{2^{-j}} * b)(x)|^2 F(x, 2^{-j})dx, \tag{7.4.6}$$

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 mean value zero, $\dot\Delta_j \operatorname*{Avg}_Q b = \Psi_{2^{-j}} * \operatorname*{Avg}_Q b = 0$. Then (7.4.6) gives

$$\mu(T(Q)) = \sum_{2^{-j}\leqslant \ell(Q)} \int_Q |\dot\Delta_j b(x)|^2 dx \leqslant 2\Sigma_1 + 2\Sigma_2,$$

where

$$\begin{aligned}
\Sigma_1 &= \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} |\dot\Delta_j((b - \operatorname*{Avg}_Q b)\chi_{Q^*})(x)|^2 dx,\\
\Sigma_2 &= \sum_{2^{-j}\leqslant \ell(Q)} \int_Q |\dot\Delta_j((b - \operatorname*{Avg}_Q b)\chi_{(Q^*)^c})(x)|^2 dx.
\end{aligned}$$

Using Plancherel's theorem twice and (7.4.2), we obtain

$$\begin{aligned}
\Sigma_1 \leqslant& \sup_\xi \sum_{j\in\mathbb{Z}} |\widehat{\Psi}(2^{-j}\xi)|^2 \int_{\mathbb{R}^n} |\widehat{((b - \operatorname*{Avg}_Q b)\chi_{Q^*})}(\eta)|^2 d\eta\\
\leqslant& B^2 \int_{Q^*} |b(x) - \operatorname*{Avg}_Q b|^2 dx\\
\leqslant& 2B^2 \int_{Q^*} |b(x) - \operatorname*{Avg}_{Q^*} b|^2 dx + 2B^2 |Q^*| |\operatorname*{Avg}_{Q^*} b - \operatorname*{Avg}_Q b|^2\\
\leqslant& 2B^2 \int_{Q^*} |b(x) - \operatorname*{Avg}_{Q^*} b|^2 dx + 2B^2 c_n \|b\|^2_{\mathrm{BMO}} |Q|\\
\leqslant& C_n B^2 c_n \|b\|^2_{\mathrm{BMO}} |Q|,
\end{aligned}$$

in view of Proposition 7.1.7 (i) and Corollary 7.2.4. To estimate Σ_2, we use the size estimate (7.4.1) of the function Ψ to obtain

$$|(\Psi_{2^{-j}} * (b - \operatorname*{Avg}_Q b)\chi_{(Q^*)^c})(x)| \leqslant \int_{(Q^*)^c} \frac{A2^{-j\delta}|b(y) - \operatorname*{Avg}_Q b|}{(2^{-j} + |x-y|)^{n+\delta}} dy. \tag{7.4.7}$$

Denote c_Q the center of Q, then for $x \in Q$ and $y \in (Q^*)^c$, we get

$$\begin{aligned}
2^{-j} + |x-y| \geqslant& |y-x| \geqslant |y - c_Q| - |c_Q - x|\\
\geqslant& \frac{1}{2}|c_Q - y| + \frac{3\sqrt{n}}{4}\ell(Q) - |c_Q - x| \; (\because |y - c_Q| \geqslant \frac{1}{2}\ell(Q^*) = \frac{3\sqrt{n}}{2}\ell(Q))\\
\geqslant& \frac{1}{2}|c_Q - y| + \frac{3\sqrt{n}}{4}\ell(Q) - \frac{\sqrt{n}}{2}\ell(Q)\\
=& \frac{1}{2}\left(|c_Q - y| + \frac{\sqrt{n}}{2}\ell(Q)\right).
\end{aligned}$$

Inserting this estimate in (7.4.7), integrating over Q, and summing over j with $2^{-j} \leqslant \ell(Q)$, we obtain

$$\begin{aligned}\Sigma_2 \leqslant & C_n \sum_{j:2^{-j}\leqslant \ell(Q)} 2^{-2j\delta} \int_Q \left(A \int_{\mathbb{R}^n} \frac{|b(y) - \operatorname{Avg}_Q b|}{(\ell(Q) + |c_Q - y|)^{n+\delta}} dy \right)^2 dx \\ \leqslant & C_n A^2 |Q| \left(A \int_{\mathbb{R}^n} \frac{\ell(Q)^\delta |b(y) - \operatorname{Avg}_Q b|}{(\ell(Q) + |c_Q - y|)^{n+\delta}} dy \right)^2 \\ \leqslant & C_{n,\delta} A^2 |Q| \|b\|_{\mathrm{BMO}}^2\end{aligned}$$

in view of (7.1.6). This proves that

$$\Sigma_1 + \Sigma_2 \leqslant C_{n,\delta}(A^2 + B^2)|Q|\|b\|_{\mathrm{BMO}}^2,$$

which implies that $\mu(T(Q)) \leqslant C_{n,\delta}(A^2 + B^2)|Q|\|b\|_{\mathrm{BMO}}^2$.

(ii) The proof can be obtained as similar fashion as in (i).

(iii) This is a generalization of (ii) and is proved 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 $R_t(\operatorname{Avg}_Q b) = 0$. We handle the term containing $R_t((b - \operatorname{Avg}_Q b)\chi_{Q^*})$ using an L^2 estimate over Q^* and condition (7.4.5), while for the term containing $R_t((b - \operatorname{Avg}_Q b)\chi_{(Q^*)^c})$, we use an L^1 estimate and condition (7.4.4). In both cases, we obtain the required conclusion in a way analogous to that in (i). ■

References

[AB12] Pascal Auscher and Lashi Bandara. *Real Harmonic Analysis*. ANU eView, 2012.

[Gra14] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[JN61] F. John and L. Nirenberg. On functions of bounded mean oscillation. *Comm. Pure Appl. Math.*, 14:415–426, 1961.

[Ste93] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

Index