Chapter 2
Maximal Functions, Fourier Transform, and Distributions

We have already seen that the convolution of a function with a fixed density is a smoothing operation that produces a certain average of the function. Averaging is an important operation in analysis and naturally arises in many situations. The study of averages of functions is better understood by the introduction of the maximal function which is defined as the largest average of a function over all balls containing a fixed point. Maximal functions are used to obtain almost everywhere convergence for certain integral averages and play an important role in this area, which is called differentiation theory. Although maximal functions do not preserve qualitative information about the given functions, they maintain crucial quantitative information, a fact of great importance in the subject of Fourier analysis.

Another important operation we study in this chapter is the Fourier transform, the father of all oscillatory integrals. This is as fundamental to Fourier analysis as marrow is to the human bone. It is a powerful transformation that carries a function from its spatial domain to its frequency domain. By doing this, it inverts the function's localization properties. If applied one more time, then magically reproduces the function composed with a reflection. It changes convolution to multiplication, translation to modulation, and expanding dilation to shrinking dilation. Its decay at infinity encodes information about the local smoothness of the function. The study of the Fourier transform also motivates the launch of a thorough study of general oscillatory integrals. We take a quick look at this topic with emphasis on one-dimensional results.

Distributions suppy a mathematical framework for many operations that do not exactly qualify to be called functions. These operations found their mathematical place in the world of functionals applied to smooth functions (called test functions). These functionals also introduced the correct interpretation for many physical objects, such as the Dirac delta function. Distributions have become an indispensable tool in analysis and have enhanced our perspective.

L. Grafakos, *Classical Fourier Analysis*, Graduate Texts in Mathematics 249,
DOI 10.1007/978-1-4939-1194-3_2, © Springer Science+Business Media New York 2014

2.1 Maximal Functions

Given a Lebesgue measurable subset A of $\mathbf{R}^n$, we denote by $|A|$ its Lebesgue measure. For $x \in \mathbf{R}^n$ and $r > 0$, we denote by $B(x,r)$ the open ball of radius r centered at x. We also use the notation $aB(x,\delta) = B(x,a\delta)$, for $a > 0$, for the ball with the same center and radius $a\delta$. Given $\delta > 0$ and f a locally integrable function on $\mathbf{R}^n$, let

$$\operatorname*{Avg}_{B(x,\delta)} |f| = \frac{1}{|B(x,\delta)|} \int_{B(x,\delta)} |f(y)|\, dy$$

denote the average of $|f|$ over the ball of radius δ centered at x.

2.1.1 The Hardy–Littlewood Maximal Operator

Definition 2.1.1. Let f be a locally integrable function on $\mathbf{R}^n$. The function

$$\mathcal{M}(f)(x) = \sup_{\delta>0} \operatorname*{Avg}_{B(x,\delta)} |f| = \sup_{\delta>0} \frac{1}{v_n \delta^n} \int_{|y|<\delta} |f(x-y)|\, dy$$

is called the *centered Hardy–Littlewood maximal function* of f.

Obviously we have $\mathcal{M}(f) = \mathcal{M}(|f|) \geq 0$; thus the maximal function is a positive operator. Information concerning cancellation of the function f is lost by passing to $\mathcal{M}(f)$. We show later that $\mathcal{M}(f)$ pointwise controls f (i.e., $\mathcal{M}(f) \geq |f|$ almost everywhere). Note that $\mathcal{M}$ maps L^∞ to itself, that is, we have

$$\big\|\mathcal{M}(f)\big\|_{L^\infty} \leq \big\|f\big\|_{L^\infty}.$$

Let us compute the Hardy–Littlewood maximal function of a specific function.

Example 2.1.2. On $\mathbf{R}$, let f be the characteristic function of the interval $[a,b]$. For $x \in (a,b)$, clearly $\mathcal{M}(f) = 1$. For $x \geq b$, a simple calculation shows that the largest average of f over all intervals $(x-\delta, x+\delta)$ is obtained when $\delta = x-a$. Similarly, when $x \leq a$, the largest average is obtained when $\delta = b-x$. Therefore,

$$\mathcal{M}(f)(x) = \begin{cases} (b-a)/2|x-b| & \text{when } x \leq a, \\ 1 & \text{when } x \in (a,b), \\ (b-a)/2|x-a| & \text{when } x \geq b. \end{cases}$$

Observe that $\mathcal{M}(f)$ has a jump at $x = a$ and $x = b$ equal to one-half that of f.

$\mathcal{M}$ is a sublinear operator, i.e., it satisfies $\mathcal{M}(f+g) \leq \mathcal{M}(f) + \mathcal{M}(g)$ and $\mathcal{M}(\lambda f) = |\lambda|\mathcal{M}(f)$ for all locally integrable functions f and g and all complex constants λ. It also has some interesting properties:

If f is locally integrable, then by considering the average of f over the ball $B(x,|x|+R)$, which contains the ball $B(0,R)$, we obtain

$$\mathcal{M}(f)(x) \geq \frac{\int_{B(0,R)} |f(y)|\,dy}{v_n(|x|+R)^n}, \tag{2.1.1}$$

for all $x \in \mathbf{R}^n$, where v_n is the volume of the unit ball in $\mathbf{R}^n$. An interesting consequence of (2.1.1) is the following: suppose that $f \neq 0$ on a set of positive measure E, then $\mathcal{M}(f)$ is not in $L^1(\mathbf{R}^n)$. In other words, if f is in $L^1_{\mathrm{loc}}(\mathbf{R}^n)$ and $\mathcal{M}(f)$ is in $L^1(\mathbf{R}^n)$, then $f=0$ a.e. To see this, integrate (2.1.1) over the ball $\mathbf{R}^n$ to deduce that $\|f\chi_{B(0,R)}\|_{L^1}=0$ and thus $f(x)=0$ for almost all x in the ball $B(0,R)$. Since this is valid for all $R=1,2,3,\dots$, it follows that $f=0$ a.e. in $\mathbf{R}^n$.

Another remarkable locality property of $\mathcal{M}$ is that if $\mathcal{M}(f)(x_0)=0$ for some x_0 in $\mathbf{R}^n$, then $f=0$ a.e. To see we take $x=x_0$ in (2.1.1) to deduce that $\|f\chi_{B(0,R)}\|_{L^1}=0$ and as before we have that $f=0$ a.e. on every ball centered at the origin, i.e., $f=0$ a.e. in $\mathbf{R}^n$.

A related analogue of $\mathcal{M}(f)$ is its uncentered version $M(f)$, defined as the supremum of all averages of f over all open balls containing a given point.

Definition 2.1.3. The *uncentered Hardy–Littlewood maximal function* of f,

$$M(f)(x) = \sup_{\substack{\delta>0 \\ |y-x|<\delta}} \operatorname*{Avg}_{B(y,\delta)} |f|,$$

is defined as the supremum of the averages of $|f|$ over all open balls $B(y,\delta)$ that contain the point x.

Clearly $\mathcal{M}(f) \leq M(f)$; in other words, M is a larger operator than $\mathcal{M}$. However, $M(f) \leq 2^n \mathcal{M}(f)$ and the boundedness properties of M are identical to those of $\mathcal{M}$.

Example 2.1.4. On $\mathbf{R}$, let f be the characteristic function of the interval $I=[a,b]$. For $x \in (a,b)$, clearly $M(f)(x)=1$. For $x>b$, a calculation shows that the largest average of f over all intervals $(y-\delta, y+\delta)$ that contain x is obtained when $\delta = \frac{1}{2}(x-a)$ and $y=\frac{1}{2}(x+a)$. Similarly, when $x<a$, the largest average is obtained when $\delta=\frac{1}{2}(b-x)$ and $y=\frac{1}{2}(b+x)$. We conclude that

$$M(f)(x) = \begin{cases} (b-a)/|x-b| & \text{when } x \leq a, \\ 1 & \text{when } x \in (a,b), \\ (b-a)/|x-a| & \text{when } x \geq b. \end{cases}$$

Observe that M does not have a jump at $x=a$ and $x=b$ and is in fact equal to the function $\left(1+\frac{\operatorname{dist}(x,I)}{|I|}\right)^{-1}$.

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

Lemma 2.1.5. *Let $\{B_1, B_2, \dots, B_k\}$ be a finite collection of open balls in $\mathbf{R}^n$. Then there exists a finite subcollection $\{B_{j_1}, \dots, B_{j_l}\}$ of pairwise disjoint balls such that*

$$\sum_{r=1}^{l} |B_{j_r}| \geq 3^{-n} \Big| \bigcup_{i=1}^{k} B_i \Big| . \tag{2.1.2}$$

Proof. Let us reindex the balls so that

$$|B_1| \geq |B_2| \geq \cdots \geq |B_k| .$$

Let $j_1 = 1$. Having chosen $j_1, j_2, \dots, j_i$, let j_{i+1} be the least index $s > j_i$ such that $\bigcup_{m=1}^{i} B_{j_m}$ is disjoint from B_s. Since we have a finite number of balls, this process will terminate, say after l steps. We have now selected pairwise disjoint balls $B_{j_1}, \dots, B_{j_l}$. If some B_m was not selected, that is, $m \notin \{j_1, \dots, j_l\}$, then B_m must intersect a selected ball B_{j_r} for some $j_r < m$. Then B_m has smaller size than B_{j_r} and we must have $B_m \subseteq 3B_{j_r}$. This shows that the union of the unselected balls is contained in the union of the triples of the selected balls. Therefore, the union of all balls is contained in the union of the triples of the selected balls. Thus

$$\Big| \bigcup_{i=1}^{k} B_i \Big| \leq \Big| \bigcup_{r=1}^{l} 3B_{j_r} \Big| \leq \sum_{r=1}^{l} |3B_{j_r}| = 3^n \sum_{r=1}^{l} |B_{j_r}| ,$$

and the required conclusion follows. □

It was noted earlier that $\mathcal{M}(f)$ and $M(f)$ never map into L^1. However, it is true that these functions are in $L^{1,\infty}$ when f is in L^1. Operators that map L^1 to $L^{1,\infty}$ are said to be *weak type* $(1,1)$. The centered and uncentered maximal functions $\mathcal{M}$ and M are of weak type $(1,1)$ as shown in the next theorem.

Theorem 2.1.6. *The uncentered and centered Hardy–Littlewood maximal operators M and $\mathcal{M}$ map $L^1(\mathbf{R}^n)$ to $L^{1,\infty}(\mathbf{R}^n)$ with constant at most 3^n and also $L^p(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n)$ for $1 < p < \infty$ with constant at most $3^{n/p} p(p-1)^{-1}$. For any $f \in L^1(\mathbf{R}^n)$ we also have*

$$\big| \{ M(f) > \alpha \} \big| \leq \frac{3^n}{\alpha} \int_{\{M(f) > \alpha\}} |f(y)| \, dy . \tag{2.1.3}$$

Proof. We claim that the set $E_\alpha = \{x \in \mathbf{R}^n : M(f)(x) > \alpha\}$ is open. Indeed, for $x \in E_\alpha$, there is an open ball B_x that contains x such that the average of $|f|$ over B_x is strictly bigger than α. Then the uncentered maximal function of any other point in B_x is also bigger than α, and thus B_x is contained in E_α. This proves that E_α is open.

Let K be a compact subset of E_α. For each $x \in K$ there exists an open ball B_x containing the point x such that

$$\int_{B_x} |f(y)| \, dy > \alpha |B_x| . \tag{2.1.4}$$

Observe that $B_x \subset E_\alpha$ for all x. By compactness there exists a finite subcover $\{B_{x_1},\dots,B_{x_k}\}$ of K. Using Lemma 2.1.5 we find a subcollection of pairwise disjoint balls $B_{x_{j_1}},\dots,B_{x_{j_l}}$ such that (2.1.2) holds. Using (2.1.4) and (2.1.2) we obtain

$$|K| \le \Big|\bigcup_{i=1}^{k} B_{x_i}\Big| \le 3^n \sum_{i=1}^{l} |B_{x_{j_i}}| \le \frac{3^n}{\alpha}\sum_{i=1}^{l}\int_{B_{x_{j_i}}} |f(y)|\,dy \le \frac{3^n}{\alpha}\int_{E_\alpha} |f(y)|\,dy,$$

since all the balls $B_{x_{j_i}}$ are disjoint and contained in E_α. Taking the supremum over all compact $K \subseteq E_\alpha$ and using the inner regularity of Lebesgue measure, we deduce (2.1.3). We have now proved that M maps $L^1 \to L^{1,\infty}$ with constant 3^n. It is a trivial fact that M maps $L^\infty \to L^\infty$ with constant 1. Since M is well defined and finite a.e. on $L^1 + L^\infty$, it is also on $L^p(\mathbf{R}^n)$ for $1 < p < \infty$. The Marcinkiewicz interpolation theorem (Theorem 1.3.2) implies that M maps $L^p(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n)$ for all $1 < p < \infty$. Using Exercise 1.3.3, we obtain the following estimate for the operator norm of M on $L^p(\mathbf{R}^n)$:

$$\big\|M\big\|_{L^p\to L^p} \le \frac{p\,3^{\frac{n}{p}}}{p-1}. \tag{2.1.5}$$

Observe that a direct application of Theorem 1.3.2 would give the slightly worse bound of $2\big(\frac{p}{p-1}\big)^{\frac{1}{p}} 3^{\frac{n}{p}}$. Finally the boundedness of $\mathcal{M}$ follows from that of M. □

Remark 2.1.7. The previous proof gives a bound on the operator norm of M on $L^p(\mathbf{R}^n)$ that grows exponentially with the dimension. One may wonder whether this bound could be improved to a better one that does not grow exponentially in the dimension n, as $n\to\infty$. This is not possible; see Exercise 2.1.8.

Example 2.1.8. Let $R>0$. Then we have

$$\frac{R^n}{(|x|+R)^n} \le M(\chi_{B(0,R)})(x) \le \frac{6^n R^n}{(|x|+R)^n}. \tag{2.1.6}$$

The lower estimate in (2.1.6), is an easy consequence of the fact that the ball $B(x,|x|+R)$ contains the ball $B(0,R)$. For the upper estimate, we first consider the case where $|x|\le 2R$, when clearly $M(\chi_{B(0,R)})(x) \le 1 \le \frac{3^n R^n}{(|x|+R)^n}$. In the case where $|x|>2R$, if the balls $B(x,r)$ and $B(0,R)$ intersect, we must have that $r>|x|-R$. But note that $|x|-R > \frac{1}{3}(|x|+R)$, since $|x|>2R$. We conclude that for $|x|>2R$ we have

$$\mathcal{M}(\chi_{B(0,R)})(x) \le \sup_{r>0}\frac{|B(x,r)\cap B(0,R)|}{|B(x,r)|} \le \sup_{r>|x|-R}\frac{v_nR^n}{v_nr^n} \le \frac{R^n}{\big(\frac{1}{3}(|x|+R)\big)^n}$$

and thus the upper estimate in (2.1.6) holds since $M(\chi_{B(0,R)}) \le 2^n\mathcal{M}(\chi_{B(0,R)})$. Thus in both cases the upper estimate in (2.1.6) is valid.

Next we estimate $M(M(\chi_{B(0,R)}))(x)$. First we write

$$\frac{R^n}{(|x|+R)^n} \le \chi_{B(0,R)} + \sum_{k=0}^{\infty}\frac{R^n}{(R+2^kR)^n}\chi_{B(0,2^{k+1}R)\setminus B(0,2^kR)}.$$

Using the upper estimate in (2.1.6) and the sublinearity of M, we obtain

$$\begin{aligned} M\bigg(\frac{R^n}{(|\cdot|+R)^n}\bigg)(x) &\le M(\chi_{B(0,R)})(x)+\sum_{k=0}^{\infty}\frac{1}{(1+2^k)^n}M(\chi_{B(0,2^{k+1}R)})(x) \\ &\le \frac{6^n\,R^n}{(|x|+R)^n}+\sum_{k=0}^{\infty}\frac{1}{2^{nk}}\frac{6^n\,(2^{k+1}R)^n}{(|x|+2^{k+1}R)^n} \\ &\le \frac{C_n\log(e+|x|/R)}{(1+|x|/R)^n}\,, \end{aligned}$$

where the last estimate follows by summing separately over k satisfying $2^{k+1}\le|x|/R$ and $2^{k+1}\ge|x|/R$. Note that the presence of the logarithm does not affect the L^p boundedness of this function when $p>1$.

2.1.2 Control of Other Maximal Operators

We now study some properties of the Hardy–Littlewood maximal function. We begin with a notational definition that we plan to use throughout this book.

Definition 2.1.9. Given a function g on $\mathbf{R}^n$ and $\varepsilon>0$, we denote by g_ε the following function:

$$g_\varepsilon(x)=\varepsilon^{-n}g(\varepsilon^{-1}x)\,. \tag{2.1.7}$$

As observed in Example 1.2.17, if g is an integrable function with integral equal to 1, then the family defined by (2.1.7) is an approximate identity. Therefore, convolution with g_ε is an averaging operation. The Hardy–Littlewood maximal function $\mathcal{M}(f)$ 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 $\mathbf{R}^n$; here v_n is the volume of the unit ball $B(0,1)$. Indeed, we have

$$\begin{aligned} \mathcal{M}(f)(x) &= \sup_{\varepsilon>0}\frac{1}{v_n\varepsilon^n}\int_{\mathbf{R}^n}|f(x-y)|\,\chi_{B(0,1)}\Big(\frac{y}{\varepsilon}\Big)\,dy \\ &= \sup_{\varepsilon>0}(|f|*k_\varepsilon)(x)\,. \end{aligned}$$

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.

It turns out that the Hardy–Littlewood maximal function controls the averages of a function with respect to any radially decreasing L^1 function. Recall that a function f on $\mathbf{R}^n$ is called *radial* if $f(x)=f(y)$ whenever $|x|=|y|$. Note that a radial function f on $\mathbf{R}^n$ has the form $f(x)=\varphi(|x|)$ for some function φ on $\mathbf{R}^+$. We have the following result.

Theorem 2.1.10. *Let $k \geq 0$ be a function on $[0,\infty)$ that is continuous except at a finite number of points. Suppose that $K(x) = k(|x|)$ is an integrable function on* $\mathbf{R}^n$ *that satisfies*

$$K(x) \geq K(y), \quad \textit{whenever } |x| \leq |y|, \tag{2.1.8}$$

i.e., k is decreasing. Then the following estimate is true:

$$\sup_{\varepsilon>0}(|f| * K_\varepsilon)(x) \leq \big\|K\big\|_{L^1}\mathcal{M}(f)(x) \tag{2.1.9}$$

for all locally integrable functions f on $\mathbf{R}^n$.

Proof. We prove (2.1.9) when K is radial, satisfies (2.1.8), and is compactly supported and continuous. When this case is established, select a sequence K_j of radial, compactly supported, continuous functions that increase to K as $j \to \infty$. This is possible, since the function k is continuous except at a finite number of points. If (2.1.9) holds for each K_j, passing to the limit implies that (2.1.9) also holds for K. Next, we observe that it suffices to prove (2.1.9) for $x = 0$. When this case is established, replacing $f(t)$ by $f(t+x)$ implies that (2.1.9) holds for all x.

Let us now fix a radial, continuous, and compactly supported function K with support in the ball $B(0,R)$, satisfying (2.1.8). Also fix an $f \in L^1_{\text{loc}}$ and take $x = 0$. Let e_1 be the vector $(1,0,0,\dots,0)$ on the unit sphere $\mathbf{S}^{n-1}$. Polar coordinates give

$$\int_{\mathbf{R}^n} |f(y)|K_\varepsilon(-y)\,dy = \int_0^\infty \int_{\mathbf{S}^{n-1}} |f(r\theta)|K_\varepsilon(re_1)r^{n-1}\,d\theta\,dr. \tag{2.1.10}$$

Define functions

$$\begin{aligned} F(r) &= \int_{\mathbf{S}^{n-1}} |f(r\theta)|\,d\theta\,, \\ G(r) &= \int_0^r F(s)s^{n-1}\,ds\,, \end{aligned}$$

where $d\theta$ denotes surface measure on $\mathbf{S}^{n-1}$. Using these functions, (2.1.10), and integration by parts, we obtain

$$\begin{aligned} \int_{\mathbf{R}^n} |f(y)|K_\varepsilon(y)\,dy &= \int_0^{\varepsilon R} F(r)r^{n-1}K_\varepsilon(re_1)\,dr \\ &= G(\varepsilon R)K_\varepsilon(\varepsilon Re_1) - G(0)K_\varepsilon(0) - \int_0^{\varepsilon R} G(r)\,dK_\varepsilon(re_1) \\ &= \int_0^\infty G(r)\,d(-K_\varepsilon(re_1))\,, \end{aligned} \tag{2.1.11}$$

where two of the integrals are of Lebesgue–Stieltjes type and we used our assumptions that $G(0) = 0$, $K_\varepsilon(0) < \infty$, $G(\varepsilon R) < \infty$, and $K_\varepsilon(\varepsilon Re_1) = 0$. Let v_n be the volume of the unit ball in $\mathbf{R}^n$. Since

$$G(r) = \int_0^r F(s)s^{n-1}\,ds = \int_{|y|\leq r} |f(y)|\,dy \leq \mathcal{M}(f)(0)v_n r^n\,,$$

it follows that the expression in (2.1.11) is dominated by

$$\begin{aligned}\mathcal{M}(f)(0)v_n\int_0^\infty r^n d(-K_\varepsilon(re_1)) &= \mathcal{M}(f)(0)\int_0^\infty nv_n r^{n-1}K_\varepsilon(re_1)\,dr\\ &= \mathcal{M}(f)(0)\big\|K\big\|_{L^1}.\end{aligned}$$

Here we used integration by parts and the fact that the surface measure of the unit sphere $\mathbf{S}^{n-1}$ is equal to nv_n. See Appendix A.3. The theorem is now proved. □

Remark 2.1.11. Theorem 2.1.10 can be generalized as follows. If K is an L^1 function on $\mathbf{R}^n$ such that $|K(x)| \le k_0(|x|) = K_0(x)$, where k_0 is a nonnegative decreasing function on $[0,\infty)$ that is continuous except at a finite number of points, then (2.1.9) holds with $\|K\|_{L^1}$ replaced by $\|K_0\|_{L^1}$. Such a K_0 is called a *radial decreasing majorant* of K. This observation is formulated as the following corollary.

Corollary 2.1.12. *If a function φ has an integrable radially decreasing majorant Φ, then the estimate*

$$\sup_{t>0}|(f*\varphi_t)(x)| \le \big\|\Phi\big\|_{L^1}\mathcal{M}(f)(x)$$

is valid for all locally integrable functions f on $\mathbf{R}^n$.

Example 2.1.13. Let

$$P(x) = \frac{c_n}{(1+|x|^2)^{\frac{n+1}{2}}},$$

where c_n is a constant such that

$$\int_{\mathbf{R}^n} P(x)\,dx = 1.$$

The function P is called the *Poisson kernel*. We define L^1 dilates P_t of the Poisson kernel P by setting

$$P_t(x) = t^{-n}P(t^{-1}x)$$

for $t > 0$. It is straightforward to verify that when $n \ge 2$,

$$\frac{d^2}{dt^2}P_t + \sum_{j=1}^n \partial_j^2 P_t = 0,$$

that is, $P_t(x_1,\dots,x_n)$ is a *harmonic function* of the variables $(x_1,\dots,x_n,t)$. Therefore, for $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$, the function

$$u(x,t) = (f*P_t)(x)$$

is harmonic in $\mathbf{R}^{n+1}_+$ and converges to $f(x)$ in $L^p(dx)$ as $t \to 0$, since $\{P_t\}_{t>0}$ is an approximate identity. If we knew that $f*P_t$ converged to f a.e. as $t \to 0$, then we could say that $u(x,t)$ solves the *Dirichlet problem*

$$\begin{aligned} \partial_t^2 u + \sum_{j=1}^{n} \partial_j^2 u &= 0 \qquad\qquad \text{on } \mathbf{R}_+^{n+1}, \\ u(x,0) &= f(x) \qquad \text{a.e. on } \mathbf{R}^n. \end{aligned} \tag{2.1.12}$$

Solving the Dirichlet problem (2.1.12) motivates the study of the almost everywhere convergence of the expressions $f * P_t$.

Let us now compute the value of the constant c_n. Denote by ω_{n-1} the surface area of $\mathbf{S}^{n-1}$. Using polar coordinates, we obtain

$$\begin{aligned} \frac{1}{c_n} &= \int_{\mathbf{R}^n} \frac{dx}{(1+|x|^2)^{\frac{n+1}{2}}} \\ &= \omega_{n-1} \int_0^\infty \frac{r^{n-1}}{(1+r^2)^{\frac{n+1}{2}}} \, dr \\ &= \omega_{n-1} \int_0^{\pi/2} (\sin\varphi)^{n-1} \, d\varphi \qquad\qquad (r = \tan\varphi) \\ &= \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \frac{1}{2} \frac{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n+1}{2})} \\ &= \frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})}, \end{aligned}$$

where we used the formula for ω_{n-1} in Appendix A.3 and an identity in Appendix A.4. We conclude that

$$c_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}$$

and that the Poisson kernel on $\mathbf{R}^n$ is given by

$$P(x) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \frac{1}{(1+|x|^2)^{\frac{n+1}{2}}}. \tag{2.1.13}$$

Theorem 2.1.10 implies that the solution of the Dirichlet problem (2.1.12) is pointwise bounded by the Hardy–Littlewood maximal function of f.

2.1.3 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.

Here is our setup. Let (X,μ), (Y,ν) be measure spaces and let $0<p\le\infty$, $0<q<\infty$. 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 $\|f-g\|_{L^p}<\delta$. 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

$$T_*(f)(y)=\sup_{\varepsilon>0}|T_\varepsilon(f)(y)| \tag{2.1.14}$$

and assume that $T_*(f)$ is ν- measurable for any $f\in L^p(X,\mu)$. We have the following.

Theorem 2.1.14. *Let $0<p<\infty$, $0<q<\infty$, and T_ε and T_* as previously. Suppose that for some $B>0$ and all $f\in L^p(X)$ we have*

$$\big\|T_*(f)\big\|_{L^{q,\infty}}\le B\big\|f\big\|_{L^p} \tag{2.1.15}$$

and that for all $f\in D$,

$$\lim_{\varepsilon\to0}T_\varepsilon(f)=T(f) \tag{2.1.16}$$

exists and is finite ν-a.e. (and defines a linear operator on D). Then for all functions f in $L^p(X,\mu)$ the limit (2.1.16) exists and is finite ν-a.e., and defines a linear operator T on $L^p(X)$ (uniquely extending T defined on D) that satisfies

$$\big\|T(f)\big\|_{L^{q,\infty}}\le B\big\|f\big\|_{L^p} \tag{2.1.17}$$

for all functions f in $L^p(X)$.

Proof. Given f in L^p, we define the *oscillation* of f:

$$O_f(y)=\limsup_{\varepsilon\to0}\limsup_{\theta\to0}|T_\varepsilon(f)(y)-T_\theta(f)(y)|\,.$$

We would like to show that for all $f\in L^p$ and $\delta>0$,

$$\nu(\{y\in Y:\ O_f(y)>\delta\})=0\,. \tag{2.1.18}$$

Once (2.1.18) is established, given $f\in L^p(X)$, 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 $T(f)(y)$ as $\varepsilon\to0$. The operator T defined this way on $L^p(X)$ 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 $\|f-g\|_{L^p}<\eta$. Since $T_\varepsilon(g)\to T(g)$ 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

$$O_f(y)\le O_g(y)+O_{f-g}(y)=O_{f-g}(y)\qquad \nu\text{-a.e.}$$

Now for any $\delta > 0$ we have

$$\begin{aligned}\nu(\{y \in Y : O_f(y) > \delta\}) &\le \nu(\{y \in Y : O_{f-g}(y) > \delta\}) \\ &\le \nu(\{y \in Y : 2T_*(f-g)(y) > \delta\}) \\ &\le \left(2B\|f-g\|_{L^p}/\delta\right)^q \\ &\le (2B\eta/\delta)^q .\end{aligned}$$

Letting $\eta \to 0$, we deduce (2.1.18). We conclude that $T_\varepsilon(f)$ is a Cauchy sequence, and hence it converges ν-a.e. to some $T(f)$. Since $|T(f)| \le |T_*(f)|$, the conclusion (2.1.17) of the theorem follows easily. $\square$

We now derive some applications. First we return to the issue of almost everywhere convergence of the expressions $f * P_y$, where P is the Poisson kernel.

Example 2.1.15. Fix $1 \le p < \infty$ and $f \in L^p(\mathbf{R}^n)$. Let

$$P(x) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \frac{1}{(1+|x|^2)^{\frac{n+1}{2}}}$$

be the Poisson kernel on $\mathbf{R}^n$ and let $P_\varepsilon(x) = \varepsilon^{-n}P(\varepsilon^{-1}x)$. We deduce from the previous theorem that the family $f * P_\varepsilon$ converges to f a.e. Let D be the set of all continuous functions with compact support on $\mathbf{R}^n$. Since the family $(P_\varepsilon)_{\varepsilon>0}$ is an approximate identity, Theorem 1.2.19 (2) implies that for f in D we have that $f * P_\varepsilon \to f$ uniformly on compact subsets of $\mathbf{R}^n$ and hence pointwise everywhere. In view of Theorem 2.1.10, the supremum of the family of linear operators $T_\varepsilon(f) = f * P_\varepsilon$ is controlled by the Hardy–Littlewood maximal function, and thus it maps L^p to $L^{p,\infty}$ for $1 \le p < \infty$. Theorem 2.1.14 now gives that $f * P_\varepsilon$ converges to f a.e. for all $f \in L^p$.

Here is another application of Theorem 2.1.14. Exercise 2.1.10 contains other applications.

Corollary 2.1.16. ***(Lebesgue's differentiation theorem)*** *For any locally integrable function f on $\mathbf{R}^n$ we have*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy = f(x) \tag{2.1.19}$$

for almost all x in $\mathbf{R}^n$. Consequently we have $|f| \le \mathcal{M}(f)$ a.e. There is also an analogous statement to (2.1.19) in which balls are replaced by cubes centered at x. Precisely, for any locally integrable function f on $\mathbf{R}^n$ we have

$$\lim_{r\to 0} \frac{1}{(2r)^n} \int_{x+[-r,r]^n} f(y)\, dy = f(x) \tag{2.1.20}$$

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

Proof. Since $\mathbf{R}^n$ is the union of the balls $B(0,N)$ for $N = 1,2,3\ldots$, it suffices to prove the required conclusion for almost all x inside a fixed ball $B(0,N)$. Given a locally integrable function f on $\mathbf{R}^n$, consider the function $f_N = f\chi_{B(0,N+1)}$. Then f_N lies in $L^1(\mathbf{R}^n)$. Let T_ε be the operator given with convolution with k_ε, where $k = v_n^{-1}\chi_{B(0,1)}$ and $0 < \varepsilon < 1$. We know that the corresponding maximal operator T_* is controlled by the centered Hardy–Littlewood maximal function $\mathcal{M}$, which maps L^1 to $L^{1,\infty}$. It is straightforward to verify that (2.1.19) holds for all continuous functions f with compact support. Since this set of functions is dense in L^1, and T_* maps L^1 to $L^{1,\infty}$, Theorem 2.1.14 implies that (2.1.19) holds for all integrable functions on $\mathbf{R}^n$, in particular for f_N. But for $0 < \varepsilon < 1$ and $x \in B(0,N)$ we have $f\chi_{B(x,\varepsilon)} = f_N\chi_{B(x,\varepsilon)}$, so it follows that

$$\lim_{\varepsilon\to 0}\frac{1}{|B(x,\varepsilon)|}\int_{B(x,\varepsilon)} f(y)\,dy = \lim_{\varepsilon\to 0}\frac{1}{|B(x,\varepsilon)|}\int_{B(x,\varepsilon)} f_N(y)\,dy = f_N(x)$$

for almost all $x \in \mathbf{R}^n$, in particular for almost all x in $B(0,N)$. But on this set $f_N = f$, so the required conclusion follows. The assertion that $|f| \le \mathcal{M}(f)$ a.e. is an easy consequence of (2.1.19) when the limit is replaced by a supremum.

Finally, with minor modifications, the proof can be adjusted to work for cubes in place of balls. To prove (2.1.20), for $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ we introduce the maximal operator

$$\mathcal{M}_c(f)(x) = \sup_{r>0}\frac{1}{(2r)^n}\int_{x+[-r,r]^n}|f(y)|\,dy\,.$$

Then Exercise 2.1.3 yields that $\mathcal{M}_c$ maps $L^1(\mathbf{R}^n)$ to weak $L^1(\mathbf{R}^n)$ and the preceding proof with $\mathcal{M}_c$ in place of $\mathcal{M}$ yields (2.1.20). □

The following corollaries were inspired by Example 2.1.15.

Corollary 2.1.17. ***(Differentiation theorem for approximate identities)*** *Let K be an L^1 function on $\mathbf{R}^n$ with integral 1 that has a continuous integrable radially decreasing majorant. Then $f*K_\varepsilon \to f$ a.e. as $\varepsilon \to 0$ for all $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$.*

Proof. It follows from Example 1.2.17 that K_ε is an approximate identity. Theorem 1.2.19 now implies that $f*K_\varepsilon \to f$ uniformly on compact sets when f is continuous. Let D be the space of all continuous functions with compact support. Then $f*K_\varepsilon \to f$ a.e. for $f \in D$. It follows from Corollary 2.1.12 that $T_*(f) = \sup_{\varepsilon>0}|f*K_\varepsilon|$ maps L^p to $L^{p,\infty}$ for $1 \le p < \infty$. Using Theorem 2.1.14, we conclude the proof of the corollary.

□

Remark 2.1.18. Fix $f \in L^p(\mathbf{R}^n)$ for some $1 \le p < \infty$. Theorem 1.2.19 implies that $f*K_\varepsilon$ converges to f in L^p and hence some subsequence $f*K_{\varepsilon_n}$ of $f*K_\varepsilon$ converges to f a.e. as $n \to \infty$, ($\varepsilon_n \to 0$). Compare this result with Corollary 2.1.17, which gives a.e. convergence for the whole family $f*K_\varepsilon$ as $\varepsilon \to 0$.

Corollary 2.1.19. ***(Differentiation theorem for multiples of approximate identities)*** *Let K be a function on $\mathbf{R}^n$ that has an integrable radially decreasing majorant.*

Let $a = \int_{\mathbf{R}^n} K(x)\,dx$. Then for all $f \in L^p(\mathbf{R}^n)$ and $1 \le p < \infty$, $(f * K_\varepsilon)(x) \to af(x)$ for almost all $x \in \mathbf{R}^n$ as $\varepsilon \to 0$.

Proof. Use Theorem 1.2.21 instead of Theorem 1.2.19 in the proof of Corollary 2.1.17. □

The following application of the Lebesgue differentiation theorem uses a simple *stopping-time argument*. This is the sort of argument in which a selection procedure stops when it is exhausted at a certain scale and is then repeated at the next scale. A certain refinement of the following proposition is of fundamental importance in the study of singular integrals given in Chapter 4.

Proposition 2.1.20. *Given a nonnegative integrable function f on $\mathbf{R}^n$ and $\alpha > 0$, there exists a collection of disjoint (possibly empty) open cubes Q_j such that for almost all $x \in \left(\bigcup_j Q_j\right)^c$ we have $f(x) \le \alpha$ and*

$$\alpha < \frac{1}{|Q_j|}\int_{Q_j} f(t)\,dt \le 2^n\alpha\,. \tag{2.1.21}$$

Proof. The proof provides an excellent paradigm of a stopping-time argument. Start by decomposing $\mathbf{R}^n$ as a union of cubes of equal size, whose interiors are disjoint, and whose diameter is so large that $|Q|^{-1}\int_Q f(x)\,dx \le \alpha$ for every Q in this mesh. This is possible since f is integrable and $|Q|^{-1}\int_Q f(x)\,dx \to 0$ as $|Q| \to \infty$. Call the union of these cubes $\mathscr{E}_0$.

Divide each cube in the mesh into 2^n congruent cubes by bisecting each of the sides. Call the new collection of cubes $\mathscr{E}_1$. Select a cube Q in $\mathscr{E}_1$ if

$$\frac{1}{|Q|}\int_Q f(x)\,dx > \alpha \tag{2.1.22}$$

and call the set of all selected cubes $\mathscr{S}_1$. Now subdivide each cube in $\mathscr{E}_1 \setminus \mathscr{S}_1$ into 2^n congruent cubes by bisecting each of the sides as before. Call this new collection of cubes $\mathscr{E}_2$. Repeat the same procedure and select a family of cubes $\mathscr{S}_2$ that satisfy (2.1.22). Continue this way ad infinitum and call the cubes in $\bigcup_{m=1}^\infty \mathscr{S}_m$ "selected." If Q was selected, then there exists Q_1 in $\mathscr{E}_{m-1}$ containing Q that was not selected at the $(m-1)th$ step for some $m \ge 1$. Therefore,

$$\alpha < \frac{1}{|Q|}\int_Q f(x)\,dx \le 2^n\frac{1}{|Q_1|}\int_{Q_1} f(x)\,dx \le 2^n\alpha\,.$$

Now call F the closure of the complement of the union of all selected cubes. If $x \in F$, then there exists a sequence of cubes containing x whose diameter shrinks down to zero such that the average of f over these cubes is less than or equal to α. By Corollary 2.1.16, it follows that $f(x) \le \alpha$ almost everywhere in F. This proves the proposition. □

In the proof of Proposition 2.1.20 it was not crucial to assume that f was defined on all $\mathbf{R}^n$, but only on a cube. We now give a local version of this result.

Corollary 2.1.21. *Let $f \geq 0$ be an integrable function over a cube Q in $\mathbf{R}^n$ and let $\alpha \geq \frac{1}{|Q|}\int_Q f\,dx$. Then there exist disjoint (possibly empty) open subcubes Q_j of Q such that for almost all $x \in Q \setminus \bigcup_j Q_j$ we have $f \leq \alpha$ and (2.1.21) holds for all Q_j.*

Proof. The proof easily follows by a simple modification of Proposition 2.1.20 in which $\mathbf{R}^n$ is replaced by the fixed cube Q. To apply Corollary 2.1.16, we extend f to be zero outside the cube Q. □

See Exercise 2.1.4 for an application of Proposition 2.1.20 involving maximal functions.

Exercises

2.1.1. A positive Borel measure μ on $\mathbf{R}^n$ is called *inner regular* if for any open subset U of $\mathbf{R}^n$ we have $\mu(U) = \sup\{\mu(K) :\ K \subsetneqq U,\ K \text{ compact}\}$ and μ is called *locally finite* if $\mu(B) < \infty$ for all balls B.
(a) Let μ be a positive inner regular locally finite measure on $\mathbf{R}^n$ that satisfies the following *doubling condition*: There exists a constant $D(\mu) > 0$ such that for all $x \in \mathbf{R}^n$ and $r > 0$ we have

$$\mu(3B(x,r)) \leq D(\mu)\,\mu(B(x,r)).$$

For $f \in L^1_{\text{loc}}(\mathbf{R}^n,\mu)$ define the uncentered maximal function $M_\mu(f)$ with respect to μ by

$$M_\mu(f)(x) \;=\; \sup_{r>0}\ \sup_{\substack{z:\,|z-x|<r\\ \mu(B(z,r))\neq 0}} \frac{1}{\mu(B(z,r))}\int_{B(z,r)} f(y)\,d\mu(y)\,.$$

Show that M_μ maps $L^1(\mathbf{R}^n,\mu)$ to $L^{1,\infty}(\mathbf{R}^n,\mu)$ with constant at most $D(\mu)$ and $L^p(\mathbf{R}^n,\mu)$ to itself with constant at most $2\big(\frac{p}{p-1}\big)^{\frac{1}{p}} D(\mu)^{\frac{1}{p}}$.

(b) Obtain as a consequence a differentiation theorem analogous to Corollary 2.1.16.
$\big[$*Hint:* Part (a): For $f \in L^1(\mathbf{R}^n,\mu)$ show that the set $E_\alpha = \{M_\mu(f) > \alpha\}$ is open. Then use the argument of the proof of Theorem 2.1.6 and the inner regularity of μ.$\big]$

2.1.2. On $\mathbf{R}$ consider the maximal function M_μ of Exercise 2.1.1.
(a) (*W. H. Young*) Prove the following covering lemma. Given a finite set $\mathscr{F}$ of open intervals in $\mathbf{R}$, prove that there exist two subfamilies each consisting of pairwise disjoint intervals such that the union of the intervals in the original family is equal to the union of the intervals of both subfamilies. Use this result to show that the maximal function M_μ of Exercise 2.1.1 maps $L^1(\mu) \to L^{1,\infty}(\mu)$ with constant at most 2.

(b) ([134]) Prove that for any σ-finite positive measure μ on $\mathbf{R}$, $\alpha > 0$, and $f \in L^1_{\mathrm{loc}}(\mathbf{R},\mu)$ we have

$$\frac{1}{\alpha}\int_A |f|\,d\mu - \mu(A) \le \frac{1}{\alpha}\int_{\{|f|>\alpha\}} |f|\,d\mu - \mu(\{|f|>\alpha\}).$$

Use this result and part (a) to prove that for all $\alpha > 0$ and all locally integrable f we have

$$\mu(\{|f|>\alpha\}) + \mu(\{M_\mu(f)>\alpha\}) \le \frac{1}{\alpha}\int_{\{|f|>\alpha\}} |f|\,d\mu + \frac{1}{\alpha}\int_{\{M_\mu(f)>\alpha\}} |f|\,d\mu$$

and note that equality is obtained when $\alpha = 1$ and $f(x) = |x|^{-1/p}$.
(c) Conclude that M_μ maps $L^p(\mu)$ to $L^p(\mu)$, $1 < p < \infty$, with bound at most the unique positive solution A_p of the equation

$$(p-1)x^p - px^{p-1} - 1 = 0.$$

(d) ([136]) If μ is the Lebesgue measure show that for $1 < p < \infty$ we have

$$\big\|M\big\|_{L^p\to L^p} = A_p,$$

where A_p is the unique positive solution of the equation in part (c).
$\big[$*Hint:* Part (a): Select a subset $\mathscr{G}$ of $\mathscr{F}$ with minimal cardinality such that $\bigcup_{J\in\mathscr{G}} J = \bigcup_{I\in\mathscr{F}} I$. Part (d): One direction follows from part (c). Conversely, $M(|x|^{-1/p})(1) = \frac{p}{p-1}\frac{\gamma^{1/p'}+1}{\gamma+1}$, where γ is the unique positive solution of the equation $\frac{p}{p-1}\frac{\gamma^{1/p'}+1}{\gamma+1} = \gamma^{-1/p}$. Conclude that $M(|x|^{-1/p})(1) = A_p$ and that $M(|x|^{-1/p}) = A_p|x|^{-1/p}$. Since this function is not in L^p, consider the family $f_\varepsilon(x) = |x|^{-1/p}\min(|x|^{-\varepsilon},|x|^{\varepsilon})$, $\varepsilon > 0$, and show that $M(f_\varepsilon)(x) \ge (1+\gamma^{\frac{1}{p'}+\varepsilon})(1+\gamma)^{-1}(\frac{1}{p'}+\varepsilon)^{-1} f_\varepsilon(x)$ for $0 < \varepsilon < p'$.$\big]$

2.1.3. Define the centered Hardy–Littlewood maximal function $\mathcal{M}_c$ and the uncentered Hardy–Littlewood maximal function M_c using cubes with sides parallel to the axes instead of balls in $\mathbf{R}^n$. Prove that

$$1 \le \frac{M(f)}{\mathcal{M}(f)} \le 2^n, \qquad \frac{1}{n^{\frac{n}{2}}}\frac{2^n}{v_n} \le \frac{M(f)}{M_c(f)} \le \frac{2^n}{v_n}, \qquad \frac{1}{n^{\frac{n}{2}}}\frac{2^n}{v_n} \le \frac{\mathcal{M}(f)}{\mathcal{M}_c(f)} \le \frac{2^n}{v_n},$$

where v_n is the volume of the unit ball in $\mathbf{R}^n$. Conclude that $\mathcal{M}_c$ and M_c are weak type $(1,1)$ and they map $L^p(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n)$ for $1 < p \le \infty$.

2.1.4. (a) Prove the estimate:

$$|\{x \in \mathbf{R}^n : M(f)(x) > 2\alpha\}| \le \frac{3^n}{\alpha}\int_{\{|f|>\alpha\}} |f(y)|\,dy$$

and conclude that M maps L^p to $L^{p,\infty}$ with norm at most $2\cdot 3^{n/p}$ for $1 \le p < \infty$.

(b) Deduce that if $f\log^+(2|f|)$ is integrable over a ball B, then $M(f)$ is integrable over the same ball B.
(c) ([375], [336]) Apply Proposition 2.1.20 to $|f|$ and $\alpha>0$ and Exercise 2.1.3 to show that with $c_n=2^n(n^{n/2}v_n)^{-1}$ we have

$$|\{x\in\mathbf{R}^n:\ M(f)(x)>c_n\,\alpha\}|\geq\frac{2^{-n}}{\alpha}\int_{\{|f|>\alpha\}}|f(y)|\,dy\,.$$

(d) Suppose that f is integrable and supported in a ball $B(0,\rho)$. Show that for x in $B(0,2\rho)\setminus B(0,\rho)$ we have $\mathcal{M}(f)(x)\leq\mathcal{M}(f)(\rho^2|x|^{-2}x)$. Conclude that

$$\int_{B(0,2\rho)}\mathcal{M}(f)\,dx\leq(4^n+1)\int_{B(0,\rho)}\mathcal{M}(f)\,dx$$

and from this deduce a similar inequality for $M(f)$.
(e) Suppose that f is integrable and supported in a ball B and that $M(f)$ is integrable over B. Let $\lambda_0=2^n|B|^{-1}\|f\|_{L^1}$. Use part (b) to prove that $f\log^+(\lambda_0^{-1}c_n|f|)$ is integrable over B.
$\big[$*Hint:* Part (a): Write $f=f\chi_{|f|>\alpha}+f\chi_{|f|\leq\alpha}$. Part (b): Show that $M(f\chi_E)$ is integrable over B, where $E=\{|f|\geq 1/2\}$. Part (c): Use Proposition 2.1.20. Part (d): Let $x'=\rho^2|x|^{-2}x$ for some $\rho<|x|<2\rho$. Show that for $R>|x|-\rho$, we have that

$$\int_{B(x,R)}|f(z)|\,dz\leq\int_{B(x',R)}|f(z)|\,dz$$

by showing that $B(x,R)\cap B(0,\rho)\subset B(x',R)$. Part (e): For $x\notin 2B$ we have $M(f)(x)\leq\lambda_0$, hence $\int_{2B}M(f)(x)\,dx\geq\int_{\lambda_0}^\infty|\{x\in 2B:\ M(f)(x)>\alpha\}|\,d\alpha.\big]$

2.1.5. (*A. Kolmogorov*) Let S be a sublinear operator that maps $L^1(\mathbf{R}^n)$ to $L^{1,\infty}(\mathbf{R}^n)$ with norm B. Suppose that $f\in L^1(\mathbf{R}^n)$. Prove that for any set A of finite Lebesgue measure and for all $0<q<1$ we have

$$\int_A|S(f)(x)|^q\,dx\leq(1-q)^{-1}B^q|A|^{1-q}\|f\|_{L^1}^q,$$

and in particular, for the Hardy–Littlewood maximal operator,

$$\int_A M(f)(x)^q\,dx\leq(1-q)^{-1}3^{nq}|A|^{1-q}\|f\|_{L^1}^q\,.$$

$\big[$*Hint:* Use the identity

$$\int_A|S(f)(x)|^q\,dx=\int_0^\infty q\,\alpha^{q-1}|\{x\in A:\,S(f)(x)>\alpha\}|\,d\alpha$$

and estimate the last measure by $\min(|A|,\frac{B}{\alpha}\|f\|_{L^1}).\big]$

2.1.6. Let $M_s(f)(x)$ be the supremum of the averages of $|f|$ over all rectangles with sides parallel to the axes containing x. The operator M_s is called the *strong maximal function*.
(a) Prove that M_s maps $L^p(\mathbf{R}^n)$ to itself.
(b) Show that the operator norm of M_s is A_p^n, where A_p is as in Exercise 2.1.2 (c).
(c) Prove that M_s is not weak type (1,1).

2.1.7. Prove that if

$$|\varphi(x_1,\dots,x_n)| \le A(1+|x_1|)^{-1-\varepsilon}\cdots(1+|x_n|)^{-1-\varepsilon}$$

for some $A,\varepsilon>0$, and $\varphi_{t_1,\dots,t_n}(x)=t_1^{-1}\cdots t_n^{-1}\varphi(t_1^{-1}x_1,\dots,t_n^{-1}x_n)$, then the maximal operator

$$f\mapsto \sup_{t_1,\dots,t_n>0}|f*\varphi_{t_1,\dots,t_n}|$$

is pointwise controlled by the strong maximal function.

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

2.1.9. (a) In $\mathbf{R}^2$ let $M_0(f)(x)$ be the maximal function obtained by taking the supremum of the averages of $|f|$ over all rectangles (of arbitrary orientation) containing x. Prove that M_0 is not bounded on $L^p(\mathbf{R}^n)$ for $p\le 2$ and conclude that M_0 is not weak type $(1,1)$.
(b) Let $M_{00}(f)(x)$ be the maximal function obtained by taking the supremum of the averages of $|f|$ over all rectangles in $\mathbf{R}^2$ of arbitrary orientation but fixed eccentricity containing x. (The eccentricity of a rectangle is the ratio of its longer side to its shorter side.) Using a covering lemma, show that M_{00} is weak type $(1,1)$ with a bound proportional to the square of the eccentricity.
(c) On $\mathbf{R}^n$ define a maximal function by taking the supremum of the averages of $|f|$ over all products of intervals $I_1\times\cdots\times I_n$ containing a point x with $|I_2|=a_2|I_1|,\dots,|I_n|=a_n|I_1|$ and $a_2,\dots,a_n>0$ fixed. Show that this maximal function is of weak type $(1,1)$ with bound independent of the numbers $a_2,\dots,a_n$.
$\big[$*Hint:* Part (b): Let b be the eccentricity. If two rectangles with the same eccentricity intersect, then the smaller one is contained in the bigger one scaled $4b$ times. Then use an argument similar to that in Lemma 2.1.5.$\big]$

2.1.10. (a) Let $0<p,q<\infty$ and let X,Y be measure spaces. Suppose that T_ε are maps from $L^p(X)$ to $L^{q,\infty}(Y)$ satisfy $|T_\varepsilon(f+g)|\le K(|T_\varepsilon(f)|+|T_\varepsilon(g)|)$ for all $\varepsilon>0$ and all $f,g\in L^p(X)$, and also $\lim_{\varepsilon\to0}T_\varepsilon(f)=0$ a.e. for all f in some dense subspace D of $L^p(X)$. Assume furthermore that the maximal operator $T_*(f)=\sup_{\varepsilon>0}|T_\varepsilon(f)|$ maps $L^p(X)$ to $L^{q,\infty}(Y)$. Prove that $\lim_{\varepsilon\to0}T_\varepsilon(f)=0$ a.e. for all f in $L^p(X)$.
(b) Use the result in part (a) to prove the following version of the Lebesgue differentiation theorem: Let $f\in L^p(\mathbf{R}^n)$ for some $0<p<\infty$. Then for almost all $x\in\mathbf{R}^n$ we have

$$\lim_{\substack{|B|\to 0\\ B\ni x}} \frac{1}{|B|}\int_B |g(y)-g(x)|^p\,dy = 0\,,$$

where the limit is taken over all open balls B containing x and shrinking to $\{x\}$.
(c) Conclude that for any f in $L^1_{\mathrm{loc}}(\mathbf{R}^n)$ and for almost all $x\in\mathbf{R}^n$ we have

$$\lim_{\substack{|B|\to 0\\ B\ni x}} \frac{1}{|B|}\int_B f(y)\,dy = f(x)\,,$$

where the limit is taken over all open balls B containing x and shrinking to $\{x\}$.
$\big[$*Hint:* (a) Define an oscillation $O_f(y) = \limsup_{\varepsilon\to 0}|T_\varepsilon(f)(y)|$. For all f in $L^p(X)$ and $g\in D$ we have that $O_f(y)\le KO_{f-g}(y)$. Then use the argument in the proof of Theorem 2.1.14. (b) Apply part (a) with

$$T_\varepsilon(f)(x) = \sup_{B(z,\varepsilon)\ni x}\left(\frac{1}{|B(z,\varepsilon)|}\int_{B(z,\varepsilon)}|f(y)-f(x)|^p\,dy\right)^{1/p},$$

observing that $T_*(f) = \sup_{\varepsilon>0}T_\varepsilon(f)\le \max(1,2^{\frac{1-p}{p}})\big(|f|+M(|f|^p)^{\frac{1}{p}}\big)$. (c) Follows from part (b) with $p=1$. Note that part (b) can be proved without part (a) but using part (c) as follows: for every rational number a there is a set E_a of Lebesgue measure zero such that for $x\in\mathbf{R}^n\setminus E_a$ we have $\lim_{B\ni x,|B|\to 0}\frac{1}{|B|}\int_B|g(y)-a|^p\,dy = |g(x)-a|^p$, since the function $y\mapsto|f(y)-a|^p$ is in $L^1_{\mathrm{loc}}(\mathbf{R}^n)$. By considering an enumeration of the rationals, find a set of measure zero E such for $x\notin E$ the preceding limit exists for all rationals a and by continuity for all real numbers a, in particular for $a=g(x)$.$\big]$

2.1.11. Let f be in $L^1(\mathbf{R})$. Define the right maximal function $M_R(f)$ and the left maximal function $M_L(f)$ as follows:

$$\begin{aligned} M_L(f)(x) &= \sup_{r>0}\frac{1}{r}\int_{x-r}^{x}|f(t)|\,dt\,,\\ M_R(f)(x) &= \sup_{r>0}\frac{1}{r}\int_{x}^{x+r}|f(t)|\,dt\,. \end{aligned}$$

(a) Show that for all $\alpha>0$ and $f\in L^1(\mathbf{R})$ we have

$$\begin{aligned} |\{x\in\mathbf{R}: M_L(f)(x)>\alpha\}| &= \frac{1}{\alpha}\int_{\{M_L(f)>\alpha\}}|f(t)|\,dt\,,\\ |\{x\in\mathbf{R}:\ M_R(f)(x)>\alpha\}| &= \frac{1}{\alpha}\int_{\{M_R(f)>\alpha\}}|f(t)|\,dt\,. \end{aligned}$$

(b) Extend the definition of $M_L(f)$ and $M_R(f)$ for $f\in L^p(\mathbf{R})$ for $1\le p\le\infty$. Show that M_L and M_R map L^p to L^p with norm at most $p/(p-1)$ for all p with $1<p<\infty$.
(c) Construct examples to show that the operator norms of M_L and M_R on $L^p(\mathbf{R})$ are exactly $p/(p-1)$ for $1<p<\infty$.

(d) Prove that $M = \max(M_R, M_L)$.
(e) Let $N = \min(M_R, M_L)$. Obtain the following consequence of part (a),

$$\int_{\mathbf{R}} M(f)^p + N(f)^p\,dx = \frac{p}{p-1}\int_{\mathbf{R}} |f|\left(M(f)^{p-1} + N(f)^{p-1}\right)dx,$$

(f) Use part (e) to prove that

$$(p-1)\big\|M(f)\big\|_{L^p}^p - p\big\|f\big\|_{L^p}\big\|M(f)\big\|_{L^p}^{p-1} - \big\|f\big\|_{L^p}^p \le 0.$$

$\big[$*Hint:* (a) Write the set $E_\alpha = \{M_R(f) > \alpha\}$ as a union of open intervals (a_j, b_j). For each x in (a_j, b_j), let $N_x = \big\{s \in \mathbf{R}:\ \int_x^s |f| > \alpha(s-x)\big\} \cap (x, b_j]$. Show that N_x is nonempty and that $\sup N_x = b_j$ for every $x \in (a_j, b_j)$. Conclude that $\int_{a_j}^{b_j} |f(t)|\,dt \ge \alpha(b_j - a_j)$, which implies that each a_j is finite. For the reverse inequality use that $a_j \notin E_\alpha$. Part (d) is due to K. L. Phillips. (e) First obtain a version of the equality with M_R in the place of M and M_L in the place of N. Then use that $M(f)^q + N(f)^q = M_L(f)^q + M_R(f)^q$ for all q. (f) Use that $|f|N(f)^{p-1} \le \frac{1}{p}|f|^p + \frac{1}{p'}N(f)^p$. This alternative proof of the result in Exercise 2.1.2(c) was suggested by J. Duoandikoetxea.$\big]$

2.1.12. A cube $Q = [a_1 2^k, (a_1+1)2^k) \times \cdots \times [a_n 2^k, (a_n+1)2^k)$ on $\mathbf{R}^n$ is called *dyadic* if $k, a_1, \dots, a_n \in \mathbf{Z}$. Observe that either two dyadic cubes are disjoint or one contains the other. Define the *dyadic maximal function*

$$M_d(f)(x) = \sup_{Q \ni x} \frac{1}{|Q|}\int_Q f(y)\,dy,$$

where the supremum is taken over all dyadic cubes Q containing x.
(a) Prove that M_d maps L^1 to $L^{1,\infty}$ with constant at most one. Presicely, show that for all $\alpha > 0$ and $f \in L^1(\mathbf{R}^n)$ we have

$$|\{x \in \mathbf{R}^n:\ M_d(f)(x) > \alpha\}| \le \frac{1}{\alpha}\int_{\{M_d(f) > \alpha\}} f(t)\,dt.$$

(b) Conclude that M_d maps $L^p(\mathbf{R}^n)$ to itself with constant at most $p/(p-1)$.

2.1.13. Observe that the proof of Theorem 2.1.6 yields the estimate

$$\lambda |\{M(f) > \lambda\}|^{\frac{1}{p}} \le 3^n |\{M(f) > \lambda\}|^{-1+\frac{1}{p}} \int_{\{M(f) > \lambda\}} |f(y)|\,dy$$

for $\lambda > 0$ and f locally integrable. Use the result of Exercise 1.1.12(a) to prove that the Hardy–Littlewood maximal operator M maps the space $L^{p,\infty}(\mathbf{R}^n)$ to itself for $1 < p < \infty$.

2.1.14. Let $K(x) = (1+|x|)^{-n-\delta}$ be defined on $\mathbf{R}^n$. Prove that there exists a constant $C_{n,\delta}$ such that for all $\varepsilon_0 > 0$ we have the estimate

$$\sup_{\varepsilon>\varepsilon_0} (|f| * K_\varepsilon)(x) \le C_{n,\delta} \sup_{\varepsilon>\varepsilon_0} \frac{1}{\varepsilon^n} \int_{|y-x|\le\varepsilon} |f(y)|\, dy,$$

for all f locally integrable on $\mathbf{R}^n$.
$\big[$*Hint:* Apply only a minor modification to the proof of Theorem 2.1.10.$\big]$

2.2 The Schwartz Class and the Fourier Transform

In this section we introduce the single most important tool in harmonic analysis, the Fourier transform. It is often the case that the Fourier transform is introduced as an operation on L^1 functions. In this exposition we first define the Fourier transform on a smaller class, the space of Schwartz functions, which turns out to be a very natural environment. Once the basic properties of the Fourier transform are derived, we extend its definition to other spaces of functions.

We begin with some preliminaries. Given $x = (x_1,\dots,x_n) \in \mathbf{R}^n$, we set $|x| = (x_1^2+\cdots+x_n^2)^{1/2}$. The partial derivative of a function f on $\mathbf{R}^n$ with respect to the jth variable x_j is denoted by $\partial_j f$ while the mth partial derivative with respect to the jth variable is denoted by $\partial_j^m f$. The *gradient* of a function f is the vector $\nabla f = (\partial_1 f,\dots,\partial_n f)$. A *multi-index* α is an ordered n-tuple of nonnegative integers. For a multi-index $\alpha = (\alpha_1,\dots,\alpha_n)$, $\partial^\alpha f$ denotes the derivative $\partial_1^{\alpha_1}\cdots\partial_n^{\alpha_n} f$. If $\alpha = (\alpha_1,\dots,\alpha_n)$ is a multi-index, $|\alpha| = \alpha_1+\cdots+\alpha_n$ denotes its size and $\alpha! = \alpha_1!\cdots\alpha_n!$ denotes the product of the factorials of its entries. The number $|\alpha|$ indicates the *total order of differentiation* of $\partial^\alpha f$. The space of functions in $\mathbf{R}^n$ all of whose derivatives of order at most $N \in \mathbf{Z}^+$ are continuous is denoted by $\mathscr{C}^N(\mathbf{R}^n)$ and the space of all *infinitely differentiable functions* on $\mathbf{R}^n$ by $\mathscr{C}^\infty(\mathbf{R}^n)$. The space of $\mathscr{C}^\infty$ functions with compact support on $\mathbf{R}^n$ is denoted by $\mathscr{C}_0^\infty(\mathbf{R}^n)$. This space is nonempty; see Exercise 2.2.1(a).

For $x \in \mathbf{R}^n$ and $\alpha = (\alpha_1,\dots,\alpha_n)$ a multi-index, we set $x^\alpha = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$. Multi-indices will be denoted by the letters $\alpha,\beta,\gamma,\delta,\dots$. It is a simple fact to verify that

$$|x^\alpha| \le c_{n,\alpha}|x|^{|\alpha|}, \tag{2.2.1}$$

for some constant that depends on the dimension n and on α. In fact, $c_{n,\alpha}$ is the maximum of the continuous function $(x_1,\dots,x_n) \mapsto |x_1^{\alpha_1}\cdots x_n^{\alpha_n}|$ on the sphere $\mathbf{S}^{n-1} = \{x \in \mathbf{R}^n : |x| = 1\}$. The converse inequality in (2.2.1) fails. However, the following substitute of the converse of (2.2.1) is of great use: for $k \in \mathbf{Z}^+$ we have

$$|x|^k \le C_{n,k} \sum_{|\beta|=k} |x^\beta| \tag{2.2.2}$$

for all $x \in \mathbf{R}^n \setminus \{0\}$. To prove (2.2.2), take $1/C_{n,k}$ to be the minimum of the function

$$x \mapsto \sum_{|\beta|=k} |x^\beta|$$

on $\mathbf{S}^{n-1}$; this minimum is positive since this function has no zeros on $\mathbf{S}^{n-1}$. A related inequality is

$$(1+|x|)^k \le 2^k(1+C_{n,k}) \sum_{|\beta|\le k} |x^\beta|\,. \tag{2.2.3}$$

This follows from (2.2.2) for $|x| \ge 1$, while for $|x| < 1$ we note that the sum in (2.2.3) is at least one since $|x^{(0,\dots,0)}| = 1$.

We end the preliminaries by noting the validity of the one-dimensional Leibniz rule

$$\frac{d^m}{dt^m}(fg) = \sum_{k=0}^{m} \binom{m}{k} \frac{d^k f}{dt^k}\, \frac{d^{m-k} g}{dt^{m-k}}\,, \tag{2.2.4}$$

for all $\mathscr{C}^m$ functions f, g on $\mathbf{R}$, and its multidimensional analogue

$$\partial^\alpha(fg) = \sum_{\beta\le\alpha} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n} (\partial^\beta f)(\partial^{\alpha-\beta} g)\,, \tag{2.2.5}$$

for f, g in $\mathscr{C}^{|\alpha|}(\mathbf{R}^n)$ for some multi-index α, where the notation $\beta \le \alpha$ in (2.2.5) means that β ranges over all multi-indices satisfying $0 \le \beta_j \le \alpha_j$ for all $1 \le j \le n$. We observe that identity (2.2.5) is easily deduced by repeated application of (2.2.4), which in turn is obtained by induction.

2.2.1 The Class of Schwartz Functions

We now introduce the class of *Schwartz functions* on $\mathbf{R}^n$. Roughly speaking, a function is Schwartz if it is smooth and all of its derivatives decay faster than the reciprocal of any polynomial at infinity. More precisely, we give the following definition.

Definition 2.2.1. A $\mathscr{C}^\infty$ complex-valued function f on $\mathbf{R}^n$ is called a Schwartz function if for every pair of multi-indices α and β there exists a positive constant $C_{\alpha,\beta}$ such that

$$\rho_{\alpha,\beta}(f) = \sup_{x\in\mathbf{R}^n} |x^\alpha \partial^\beta f(x)| = C_{\alpha,\beta} < \infty. \tag{2.2.6}$$

The quantities $\rho_{\alpha,\beta}(f)$ are called the *Schwartz seminorms* of f. The set of all Schwartz functions on $\mathbf{R}^n$ is denoted by $\mathscr{S}(\mathbf{R}^n)$.

Example 2.2.2. The function $e^{-|x|^2}$ is in $\mathscr{S}(\mathbf{R}^n)$ but $e^{-|x|}$ is not, since it fails to be differentiable at the origin. The $\mathscr{C}^\infty$ function $g(x) = (1+|x|^4)^{-a}$, $a > 0$, is not in $\mathscr{S}$ since it decays only like the reciprocal of a fixed polynomial at infinity. The set of all smooth functions with compact support, $\mathscr{C}_0^\infty(\mathbf{R}^n)$, is contained in $\mathscr{S}(\mathbf{R}^n)$.

Remark 2.2.3. If f_1 is in $\mathscr{S}(\mathbf{R}^n)$ and f_2 is in $\mathscr{S}(\mathbf{R}^m)$, then the function of $m+n$ variables $f_1(x_1,\dots,x_n)f_2(x_{n+1},\dots,x_{n+m})$ is in $\mathscr{S}(\mathbf{R}^{n+m})$. If f is in $\mathscr{S}(\mathbf{R}^n)$ and $P(x)$ is a polynomial of n variables, then $P(x)f(x)$ is also in $\mathscr{S}(\mathbf{R}^n)$. If α is a multi-index and f is in $\mathscr{S}(\mathbf{R}^n)$, then $\partial^\alpha f$ is in $\mathscr{S}(\mathbf{R}^n)$. Also note that

$$f \in \mathscr{S}(\mathbf{R}^n) \iff \sup_{x\in\mathbf{R}^n} |\partial^\alpha (x^\beta f(x))| < \infty \qquad \text{for all multi-indices } \alpha, \beta.$$

Remark 2.2.4. The following alternative characterization of Schwartz functions is very useful. A $\mathscr{C}^\infty$ function f is in $\mathscr{S}(\mathbf{R}^n)$ if and only if for all positive integers N and all multi-indices α there exists a positive constant $C_{\alpha,N}$ such that

$$|(\partial^\alpha f)(x)| \le C_{\alpha,N}(1+|x|)^{-N}. \tag{2.2.7}$$

The simple proofs are omitted. We now discuss convergence in $\mathscr{S}(\mathbf{R}^n)$.

Definition 2.2.5. Let f_k, f be in $\mathscr{S}(\mathbf{R}^n)$ for $k = 1,2,\dots$. We say that the sequence f_k converges to f in $\mathscr{S}(\mathbf{R}^n)$ if for all multi-indices α and β we have

$$\rho_{\alpha,\beta}(f_k - f) = \sup_{x\in\mathbf{R}^n} |x^\alpha (\partial^\beta (f_k - f))(x)| \to 0 \qquad \text{as} \quad k\to\infty.$$

For instance, for any fixed $x_0 \in \mathbf{R}^n$, $f(x+x_0/k) \to f(x)$ in $\mathscr{S}(\mathbf{R}^n)$ for any f in $\mathscr{S}(\mathbf{R}^n)$ as $k\to\infty$.

This notion of convergence is compatible with a topology on $\mathscr{S}(\mathbf{R}^n)$ under which the operations $(f,g)\mapsto f+g$, $(a,f)\to af$, and $f\mapsto \partial^\alpha f$ are continuous for all complex scalars a and multi-indices α ($f,g\in\mathscr{S}(\mathbf{R}^n)$). A subbasis for open sets containing 0 in this topology is

$$\{f\in\mathscr{S} :\ \rho_{\alpha,\beta}(f) < r\},$$

for all α, β multi-indices and all $r\in\mathbf{Q}^+$. Observe the following: If $\rho_{\alpha,\beta}(f) = 0$, then $f = 0$. This means that $\mathscr{S}(\mathbf{R}^n)$ is a locally convex topological vector space equipped with the family of seminorms $\rho_{\alpha,\beta}$ that separate points. We refer to Reed and Simon [286] for the pertinent definitions. Since the origin in $\mathscr{S}(\mathbf{R}^n)$ has a countable base, this space is metrizable. In fact, the following is a metric on $\mathscr{S}(\mathbf{R}^n)$:

$$d(f,g) = \sum_{j=1}^{\infty} 2^{-j} \frac{\rho_j(f-g)}{1+\rho_j(f-g)},$$

where ρ_j is an enumeration of all the seminorms $\rho_{\alpha,\beta}$, α and β multi-indices. One may easily verify that $\mathscr{S}$ is complete with respect to the metric d. Indeed, a Cauchy sequence $\{h_j\}_j$ in $\mathscr{S}$ would have to be Cauchy in L^∞ and therefore it would converge uniformly to some function h. The same is true for the sequences $\{\partial^\beta h_j\}_j$ and $\{x^\alpha h_j(x)\}_j$, and the limits of these sequences can be shown to be the functions $\partial^\beta h$ and $x^\alpha h(x)$, respectively. It follows that the sequence $\{h_j\}$ converges to h in $\mathscr{S}$. Therefore, $\mathscr{S}(\mathbf{R}^n)$ is a *Fréchet space (complete metrizable locally convex space).*

We note that convergence in $\mathscr{S}$ is stronger than convergence in all L^p. We have the following.

Proposition 2.2.6. *Let f, f_k, $k = 1,2,3,\dots$, be in $\mathscr{S}(\mathbf{R}^n)$. If $f_k\to f$ in $\mathscr{S}$ then $f_k \to f$ in L^p for all $0<p\le\infty$. Moreover, there exists a $C_{p,n}>0$ such that*

$$\big\|\partial^\beta f\big\|_{L^p} \le C_{p,n} \sum_{|\alpha|\le[\frac{n+1}{p}]+1} \rho_{\alpha,\beta}(f) \tag{2.2.8}$$

for all f for which the right-hand side is finite.

Proof. Observe that when $p<\infty$ we have

$$\begin{aligned}\big\|\partial^\beta f\big\|_{L^p} &\le \left[\int_{|x|\le 1} |\partial^\beta f(x)|^p\,dx + \int_{|x|\ge 1} |x|^{n+1}|\partial^\beta f(x)|^p |x|^{-(n+1)}\,dx\right]^{1/p}\\ &\le \left[v_n\big\|\partial^\beta f\big\|_{L^\infty}^p + \Big(\sup_{|x|\ge 1}|x|^{n+1}|\partial^\beta f(x)|^p\Big)\int_{|x|\ge 1}|x|^{-(n+1)}\,dx\right]^{1/p}\\ &\le C_{p,n}\Big(\big\|\partial^\beta f\big\|_{L^\infty} + \sup_{|x|\ge 1}\big(|x|^{[\frac{n+1}{p}]+1}|\partial^\beta f(x)|\big)\Big).\end{aligned}$$

The preceding inequality is also trivially valid when $p=\infty$. Now set $m=\left[\frac{n+1}{p}\right]+1$ and use (2.2.2) to obtain

$$\sup_{|x|\ge 1}|x|^m|\partial^\beta f(x)| \le \sup_{|x|\ge 1} C_{n,m}\sum_{|\alpha|=m}|x^\alpha\partial^\beta f(x)| \le C_{n,m}\sum_{|\alpha|\le m}\rho_{\alpha,\beta}(f).$$

Conclusion (2.2.8) now follows immediately. This shows that convergence in $\mathscr{S}$ implies convergence in L^p. □

We now show that the Schwartz class is closed under certain operations.

Proposition 2.2.7. *Let f, g be in $\mathscr{S}(\mathbf{R}^n)$. Then fg and $f*g$ are in $\mathscr{S}(\mathbf{R}^n)$. Moreover,*

$$\partial^\alpha(f*g) = (\partial^\alpha f)*g = f*(\partial^\alpha g) \tag{2.2.9}$$

for all multi-indices α.

Proof. Fix f and g in $\mathscr{S}(\mathbf{R}^n)$. Let e_j be the unit vector $(0,\dots,1,\dots,0)$ with 1 in the jth entry and zeros in all the other entries. Since

$$\frac{f(y+he_j)-f(y)}{h} - (\partial_j f)(y) \to 0 \tag{2.2.10}$$

as $h\to 0$, and since the expression in (2.2.10) is pointwise bounded by some constant depending on f, the integral of the expression in (2.2.10) with respect to the measure $g(x-y)\,dy$ converges to zero as $h\to 0$ by the Lebesgue dominated convergence theorem. This proves (2.2.9) when $\alpha=(0,\dots,1,\dots,0)$. The general case follows by repeating the previous argument and using induction.

We now show that the convolution of two functions in $\mathscr{S}$ is also in $\mathscr{S}$. For each $N>0$ there is a constant C_N such that

$$\left|\int_{\mathbf{R}^n} f(x-y)g(y)\,dy\right| \le C_N\int_{\mathbf{R}^n}(1+|x-y|)^{-N}(1+|y|)^{-N-n-1}dy. \tag{2.2.11}$$

Inserting the simple estimate

$$(1+|x-y|)^{-N} \le (1+|y|)^{N}(1+|x|)^{-N}$$

in (2.2.11) we obtain that

$$|(f*g)(x)| \le C_N(1+|x|)^{-N}\int_{\mathbf{R}^n}(1+|y|)^{-n-1}dy = C_N'(1+|x|)^{-N}.$$

This shows that $f*g$ decays like $(1+|x|)^{-N}$ at infinity, but since $N>0$ is arbitrary it follows that $f*g$ decays faster than the reciprocal of any polynomial.

Since $\partial^\alpha(f*g) = (\partial^\alpha f)*g$, replacing f by $\partial^\alpha f$ in the previous argument, we also conclude that all the derivatives of $f*g$ decay faster than the reciprocal of any polynomial at infinity. Using (2.2.7), we conclude that $f*g$ is in $\mathscr{S}$. Finally, the fact that fg is in $\mathscr{S}$ follows directly from Leibniz's rule (2.2.5) and (2.2.7). □

2.2.2 The Fourier Transform of a Schwartz Function

The Fourier transform is often introduced as an operation on L^1. In that setting, problems of convergence arise when certain manipulations of functions are performed. Also, Fourier inversion requires the additional assumption that the Fourier transform is in L^1. Here we initially introduce the Fourier transform on the space of Schwartz functions. The rapid decay of Schwartz functions at infinity allows us to develop its fundamental properties without encountering any convergence problems. The Fourier transform is a homeomorphism of the Schwartz class and Fourier inversion holds in it. For these reasons, this class is a natural environment for it.

For $x=(x_1,\dots,x_n)$, $y=(y_1,\dots,y_n)$ in $\mathbf{R}^n$ we use the notation

$$x\cdot y = \sum_{j=1}^{n} x_j y_j .$$

Definition 2.2.8. Given f in $\mathscr{S}(\mathbf{R}^n)$ we define

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

We call $\widehat{f}$ the Fourier transform of f.

Example 2.2.9. If $f(x)=e^{-\pi|x|^2}$ defined on $\mathbf{R}^n$, then $\widehat{f}(\xi)=f(\xi)$. To prove this, observe that the function

$$s \mapsto \int_{-\infty}^{+\infty} e^{-\pi(t+is)^2}dt, \qquad s\in\mathbf{R},$$

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

$$\int_{-\infty}^{+\infty} -2\pi i(t+is)e^{-\pi(t+is)^2}\,dt = \int_{-\infty}^{+\infty} i\frac{d}{dt}(e^{-\pi(t+is)^2})\,dt = 0\,.$$

Using this fact, we calculate the Fourier transform of the function $t \mapsto e^{-\pi t^2}$ on $\mathbf{R}$ by completing the squares as follows:

$$\int_{\mathbf{R}} e^{-\pi t^2} e^{-2\pi i t\tau}\,dt = \int_{\mathbf{R}} e^{-\pi(t+i\tau)^2} e^{\pi(i\tau)^2}\,dt = \left(\int_{-\infty}^{+\infty} e^{-\pi t^2}\,dt\right) e^{-\pi\tau^2} = e^{-\pi\tau^2},$$

where $\tau \in \mathbf{R}$, and we used that

$$\int_{-\infty}^{+\infty} e^{-t^2}\,dt = \sqrt{\pi}\,, \tag{2.2.12}$$

a fact that can be found in Appendix A.1.

Remark 2.2.10. It follows from the definition of the Fourier transform that if f is in $\mathscr{S}(\mathbf{R}^n)$ and g is in $\mathscr{S}(\mathbf{R}^m)$, then

$$[f(x_1,\dots,x_n)g(x_{n+1},\dots,x_{n+m})]\widehat{\ }= \widehat{f}(\xi_1,\dots,\xi_n)\widehat{g}(\xi_{n+1},\dots,\xi_{n+m}),$$

where the first $\widehat{\ }$ denotes the Fourier transform on $\mathbf{R}^{n+m}$. In other words, the Fourier transform preserves separation of variables. Combining this observation with the result in Example 2.2.9, we conclude that the function $f(x) = e^{-\pi|x|^2}$ defined on $\mathbf{R}^n$ is equal to its Fourier transform.

We now continue with some properties of the Fourier transform. Before we do this we introduce some notation. For a measurable function f on $\mathbf{R}^n$, $x \in \mathbf{R}^n$, and $a > 0$ we define the *translation*, *dilation*, and *reflection* of f by

$$\begin{aligned} (\tau^y f)(x) &= f(x-y) \\ (\delta^a f)(x) &= f(ax) \\ \widetilde{f}(x) &= f(-x). \end{aligned} \tag{2.2.13}$$

Also recall the notation $f_a = a^{-n}\delta^{1/a}(f)$ introduced in Definition 2.1.9.

Proposition 2.2.11. *Given f, g in $\mathscr{S}(\mathbf{R}^n)$, $y \in \mathbf{R}^n$, $b \in \mathbf{C}$, α a multi-index, and $t > 0$, we have*

(1) $\|\widehat{f}\|_{L^\infty} \le \|f\|_{L^1}$,

(2) $\widehat{f+g} = \widehat{f} + \widehat{g}$,

(3) $\widehat{bf} = b\widehat{f}$,

(4) $\widehat{\widetilde{f}} = \widetilde{\widehat{f}}$,

(5) $\widehat{\overline{f}} = \overline{\widetilde{\widehat{f}}}$,

(6) $\widehat{\tau^y f}(\xi) = e^{-2\pi i y\cdot\xi}\widehat{f}(\xi)$,

(7) $(e^{2\pi i x\cdot y} f(x))\widehat{\ }(\xi) = \tau^y(\widehat{f})(\xi)$,

(8) $(\delta^t f)\widehat{\ } = t^{-n}\delta^{t^{-1}}\widehat{f} = (\widehat{f})_t$,

(9) $(\partial^\alpha f)\widehat{\ }(\xi) = (2\pi i\xi)^\alpha \widehat{f}(\xi)$,

(10) $(\partial^\alpha \widehat{f})(\xi) = ((-2\pi i x)^\alpha f(x))\widehat{\ }(\xi)$,

(11) $\widehat{f} \in \mathscr{S}$,

(12) $\widehat{f * g} = \widehat{f}\,\widehat{g}$,

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

Proof. Property (1) follows directly from Definition 2.2.8. Properties (2)–(5) are trivial. Properties (6)–(8) require a suitable change of variables but they are omitted. Property (9) is proved by integration by parts (which is justified by the rapid decay of the integrands):

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

To prove (10), let $\alpha = e_j = (0,\dots,1,\dots,0)$, where all entries are zero except for the jth entry, which is 1. Since

$$\frac{e^{-2\pi i x\cdot(\xi+he_j)} - e^{-2\pi i x\cdot\xi}}{h} - (-2\pi i x_j)e^{-2\pi i x\cdot\xi} \to 0 \tag{2.2.14}$$

as $h\to 0$ and the preceding function is bounded by $C|x|$ for all h and ξ, the Lebesgue dominated convergence theorem implies that the integral of the function in (2.2.14) with respect to the measure $f(x)dx$ converges to zero. This proves (10) for $\alpha = e_j$. For other α's use induction. To prove (11) we use (9), (10), and (1) in the following way:

$$\big\|x^\alpha(\partial^\beta \widehat{f})(x)\big\|_{L^\infty} = \frac{(2\pi)^{|\beta|}}{(2\pi)^{|\alpha|}}\big\|(\partial^\alpha(x^\beta f(x)))\widehat{\ }\big\|_{L^\infty} \le \frac{(2\pi)^{|\beta|}}{(2\pi)^{|\alpha|}}\big\|\partial^\alpha(x^\beta f(x))\big\|_{L^1} < \infty\,.$$

Identity (12) follows from the following calculation:

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

where the application of Fubini's theorem is justified by the absolute convergence of the integrals. Finally, we prove (13). We have

$$
\begin{aligned}
\widehat{f\circ A}(\xi) &= \int_{\mathbf{R}^n} f(Ax)e^{-2\pi i x\cdot\xi}\,dx \\
&= \int_{\mathbf{R}^n} f(y)e^{-2\pi i A^{-1}y\cdot\xi}\,dy \\
&= \int_{\mathbf{R}^n} f(y)e^{-2\pi i A^t y\cdot\xi}\,dy \\
&= \int_{\mathbf{R}^n} f(y)e^{-2\pi i y\cdot A\xi}\,dy \\
&= \widehat{f}(A\xi),
\end{aligned}
$$

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

Corollary 2.2.12. *The Fourier transform of a radial function is radial. Products and convolutions of radial functions are radial.*

Proof. Let ξ_1, ξ_2 in $\mathbf{R}^n$ with $|\xi_1| = |\xi_2|$. Then for some orthogonal matrix A we have $A\xi_1 = \xi_2$. Since f is radial, we have $f = f\circ A$. Then

$$\widehat{f}(\xi_2) = \widehat{f}(A\xi_1) = \widehat{f\circ A}(\xi_1) = \widehat{f}(\xi_1),$$

where we used (13) in Proposition 2.2.11 to justify the second equality. Products and convolutions of radial functions are easily seen to be radial. □

2.2.3 The Inverse Fourier Transform and Fourier Inversion

We now define the inverse Fourier transform.

Definition 2.2.13. Given a Schwartz function f, we define

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

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

$$f \mapsto f^\vee$$

is called the *inverse Fourier transform*.

It is straightforward that the inverse Fourier transform shares the same properties as the Fourier transform. One may want to list (and prove) properties for the inverse Fourier transform analogous to those in Proposition 2.2.11.

We now investigate the relation between the Fourier transform and the inverse Fourier transform. In the next theorem, we prove that one is the inverse operation of the other. This property is referred to as *Fourier inversion*.

Theorem 2.2.14. *Given f, g, and h in $\mathscr{S}(\mathbf{R}^n)$, we have*

(1) $\displaystyle\int_{\mathbf{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbf{R}^n} \widehat{f}(x)g(x)\,dx,$

(2) ***(Fourier Inversion)*** $\quad (\widehat{f})^\vee = f = (f^\vee)^\wedge\,,$

(3) ***(Parseval's relation)*** $\displaystyle\int_{\mathbf{R}^n} f(x)\overline{h}(x)\,dx = \int_{\mathbf{R}^n} \widehat{f}(\xi)\overline{\widehat{h}(\xi)}\,d\xi\,,$

(4) ***(Plancherel's identity)*** $\quad \|f\|_{L^2} = \|\widehat{f}\|_{L^2} = \|f^\vee\|_{L^2}\,,$

(5) $\displaystyle\int_{\mathbf{R}^n} f(x)h(x)\,dx = \int_{\mathbf{R}^n} \widehat{f}(x)h^\vee(x)\,dx.$

Proof. (1) follows immediately from the definition of the Fourier transform and Fubini's theorem. To prove (2) we use (1) with

$$g(\xi) = e^{2\pi i\xi\cdot t}e^{-\pi|\varepsilon\xi|^2}\,.$$

By Proposition 2.2.11 (7) and (8) and Example 2.2.9, we have that

$$\widehat{g}(x) = \frac{1}{\varepsilon^n}e^{-\pi|(x-t)/\varepsilon|^2},$$

which is an approximate identity. Now (1) gives

$$\int_{\mathbf{R}^n} f(x)\varepsilon^{-n}e^{-\pi\varepsilon^{-2}|x-t|^2}\,dx = \int_{\mathbf{R}^n} \widehat{f}(\xi)e^{2\pi i\xi\cdot t}e^{-\pi|\varepsilon\xi|^2}\,d\xi\,. \tag{2.2.15}$$

Now let $\varepsilon \to 0$ in (2.2.15). The left-hand side of (2.2.15) converges to $f(t)$ uniformly on compact sets by Theorem 1.2.19. The right-hand side of (2.2.15) converges to $(\widehat{f})^\vee(t)$ as $\varepsilon \to 0$ by the Lebesgue dominated convergence theorem. We conclude that $(\widehat{f})^\vee = f$ on $\mathbf{R}^n$. Replacing f by $\widetilde{f}$ and using the result just proved, we conclude that $(f^\vee)^\wedge = f$.

Note that if $g = \overline{\widehat{h}}$, then Proposition 2.2.11 (5) and identity (2) imply that $\widehat{g} = \overline{h}$. Then (3) follows from (1) by expressing h in terms of g. Identity (4) is a trivial consequence of (3). (Sometimes the polarized identity (3) is also referred to as Plancherel's identity.) Finally, (5) easily follows from (1) and (2) with $\widehat{g} = h$. $\square$

Next we have the following simple corollary of Theorem 2.2.14.

Corollary 2.2.15. *The Fourier transform is a homeomorphism from $\mathscr{S}(\mathbf{R}^n)$ onto itself.*

Proof. The continuity of the Fourier transform (and its inverse) follows from Exercise 2.2.2, while Fourier inversion yields that this map is bijective. $\square$

2.2.4 The Fourier Transform on $L^1 + L^2$

We have defined the Fourier transform on $\mathscr{S}(\mathbf{R}^n)$. We now extend this definition to the space $L^1(\mathbf{R}^n) + L^2(\mathbf{R}^n)$.

We begin by observing that the Fourier transform given in Definition 2.2.8,

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

makes sense as a convergent integral for functions $f \in L^1(\mathbf{R}^n)$. This allows us to extend the definition of the Fourier transform on L^1. Moreover, this operator satisfies properties (1)–(8) as well as (12) and (13) in Proposition 2.2.11, with f, g integrable. We also define the inverse Fourier transform on L^1 by setting $f^\vee(x) = \widehat{f}(-x)$ for $f \in L^1(\mathbf{R}^n)$ and we note that analogous properties hold for it. One problem in this generality is that when f is integrable, one may not necessarily have $(\widehat{f})^\vee = f$ a.e. This inversion is possible when $\widehat{f}$ is also integrable; see Exercise 2.2.6.

The integral defining the Fourier transform does not converge absolutely for functions in $L^2(\mathbf{R}^n)$; however, the Fourier transform has a natural definition in this space accompanied by an elegant theory. In view of the result in Exercise 2.2.8, the Fourier transform is an L^2 isometry on $L^1 \cap L^2$, which is a dense subspace of L^2. By density, there is a unique bounded extension of the Fourier transform on L^2. Let us denote this extension by $\mathscr{F}$. Then $\mathscr{F}$ is also an isometry on L^2, i.e.,

$$\big\|\mathscr{F}(f)\big\|_{L^2} = \big\|f\big\|_{L^2}$$

for all $f \in L^2(\mathbf{R}^n)$, and any sequence of functions $f_N \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ converging to a given f in $L^2(\mathbf{R}^n)$ satisfies

$$\big\|\widehat{f_N} - \mathscr{F}(f)\big\|_{L^2} \to 0\,, \tag{2.2.16}$$

as $N \to \infty$. In particular, the sequence of functions $f_N(x) = f(x)\chi_{|x|\le N}$ yields that

$$\widehat{f_N}(\xi) = \int_{|x|\le N} f(x) e^{-2\pi i x\cdot\xi}\, dx \tag{2.2.17}$$

converges to $\mathscr{F}(f)(\xi)$ in L^2 as $N \to \infty$. If f is both integrable and square integrable, the expressions in (2.2.17) also converge to $\widehat{f}(\xi)$ pointwise. Also, in view of Theorem 1.1.11 and (2.2.16), there is a subsequence of $\widehat{f_N}$ that converges to $\mathscr{F}(f)$ pointwise a.e. Consequently, for f in $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ the expressions $\widehat{f}$ and $\mathscr{F}(f)$ coincide pointwise a.e. For this reason we often adopt the notation $\widehat{f}$ to denote the Fourier transform of functions f in L^2 as well.

In a similar fashion, we let $\mathscr{F}'$ be the isometry on $L^2(\mathbf{R}^n)$ that extends the operator $f \mapsto f^\vee$, which is an L^2 isometry on $L^1 \cap L^2$; the last statement follows by adapting the result of Exercise 2.2.8 to the inverse Fourier transform. Since $\varphi^\vee(x) = \widehat{\varphi}(-x)$ for φ in the Schwartz class, which is dense in L^2 (Exercise 2.2.5), it follows that $\mathscr{F}'(f)(x) = \mathscr{F}(f)(-x)$ for all $f \in L^2$ and almost all $x \in \mathbf{R}^n$. The operators $\mathscr{F}$ and $\mathscr{F}'$ are L^2-isometries that satisfy $\mathscr{F}' \circ \mathscr{F} = \mathscr{F} \circ \mathscr{F}' = \mathrm{Id}$ on the Schwartz space. By density this identity also holds for L^2 functions and implies that $\mathscr{F}$ and $\mathscr{F}'$ are injective and surjective mappings from L^2 to itself; consequently, $\mathscr{F}'$ coincides with the inverse operator $\mathscr{F}^{-1}$ of $\mathscr{F} : L^2 \to L^2$, and Fourier inversion

$$f = \mathscr{F}^{-1} \circ \mathscr{F}(f) = \mathscr{F} \circ \mathscr{F}^{-1}(f) \qquad \text{a.e.}$$

holds on L^2.

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 $L^p(\mathbf{R}^n)$, with $1 < p < 2$, we define $\widehat{f} = \widehat{f_1} + \widehat{f_2}$, where $f_1 \in L^1(\mathbf{R}^n)$, $f_2 \in L^2(\mathbf{R}^n)$, and $f = f_1 + f_2$; we may take, for instance, $f_1 = f\chi_{|f|>1}$ and $f_2 = f\chi_{|f|\le 1}$. The definition of $\widehat{f}$ is independent of the choice of f_1 and f_2, for if $f_1 + f_2 = h_1 + h_2$ for $f_1, h_1 \in L^1(\mathbf{R}^n)$ and $f_2, h_2 \in L^2(\mathbf{R}^n)$, we have $f_1 - h_1 = h_2 - f_2 \in L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$. Since these functions are equal on $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$, their Fourier transforms are also equal, and we obtain $\widehat{f_1} - \widehat{h_1} = \widehat{h_2} - \widehat{f_2}$, 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.

Proposition 2.2.16. ***(Hausdorff–Young inequality)*** *For every function f in $L^p(\mathbf{R}^n)$ we have the estimate*

$$\|\widehat{f}\|_{L^{p'}} \le \|f\|_{L^p}$$

whenever $1 \le p \le 2$.

Proof. This follows easily from Theorem 1.3.4. Interpolate between the estimates $\|\widehat{f}\|_{L^\infty} \le \|f\|_{L^1}$ (Proposition 2.2.11 (1)) and $\|\widehat{f}\|_{L^2} \le \|f\|_{L^2}$ to obtain $\|\widehat{f}\|_{L^{p'}} \le \|f\|_{L^p}$. We conclude that the Fourier transform is a bounded operator from $L^p(\mathbf{R}^n)$ to $L^{p'}(\mathbf{R}^n)$ with norm at most 1 when $1 \le p \le 2$. □

Next, we are concerned with the behavior of the Fourier transform at infinity.

Proposition 2.2.17. ***(Riemann–Lebesgue lemma)*** *For a function f in $L^1(\mathbf{R}^n)$ we have that*

$$|\widehat{f}(\xi)| \to 0 \qquad \textit{as} \qquad |\xi| \to \infty.$$

Proof. Consider the function $\chi_{[a,b]}$ on $\mathbf{R}$. A simple computation gives

$$\widehat{\chi_{[a,b]}}(\xi) = \int_a^b e^{-2\pi i x\xi}\,dx = \frac{e^{-2\pi i\xi a} - e^{-2\pi i\xi b}}{2\pi i\xi},$$

which tends to zero as $|\xi| \to \infty$. Likewise, if $g = \prod_{j=1}^n \chi_{[a_j,b_j]}$ on $\mathbf{R}^n$, then

$$\widehat{g}(\xi) = \prod_{j=1}^n \frac{e^{-2\pi i\xi_j a_j} - e^{-2\pi i\xi_j b_j}}{2\pi i\xi_j}.$$

Given a $\xi = (\xi_1,\dots,\xi_n) \neq 0$, there is j_0 such that $|\xi_{j_0}| > |\xi|/\sqrt{n}$. Then

$$\left|\prod_{j=1}^n \frac{e^{-2\pi i\xi_j a_j} - e^{-2\pi i\xi_j b_j}}{2\pi i\xi_j}\right| \le \frac{2\sqrt{n}}{2\pi|\xi|} \sup_{1\le j_0\le n} \prod_{j\neq j_0}(b_j - a_j)$$

which also tends to zero as $|\xi| \to \infty$ in $\mathbf{R}^n$.

Given a general integrable function f on $\mathbf{R}^n$ and $\varepsilon > 0$, there is a simple function h, which is a finite linear combination of characteristic functions of rectangles (like g), such that $\|f-h\|_{L^1} < \frac{\varepsilon}{2}$. Then there is an M is such that for $|\xi| > M$ we have $|\widehat{h}(\xi)| < \frac{\varepsilon}{2}$. It follows that

$$|\widehat{f}(\xi)| \le |\widehat{f}(\xi) - \widehat{h}(\xi)| + |\widehat{h}(\xi)| \le \big\|f-h\big\|_{L^1} + |\widehat{h}(\xi)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2},$$

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

A different proof can be given by taking the function h in the preceding paragraph to be a Schwartz function and using that Schwartz functions are dense in $L^1(\mathbf{R}^n)$; see Exercise 2.2.5 about the last assertion. □

We end this section with an example that illustrates some of the practical uses of the Fourier transform.

Example 2.2.18. We would like to find a Schwartz function $f(x_1,x_2,x_3)$ on $\mathbf{R}^3$ that satisfies the partial differential equation

$$f(x) + \partial_1^2\partial_2^2\partial_3^4 f(x) + 4i\partial_1^2 f(x) + \partial_2^7 f(x) = e^{-\pi|x|^2}.$$

Taking the Fourier transform on both sides of this identity and using Proposition 2.2.11 (2), (9) and the result of Example 2.2.9, we obtain

$$\widehat{f}(\xi)\Big[1 + (2\pi i\xi_1)^2(2\pi i\xi_2)^2(2\pi i\xi_3)^4 + 4i(2\pi i\xi_1)^2 + (2\pi i\xi_2)^7\Big] = e^{-\pi|\xi|^2}.$$

Let $p(\xi) = p(\xi_1,\xi_2,\xi_3)$ be the polynomial inside the square brackets. We observe that $p(\xi)$ has no real zeros and we may therefore write

$$\widehat{f}(\xi) = e^{-\pi|\xi|^2}p(\xi)^{-1} \implies f(x) = \big(e^{-\pi|\xi|^2}p(\xi)^{-1}\big)^{\vee}(x).$$

In general, let

$$P(\xi) = \sum_{|\alpha| \le N} C_\alpha \xi^\alpha$$

be a polynomial in $\mathbf{R}^n$ with constant complex coefficients C_α indexed by multi-indices α. If $P(2\pi i\xi)$ has no real zeros, and u is in $\mathscr{S}(\mathbf{R}^n)$, then the partial differential equation

$$P(\partial)f = \sum_{|\alpha| \le N} C_\alpha \partial^\alpha f = u$$

is solved as before to give

$$f = \big(\widehat{u}(\xi)P(2\pi i\xi)^{-1}\big)^{\vee}.$$

Since $P(2\pi i\xi)$ has no real zeros and $u \in \mathscr{S}(\mathbf{R}^n)$, the function

$$\widehat{u}(\xi)P(2\pi i\xi)^{-1}$$

is smooth and therefore a Schwartz function. Then f is also in $\mathscr{S}(\mathbf{R}^n)$ by Proposition 2.2.11 (11).

Exercises

2.2.1. (a) Construct a Schwartz function supported in the unit ball of $\mathbf{R}^n$.
(b) Construct a $\mathscr{C}_0^\infty(\mathbf{R}^n)$ function equal to 1 on the annulus $1 \le |x| \le 2$ and vanishing off the annulus $1/2 \le |x| \le 4$.
(c) Construct a nonnegative nonzero Schwartz function f on $\mathbf{R}^n$ whose Fourier transform is nonnegative and compactly supported.
[*Hint:* Part (a): Try the construction in dimension one first using the $\mathscr{C}^\infty$ function $\eta(x) = e^{-1/x}$ for $x > 0$ and $\eta(x) = 0$ for $x < 0$. Part (c): Take $f = |\phi * \widetilde{\phi}|^2$, where $\widehat{\phi}$ is odd, real-valued, and compactly supported; here $\widetilde{\phi}(x) = \phi(-x)$.]

2.2.2. If $f_k, f \in \mathscr{S}(\mathbf{R}^n)$ and $f_k \to f$ in $\mathscr{S}(\mathbf{R}^n)$, then $\widehat{f_k} \to \widehat{f}$ and $f_k^\vee \to f^\vee$ in $\mathscr{S}(\mathbf{R}^n)$.

2.2.3. Find the *spectrum* (i.e., the set of all *eigenvalues* of the Fourier transform), that is, all complex numbers λ for which there exist nonzero functions f such that

$$\widehat{f} = \lambda f.$$

[*Hint:* Apply the Fourier transform three times to the preceding identity. Consider the functions $xe^{-\pi x^2}$, $(a + bx^2)e^{-\pi x^2}$, and $(cx + dx^3)e^{-\pi x^2}$ for suitable a, b, c, d to show that all fourth roots of unity are indeed eigenvalues of the Fourier transform.]

2.2.4. Use the idea of the proof of Proposition 2.2.7 to show that if the functions f, g defined on $\mathbf{R}^n$ satisfy $|f(x)| \le A(1+|x|)^{-M}$ and $|g(x)| \le B(1+|x|)^{-N}$ for some $M, N > n$, then

$$|(f*g)(x)| \le ABC(1+|x|)^{-L},$$

where $L = \min(N,M)$ and $C = C(N,M) > 0$.

2.2.5. Show that $\mathscr{C}_0^\infty(\mathbf{R}^n)$ is dense on $L^p(\mathbf{R}^n)$ for $0 < p < \infty$ but not for $p = \infty$.
$\big[$*Hint:* Use a smooth approximate identity when $p \ge 1$. Reduce the case $p < 1$ to $p = 1$.$\big]$

2.2.6. (a) Prove that if $f \in L^1$, then $\widehat{f}$ is uniformly continuous on $\mathbf{R}^n$.
(b) Prove that for $f, g \in L^1(\mathbf{R}^n)$ we have

$$\int_{\mathbf{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbf{R}^n} \widehat{f}(y)g(y)\,dy.$$

(c) Take $\widehat{g}(x) = \varepsilon^{-n} e^{-\pi\varepsilon^{-2}|x-t|^2}$ in (b) and let $\varepsilon \to 0$ to prove that if f and $\widehat{f}$ are both in L^1, then $(\widehat{f})^\vee = f$ a.e. This fact is called *Fourier inversion on* L^1.

2.2.7. (a) Prove that if a function f in $L^1(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ is continuous at 0, then

$$\lim_{\varepsilon \to 0} \int_{\mathbf{R}^n} \widehat{f}(x) e^{-\pi|\varepsilon x|^2}\,dx = f(0).$$

(b) Let $f \in L^1(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ be continuous at zero and satisfy $\widehat{f} \ge 0$. Show that $\widehat{f}$ is in L^1 and conclude that Fourier inversion holds at zero $f(0) = \|\widehat{f}\|_{L^1}$, and also $f = (\widehat{f})^\vee$ a.e. in general.
$\big[$*Hint:* Part (a): Let $g(x) = e^{-\pi|\varepsilon x|^2}$ in Exercise 2.2.6(b) and use Theorem 1.2.19 (2).$\big]$

2.2.8. Given f in $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$, without appealing to density, prove that

$$\|\widehat{f}\|_{L^2} = \|f\|_{L^2}.$$

$\big[$*Hint:* Let $h = f * \overline{\widetilde{f}}$, where $\widetilde{f}(x) = f(-x)$ and the bar indicates complex conjugation. Then $h \in L^1(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$, $\widehat{h} = |\widehat{f}|^2 \ge 0$, and h is continuous at zero. Exercise 2.2.7(b) yields $\|\widehat{f}\|_{L^2}^2 = \|\widehat{h}\|_{L^1} = h(0) = \int_{\mathbf{R}^n} f(x)\overline{\widetilde{f}(-x)}\,dx = \|f\|_{L^2}^2$.$\big]$

2.2.9. (a) Prove that for all $0 < \varepsilon < t < \infty$ we have

$$\left| \int_\varepsilon^t \frac{\sin(\xi)}{\xi}\,d\xi \right| \le 4.$$

(b) If f is an odd L^1 function on the line, conclude that for all $t > \varepsilon > 0$ we have

$$\left| \int_\varepsilon^t \frac{\widehat{f}(\xi)}{\xi}\,d\xi \right| \le 4\|f\|_{L^1}.$$

(c) Let $g(\xi)$ be a continuous odd function that is equal to $1/\log(\xi)$ for $\xi \ge 2$. Show that there does not exist an L^1 function whose Fourier transform is g.

2.2.10. Let f be in $L^1(\mathbf{R})$. Prove that

$$\int_{-\infty}^{+\infty} f\Big(x-\frac{1}{x}\Big)\,dx = \int_{-\infty}^{+\infty} f(u)\,du\,.$$

$\big[$*Hint:* For $x\in(-\infty,0)$ use the change of variables $u=x-\frac{1}{x}$ or $x=\frac{1}{2}\big(u-\sqrt{4+u^2}\,\big)$. For $x\in(0,\infty)$ use the change of variables $u=x-\frac{1}{x}$ or $x=\frac{1}{2}\big(u+\sqrt{4+u^2}\,\big)$.$\big]$

2.2.11. (a) Use Exercise 2.2.10 with $f(x)=e^{-tx^2}$ to obtain the *subordination* identity

$$e^{-2t}=\frac{1}{\sqrt{\pi}}\int_0^\infty e^{-y-t^2/y}\,\frac{dy}{\sqrt{y}}\,,\qquad \text{where } t>0.$$

(b) Set $t=\pi|x|$ and integrate with respect to $e^{-2\pi i\xi\cdot x}dx$ to prove that

$$(e^{-2\pi|x|})\widehat{\ }(\xi)=\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\frac{1}{(1+|\xi|^2)^{\frac{n+1}{2}}}\,.$$

This calculation gives the Fourier transform of the Poisson kernel.

2.2.12. Let $1\le p\le\infty$ and let p' be its dual index.
(a) Prove that Schwartz functions f on the line satisfy the estimate

$$\big\|f\big\|_{L^\infty}^2\le 2\big\|f\big\|_{L^p}\big\|f'\big\|_{L^{p'}}\,.$$

(b) Prove that all Schwartz functions f on $\mathbf{R}^n$ satisfy the estimate

$$\big\|f\big\|_{L^\infty}^2\le\sum_{|\alpha+\beta|=n}\big\|\partial^\alpha f\big\|_{L^p}\big\|\partial^\beta f\big\|_{L^{p'}}\,,$$

where the sum is taken over all pairs of multi-indices α and β whose sum has size n.
$\big[$*Hint:* Part (a): Write $f(x)^2=\int_{-\infty}^x\frac{d}{dt}f(t)^2\,dt$.$\big]$

2.2.13. The *uncertainty principle* says that the position and the momentum of a particle cannot be simultaneously localized. Prove the following inequality, which presents a quantitative version of this principle:

$$\big\|f\big\|_{L^2(\mathbf{R}^n)}^2\le\frac{4\pi}{n}\inf_{y\in\mathbf{R}^n}\bigg[\int_{\mathbf{R}^n}|x-y|^2|f(x)|^2\,dx\bigg]^{\frac12}\inf_{z\in\mathbf{R}^n}\bigg[\int_{\mathbf{R}^n}|\xi-z|^2|\widehat{f}(\xi)|^2\,d\xi\bigg]^{\frac12},$$

where f is a Schwartz function on $\mathbf{R}^n$ (or an L^2 function with sufficient decay at infinity).
$\big[$*Hint:* Let y be in $\mathbf{R}^n$. Start with

$$\big\|f\big\|_{L^2}^2=\frac{1}{n}\int_{\mathbf{R}^n}f(x)\overline{f(x)}\sum_{j=1}^n\frac{\partial}{\partial x_j}(x_j-y_j)\,dx\,,$$

integrate by parts, apply the Cauchy–Schwarz inequality, Plancherel's identity, and the identity $\sum_{j=1}^n |\widehat{\partial_j f}(\xi)|^2 = 4\pi^2|\xi|^2|\widehat{f}(\xi)|^2$ for all $\xi \in \mathbf{R}^n$. Then replace $f(x)$ by $f(x)e^{2\pi i x\cdot z}$.$\big]$

2.2.14. Let $-\infty < \alpha < \frac{n}{2} < \beta < +\infty$. Prove the validity of the following inequality:

$$\big\|g\big\|_{L^1(\mathbf{R}^n)} \le C\,\big\||x|^\alpha g(x)\big\|_{L^2(\mathbf{R}^n)}^{\frac{\beta-n/2}{\beta-\alpha}}\big\||x|^\beta g(x)\big\|_{L^2(\mathbf{R}^n)}^{\frac{n/2-\alpha}{\beta-\alpha}}$$

for some constant $C = C(n,\alpha,\beta)$ independent of g.
$\big[$*Hint:* First prove $\|g\|_{L^1} \le C\|\,|x|^\alpha g(x)\|_{L^2} + \|\,|x|^\beta g(x)\|_{L^2}$ and then replace $g(x)$ by $g(\lambda x)$ for some suitable $\lambda > 0$.$\big]$

2.3 The Class of Tempered Distributions

The fundamental idea of the theory of distributions is that it is generally easier to work with linear functionals acting on spaces of "nice" functions than to work with "bad" functions directly. The set of "nice" functions we consider is closed under the basic operations in analysis, and these operations are extended to distributions by duality. This wonderful interpretation has proved to be an indispensable tool that has clarified many situations in analysis.

2.3.1 Spaces of Test Functions

We recall the space $\mathscr{C}_0^\infty(\mathbf{R}^n)$ of all smooth functions with compact support, and $\mathscr{C}^\infty(\mathbf{R}^n)$ of all smooth functions on $\mathbf{R}^n$. We are mainly interested in the three spaces of "nice" functions on $\mathbf{R}^n$ that are nested as follows:

$$\mathscr{C}_0^\infty(\mathbf{R}^n) \subseteq \mathscr{S}(\mathbf{R}^n) \subseteq \mathscr{C}^\infty(\mathbf{R}^n).$$

Here $\mathscr{S}(\mathbf{R}^n)$ is the space of Schwartz functions introduced in Section 2.2.

Definition 2.3.1. We define convergence of sequences in these spaces. We say that

$$\begin{aligned}
f_k \to f \text{ in } \mathscr{C}^\infty \quad &\iff \quad f_k, f \in \mathscr{C}^\infty \text{ and } \lim_{k\to\infty} \sup_{|x|\le N} |\partial^\alpha (f_k - f)(x)| = 0 \\
&\qquad\quad \forall\, \alpha \text{ multi-indices and all } N = 1,2,\dots. \\
f_k \to f \text{ in } \mathscr{S} \quad &\iff \quad f_k, f \in \mathscr{S} \text{ and } \lim_{k\to\infty} \sup_{x\in\mathbf{R}^n} |x^\alpha \partial^\beta (f_k - f)(x)| = 0 \\
&\qquad\quad \forall\, \alpha,\beta \text{ multi-indices.} \\
f_k \to f \text{ in } \mathscr{C}_0^\infty \quad &\iff \quad f_k, f \in \mathscr{C}_0^\infty, \text{ support}(f_k) \subseteq B \text{ for all } k,\ B \text{ compact,} \\
&\qquad\quad \text{and } \lim_{k\to\infty} \big\|\partial^\alpha (f_k - f)\big\|_{L^\infty} = 0\ \forall\, \alpha \text{ multi-indices.}
\end{aligned}$$

It follows that convergence in $\mathscr{C}_0^\infty(\mathbf{R}^n)$ implies convergence in $\mathscr{S}(\mathbf{R}^n)$, which in turn implies convergence in $\mathscr{C}^\infty(\mathbf{R}^n)$.

Example 2.3.2. Let φ be a nonzero $\mathscr{C}_0^\infty$ function on $\mathbf{R}$. We call such functions *smooth bumps*. Define the sequence of smooth bumps $\varphi_k(x) = \varphi(x-k)/k$. Then $\varphi_k(x)$ does not converge to zero in $\mathscr{C}_0^\infty(\mathbf{R})$, even though φ_k (and all of its derivatives) converge to zero uniformly. Furthermore, we see that φ_k does not converge to any function in $\mathscr{S}(\mathbf{R})$. Clearly $\varphi_k \to 0$ in $\mathscr{C}^\infty(\mathbf{R})$.

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

$$\widetilde{\rho}_{\alpha,N}(f) = \sup_{|x|\le N} |(\partial^\alpha f)(x)|, \tag{2.3.1}$$

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

The topology of $\mathscr{C}_0^\infty$ given in Definition 2.3.1 is the *inductive limit topology,* and under this topology it is complete. Indeed, letting $\mathscr{C}_0^\infty(B(0,k))$ be the space of all smooth functions with support in $B(0,k)$, then $\mathscr{C}_0^\infty(\mathbf{R}^n)$ is equal to $\bigcup_{k=1}^\infty \mathscr{C}_0^\infty(B(0,k))$ and each space $\mathscr{C}_0^\infty(B(0,k))$ is complete with respect to the topology generated by the family of seminorms $\widetilde{\rho}_{\alpha,N}$; hence so is $\mathscr{C}_0^\infty(\mathbf{R}^n)$. Nevertheless, $\mathscr{C}_0^\infty(\mathbf{R}^n)$ is not metrizable. Details on the topologies of these spaces can be found in [286].

2.3.2 *Spaces of Functionals on Test Functions*

The dual spaces (i.e., the spaces of continuous linear functionals on the sets of test functions) we introduced is denoted by

$$\begin{aligned}
(\mathscr{C}_0^\infty(\mathbf{R}^n))' &= \mathscr{D}'(\mathbf{R}^n)\,,\\
(\mathscr{S}(\mathbf{R}^n))' &= \mathscr{S}'(\mathbf{R}^n)\,,\\
(\mathscr{C}^\infty(\mathbf{R}^n))' &= \mathscr{E}'(\mathbf{R}^n)\,.
\end{aligned}$$

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

$$\begin{aligned}
T_k \to T \quad \text{in } \mathscr{D}' &\iff \quad T_k, T \in \mathscr{D}' \text{ and } T_k(f) \to T(f) \text{ for all } f \in \mathscr{C}_0^\infty.\\
T_k \to T \quad \text{in } \mathscr{S}' &\iff \quad T_k, T \in \mathscr{S}' \text{ and } T_k(f) \to T(f) \text{ for all } f \in \mathscr{S}.\\
T_k \to T \quad \text{in } \mathscr{E}' &\iff \quad T_k, T \in \mathscr{E}' \text{ and } T_k(f) \to T(f) \text{ for all } f \in \mathscr{C}^\infty.
\end{aligned}$$

The dual spaces are nested as follows:

$$\mathscr{E}'(\mathbf{R}^n) \subseteq \mathscr{S}'(\mathbf{R}^n) \subseteq \mathscr{D}'(\mathbf{R}^n)\,.$$

Definition 2.3.3. Elements of the space $\mathscr{D}'(\mathbf{R}^n)$ are called *distributions*. Elements of $\mathscr{S}'(\mathbf{R}^n)$ are called *tempered distributions*. Elements of the space $\mathscr{E}'(\mathbf{R}^n)$ 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:

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

Proposition 2.3.4. *(a) A linear functional u on $\mathscr{C}_0^\infty(\mathbf{R}^n)$ is a distribution if and only if for every compact $K \subseteq \mathbf{R}^n$, there exist $C > 0$ and an integer m such that*

$$\big|\langle u, f \rangle\big| \le C \sum_{|\alpha| \le m} \big\|\partial^\alpha f\big\|_{L^\infty}, \quad \textit{for all } f \in \mathscr{C}^\infty \textit{ with support in } K. \tag{2.3.2}$$

(b) A linear functional u on $\mathscr{S}(\mathbf{R}^n)$ is a tempered distribution if and only if there exist $C > 0$ and k, m integers such that

$$\big|\langle u, f \rangle\big| \le C \sum_{\substack{|\alpha| \le m \\ |\beta| \le k}} \rho_{\alpha,\beta}(f), \qquad \textit{for all } f \in \mathscr{S}(\mathbf{R}^n). \tag{2.3.3}$$

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

$$\big|\langle u, f \rangle\big| \le C \sum_{|\alpha| \le m} \widetilde{\rho}_{\alpha,N}(f), \qquad \textit{for all } f \in \mathscr{C}^\infty(\mathbf{R}^n). \tag{2.3.4}$$

The seminorms $\rho_{\alpha,\beta}$ and $\widetilde{\rho}_{\alpha,N}$ are defined in (2.2.6) and (2.3.1), respectively.

Proof. We prove only (2.3.3), since the proofs of (2.3.2) and (2.3.4) are similar. It is clear that (2.3.3) implies continuity of u. Conversely, it was pointed out in Section 2.2 that the family of sets $\{f \in \mathscr{S}(\mathbf{R}^n) : \rho_{\alpha,\beta}(f) < \delta\}$, where α, β are multi-indices and $\delta > 0$, forms a subbasis for the topology of $\mathscr{S}$. Thus if u is a continuous functional on $\mathscr{S}$, there exist integers k, m and a $\delta > 0$ such that

$$|\alpha| \le m, \ |\beta| \le k, \quad \text{and } \rho_{\alpha,\beta}(f) < \delta \implies \big|\langle u, f \rangle\big| \le 1. \tag{2.3.5}$$

We see that (2.3.3) follows from (2.3.5) with $C = 1/\delta$. □

Examples 2.3.5. We now discuss some important examples.

1. The *Dirac mass* at the origin δ_0. This is defined for $\varphi \in \mathscr{C}^\infty(\mathbf{R}^n)$ by

$$\langle \delta_0, \varphi \rangle = \varphi(0).$$

We claim that δ_0 is in $\mathscr{E}'$. To see this we observe that if $\varphi_k \to \varphi$ in $\mathscr{C}^\infty$ then $\langle \delta_0, \varphi_k \rangle \to \langle \delta_0, \varphi \rangle$. The Dirac mass at a point $a \in \mathbf{R}^n$ is defined similarly by

$$\langle \delta_a, \varphi \rangle = \varphi(a).$$

2. Some functions g can be thought of as distributions via the identification $g \mapsto L_g$, where L_g is the functional

$$L_g(\varphi) = \int_{\mathbf{R}^n} \varphi(x) g(x)\, dx.$$

Here are some examples: The function 1 is in $\mathscr{S}'$ but not in $\mathscr{E}'$. Compactly supported integrable functions are in $\mathscr{E}'$. The function $e^{|x|^2}$ is in $\mathscr{D}'$ but not in $\mathscr{S}'$.
3. Functions in L^1_{loc} are distributions. To see this, first observe that if $g \in L^1_{\mathrm{loc}}$, then the integral

$$L_g(\varphi) = \int_{\mathbf{R}^n} \varphi(x) g(x)\, dx$$

is well defined for all $\varphi \in \mathscr{D}$ and satisfies $|L_g(\varphi)| \le \big(\int_K |g(x)|\, dx\big) \|\varphi\|_{L^\infty}$ for all smooth functions φ supported in the compact set K.
4. Functions in L^p, $1 \le p \le \infty$, are tempered distributions, but may not in $\mathscr{E}'$ unless they have compact support.
5. Any finite Borel measure μ is a tempered distribution via the identification

$$L_\mu(\varphi) = \int_{\mathbf{R}^n} \varphi(x)\, d\mu(x).$$

To see this, observe that $\varphi_k \to \varphi$ in $\mathscr{S}$ implies that $L_\mu(\varphi_k) \to L_\mu(\varphi)$. Finite Borel measures may not be distributions with compact support.
6. Every function g that satisfies $|g(x)| \le C(1+|x|)^k$, for some real number k, is a tempered distribution. To see this, observe that

$$\big|L_g(\varphi)\big| \le \sup_{x \in \mathbf{R}^n} (1+|x|)^m |\varphi(x)| \int_{\mathbf{R}^n} (1+|x|)^{k-m} dx,$$

where $m > n+k$ and the expression $\sup_{x\in\mathbf{R}^n}(1+|x|)^m|\varphi(x)|$ is bounded by a finite sum of Schwartz seminorms $\rho_{\alpha,\beta}(\varphi)$.
7. The function $\log|x|$ is a tempered distribution; indeed for any $\varphi \in \mathscr{S}(\mathbf{R}^n)$, the integral of $\varphi(x)\log|x|$ is bounded by a finite number of Schwartz seminorms of φ. More generally, any function that is integrable on a ball $|x| \le M$ and for some $C > 0$ satisfies $|g(x)| \le C(1+|x|)^k$ for $|x| \ge M$, is a tempered distribution.
8. Here is an example of a compactly supported distribution on $\mathbf{R}$ that is neither a locally integrable function nor a finite Borel measure:

$$\langle u, \varphi \rangle = \lim_{\varepsilon \to 0} \int_{\varepsilon \le |x| \le 1} \varphi(x) \frac{dx}{x} = \lim_{\varepsilon \to 0} \int_{\varepsilon \le |x| \le 1} (\varphi(x) - \varphi(0)) \frac{dx}{x}.$$

We have that $|\langle u, \varphi \rangle| \le 2\|\varphi'\|_{L^\infty([-1,1])}$ and notice that $\|\varphi'\|_{L^\infty([-1,1])}$ is a $\widetilde{\rho}_{\alpha,N}$ seminorm of φ.

2.3.3 The Space of Tempered Distributions

Having set down the basic definitions of distributions, we now focus our study on the space of tempered distributions. These distributions are the most useful in harmonic analysis. The main reason for this is that the subject is concerned with boundedness of translation-invariant operators, and every such bounded operator from $L^p(\mathbf{R}^n)$ to $L^q(\mathbf{R}^n)$ is given by convolution with a tempered distribution. This fact is shown in Section 2.5.

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

$$\int_{\mathbf{R}^n} (\partial^\alpha f)(x) g(x)\, dx = (-1)^{|\alpha|} \int_{\mathbf{R}^n} f(x)(\partial^\alpha g)(x)\, dx. \tag{2.3.6}$$

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 (2.3.6) for g in $\mathscr{S}'$ and $f \in \mathscr{S}$ if the integrals in (2.3.6) are interpreted as actions of distributions on functions. We simply use equation (2.3.6) to define the derivative of a distribution.

Definition 2.3.6. Let $u \in \mathscr{S}'$ and α a multi-index. Define

$$\langle \partial^\alpha u, f \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha f \rangle . \tag{2.3.7}$$

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

In view of Theorem 2.2.14, it is natural to give the following:

Definition 2.3.7. Let $u \in \mathscr{S}'$. We define the Fourier transform $\widehat{u}$ and the inverse Fourier transform $u^\vee$ of a tempered distribution u by

$$\langle \widehat{u}, f \rangle = \langle u, \widehat{f} \rangle \qquad \text{and} \qquad \langle u^\vee, f \rangle = \langle u, f^\vee \rangle, \tag{2.3.8}$$

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

Example 2.3.8. We observe that $\widehat{\delta_0} = 1$. More generally, for any multi-index α we have

$$(\partial^\alpha \delta_0)^\wedge = (2\pi i x)^\alpha .$$

To see this, observe that for all $f \in \mathscr{S}$ we have

$$\begin{aligned}
\langle (\partial^\alpha \delta_0)^\wedge, f \rangle &= \langle \partial^\alpha \delta_0, \widehat{f} \rangle \\
&= (-1)^{|\alpha|} \langle \delta_0, \partial^\alpha \widehat{f} \rangle \\
&= (-1)^{|\alpha|} \langle \delta_0, ((-2\pi i x)^\alpha f(x))^\wedge \rangle \\
&= (-1)^{|\alpha|} ((-2\pi i x)^\alpha f(x))^\wedge (0)
\end{aligned}$$

$$= (-1)^{|\alpha|} \int_{\mathbf{R}^n} (-2\pi i x)^\alpha f(x)\,dx$$
$$= \int_{\mathbf{R}^n} (2\pi i x)^\alpha f(x)\,dx.$$

This calculation indicates that $(\partial^\alpha \delta_0)^\wedge$ can be identified with the function $(2\pi i x)^\alpha$.

Example 2.3.9. Recall that for $x_0 \in \mathbf{R}^n$, $\delta_{x_0}(f) = \langle \delta_{x_0}, f \rangle = f(x_0)$. Then

$$\langle \widehat{\delta_{x_0}}, h \rangle = \langle \delta_{x_0}, \widehat{h} \rangle = \widehat{h}(x_0) = \int_{\mathbf{R}^n} h(x) e^{-2\pi i x \cdot x_0}\,dx, \qquad h \in \mathscr{S}(\mathbf{R}^n),$$

that is, $\widehat{\delta_{x_0}}$ can be identified with the function $x \mapsto e^{-2\pi i x \cdot x_0}$. In particular, $\widehat{\delta_0} = 1$.

Example 2.3.10. The function $e^{|x|^2}$ is not in $\mathscr{S}'(\mathbf{R}^n)$ and therefore its Fourier transform is not defined as a distribution. However, the Fourier transform of any locally integrable function with polynomial growth at infinity is defined as a tempered distribution.

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

$$\begin{aligned} \int_{\mathbf{R}^n} g(x) f(x-t)\,dx &= \int_{\mathbf{R}^n} g(x+t) f(x)\,dx, \\ \int_{\mathbf{R}^n} g(ax) f(x)\,dx &= \int_{\mathbf{R}^n} g(x) a^{-n} f(a^{-1}x)\,dx, \\ \int_{\mathbf{R}^n} \widetilde{g}(x) f(x)\,dx &= \int_{\mathbf{R}^n} g(x) \widetilde{f}(x)\,dx, \end{aligned} \tag{2.3.9}$$

for all $t \in \mathbf{R}^n$ and $a > 0$. Recall now the definitions of τ^t, δ^a, and $\widetilde{\ }$ given in (2.2.13). Motivated by (2.3.9), we give the following:

Definition 2.3.11. The *translation* $\tau^t u$, the *dilation* $\delta^a u$, and the *reflection* $\widetilde{u}$ of a tempered distribution u are defined as follows:

$$\langle \tau^t u, f \rangle = \langle u, \tau^{-t} f \rangle, \tag{2.3.10}$$
$$\langle \delta^a u, f \rangle = \langle u, a^{-n} \delta^{1/a} f \rangle, \tag{2.3.11}$$
$$\langle \widetilde{u}, f \rangle = \langle u, \widetilde{f} \rangle, \tag{2.3.12}$$

for all $t \in \mathbf{R}^n$ and $a > 0$. Let A be an invertible matrix. The composition of a distribution u with an invertible matrix A is the distribution

$$\langle u^A, \varphi \rangle = |\det A|^{-1} \langle u, \varphi^{A^{-1}} \rangle, \tag{2.3.13}$$

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 2.3.12. The Dirac mass at the origin δ_0 is equal to its reflection, while $\delta^a\delta_0 = a^{-n}\delta_0$. Also, $\tau^x\delta_0 = \delta_x$ for any $x \in \mathbf{R}^n$.

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

$$\int_{\mathbf{R}^n} (h*g)(x)f(x)\,dx = \int_{\mathbf{R}^n} g(x)(\widetilde{h}*f)(x)\,dx. \tag{2.3.14}$$

Motivated by (2.3.14), we define the convolution of a function with a tempered distribution as follows:

Definition 2.3.13. Let $u \in \mathscr{S}'$ and $h \in \mathscr{S}$. Define the convolution $h*u$ by

$$\langle h*u, f\rangle = \langle u, \widetilde{h}*f\rangle, \qquad f \in \mathscr{S}. \tag{2.3.15}$$

Example 2.3.14. Let $u = \delta_{x_0}$ and $f \in \mathscr{S}$. Then $f*\delta_{x_0}$ is the function $x \mapsto f(x-x_0)$, for when $h \in \mathscr{S}$, we have

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

It follows that convolution with δ_0 is the identity operator.

We now define the product of a function and a distribution.

Definition 2.3.15. Let $u \in \mathscr{S}'$ and let h be a $\mathscr{C}^\infty$ 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 $|(\partial^\alpha h)(x)| \le C_\alpha(1+|x|)^{k_\alpha}$ for some $C_\alpha, k_\alpha > 0$. Then define the product hu of h and u by

$$\langle hu, f\rangle = \langle u, hf\rangle, \qquad f \in \mathscr{S}. \tag{2.3.16}$$

Note that hf is in $\mathscr{S}$ and thus (2.3.16) is well defined. The product of an arbitrary $\mathscr{C}^\infty$ function with a tempered distribution is not defined.

We observe that if a function g is supported in a set K, then for all $f \in \mathscr{C}_0^\infty(K^c)$ we have

$$\int_{\mathbf{R}^n} f(x)g(x)\,dx = 0. \tag{2.3.17}$$

Moreover, the support of g is the intersection of all closed sets K with the property (2.3.17) for all f in $\mathscr{C}_0^\infty(K^c)$. Motivated by the preceding observation we give the following:

Definition 2.3.16. Let u be in $\mathscr{D}'(\mathbf{R}^n)$. The *support* of u ($\operatorname{supp} u$) is the intersection of all closed sets K with the property

$$\varphi \in \mathscr{C}_0^\infty(\mathbf{R}^n), \qquad \operatorname{supp}\varphi \subseteq \mathbf{R}^n \setminus K \implies \langle u, \varphi\rangle = 0. \tag{2.3.18}$$

Distributions with compact support are exactly those whose support (as defined in the previous definition) is a compact set. To prove this assertion, we start with a distribution u with compact support as defined in Definition 2.3.3. Then there exist $C,N,m>0$ such that (2.3.4) holds. For a $\mathscr{C}^\infty$ function f whose support is contained in $B(0,N)^c$, the expression on the right in (2.3.4) 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 2.3.16 is a compact set, then there exists an $N>0$ such that $\operatorname{supp} u$ is contained in $B(0,N)$. We take a smooth function η that is equal to 1 on $B(0,N)$ and vanishes off $B(0,N+1)$. Then for $h \in \mathscr{C}_0^\infty$ the support of $h(1-\eta)$ does not meet the support of u, and we must have

$$\big\langle u,h\big\rangle = \big\langle u,h\eta\big\rangle + \big\langle u,h(1-\eta)\big\rangle = \big\langle u,h\eta\big\rangle .$$

The distribution u can be thought of as an element of $\mathscr{E}'$ by defining for $f \in \mathscr{C}^\infty(\mathbf{R}^n)$

$$\big\langle u,f\big\rangle = \big\langle u,f\eta\big\rangle .$$

Taking m to be the integer that corresponds to the compact set $K = \overline{B(0,N+1)}$ in (2.3.2), and using that the L^∞ norm of $\partial^\alpha(f\eta)$ is controlled by a finite sum of seminorms $\widetilde{\rho}_{\alpha,N+1}(f)$ with $|\alpha| \le m$, we obtain the validity of (2.3.4) for $f \in \mathscr{C}^\infty$.

Example 2.3.17. The support of the Dirac mass at x_0 is the set $\{x_0\}$.

Along the same lines, we give the following definition:

Definition 2.3.18. We say that a distribution u in $\mathscr{D}'(\mathbf{R}^n)$ coincides with the function h on an open set Ω if

$$\big\langle u,f\big\rangle = \int_{\mathbf{R}^n} f(x)h(x)\,dx \qquad \text{for all } f \text{ in } \mathscr{C}_0^\infty(\Omega). \tag{2.3.19}$$

When (2.3.19) occurs we often say that u agrees with h away from Ω^c.

This definition implies that the support of the distribution $u-h$ is contained in the set Ω^c.

Example 2.3.19. The distribution $|x|^2 + \delta_{a_1} + \delta_{a_2}$, where a_1, a_2 are in $\mathbf{R}^n$, coincides with the function $|x|^2$ on any open set not containing the points a_1 and a_2. Also, the distribution in Example 2.3.5 (8) coincides with the function $x^{-1}\chi_{|x|\le 1}$ away from the origin in the real line.

Having ended the streak of definitions regarding operations with distributions, we now discuss properties of convolutions and Fourier transforms.

Theorem 2.3.20. *If $u \in \mathscr{S}'$ and $\varphi \in \mathscr{S}$, then $\varphi * u$ is a $\mathscr{C}^\infty$ function and*

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

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

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

Furthermore, if u has compact support, then $\varphi * u$ is a Schwartz function.

Proof. Let ψ be in $\mathscr{S}(\mathbf{R}^n)$. We have

$$\begin{aligned}
\langle \varphi * u, \psi \rangle &= \langle u, \widetilde{\varphi} * \psi \rangle \\
&= u\left(\int_{\mathbf{R}^n} \widetilde{\varphi}(\cdot - y)\psi(y)\,dy \right) \\
&= u\left(\int_{\mathbf{R}^n} (\tau^y \widetilde{\varphi})(\cdot)\psi(y)\,dy \right) \qquad (2.3.20)\\
&= \int_{\mathbf{R}^n} \langle u, \tau^y \widetilde{\varphi} \rangle \psi(y)\,dy,
\end{aligned}$$

where the last step is justified by the continuity of u and by the fact that the Riemann sums of the inner integral in (2.3.20) converge to that integral in the topology of $\mathscr{S}$, a fact that will be justified later. This calculation identifies the function $\varphi * u$ as

$$(\varphi * u)(x) = \langle u, \tau^x \widetilde{\varphi} \rangle. \qquad (2.3.21)$$

We now show that $(\varphi * u)(x)$ is a $\mathscr{C}^\infty$ function. Let $e_j = (0,\dots,1,\dots,0)$ with 1 in the jth entry and zero elsewhere. Then

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

by the continuity of u and the fact that $\big(\tau^{-he_j}(\tau^x \widetilde{\varphi}) - \tau^x \widetilde{\varphi}\big)/h$ tends to $\partial_j \tau^x \widetilde{\varphi} = \tau^x(\partial_j \widetilde{\varphi})$ in $\mathscr{S}$ as $h \to 0$; see Exercise 2.3.5 (a). The same calculation for higher-order derivatives shows that $\varphi * u \in \mathscr{C}^\infty$ and that $\partial^\gamma(\varphi * u) = (\partial^\gamma \varphi) * u$ for all multi-indices γ. It follows from (2.3.3) that for some C, m, and k we have

$$\begin{aligned}
|\partial^\alpha(\varphi * u)(x)| &\le C \sum_{\substack{|\gamma|\le m\\ |\beta|\le k}} \sup_{y\in\mathbf{R}^n} |y^\gamma \tau^x(\partial^{\alpha+\beta}\widetilde{\varphi})(y)| \\
&= C \sum_{\substack{|\gamma|\le m\\ |\beta|\le k}} \sup_{y\in\mathbf{R}^n} |(x+y)^\gamma (\partial^{\alpha+\beta}\widetilde{\varphi})(y)| \qquad (2.3.22)\\
&\le C_m \sum_{|\beta|\le k} \sup_{y\in\mathbf{R}^n} (1+|x|^m+|y|^m)|(\partial^{\alpha+\beta}\widetilde{\varphi})(y)|,
\end{aligned}$$

and this clearly implies that $\partial^\alpha(\varphi * u)$ grows at most polynomially at infinity.

We now indicate why $\varphi * u$ is Schwartz whenever u has compact support. Applying estimate (2.3.4) to the function $y \mapsto \varphi(x-y)$ yields that

$$\big|\big\langle u, \varphi(x-\cdot)\big\rangle\big| = |(\varphi * u)(x)| \le C \sum_{|\alpha|\le m} \sup_{|y|\le N} |\partial_y^\alpha \varphi(x-y)|$$

for some constants C, m, N. Since for $|x| \ge 2N$ we have

$$|\partial_y^\alpha \varphi(x-y)| \le C_{\alpha,M}(1+|x-y|)^{-M} \le C_{\alpha,M,N}(1+|x|)^{-M},$$

it follows that $\varphi * u$ decays rapidly at infinity. Since $\partial^\gamma(\varphi * u) = (\partial^\gamma \varphi) * u$, the same argument yields that all the derivatives of $\varphi * u$ decay rapidly at infinity; hence $\varphi * u$ is a Schwartz function. Incidentally, this argument actually shows that any Schwartz seminorm of $\varphi * u$ is controlled by a finite sum of Schwartz seminorms of φ.

We now return to the point left open concerning the convergence of the Riemann sums in (2.3.20) in the topology of $\mathscr{S}(\mathbf{R}^n)$. For each $N = 1, 2, \dots$, 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

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

converges to zero in $L^\infty(\mathbf{R}^n)$ as $N \to \infty$. We have

$$\begin{aligned} x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y_m)\psi(y_m)|Q_m| - \int_{Q_m} x^\alpha \partial_x^\beta \widetilde{\varphi}(x-y)\psi(y)\,dy \\ = \int_{Q_m} x^\alpha (y-y_m)\cdot \nabla\big(\partial_x^\beta \widetilde{\varphi}(x-\cdot)\psi\big)(\xi)\,dy \end{aligned}$$

for some $\xi = y + \theta(y_m - y)$, where $\theta \in [0,1]$. Distributing the gradient to both factors, we see that the last integrand is at most

$$C\,|x|^{|\alpha|}\,\frac{\sqrt{n}}{N}\,\frac{1}{(1+|x-\xi|)^{M/2}}\,\frac{1}{(2+|\xi|)^M}$$

for M large (pick $M > 2|\alpha|$), which in turn is at most

$$C'\,|x|^{|\alpha|}\,\frac{\sqrt{n}}{N}\,\frac{1}{(1+|x|)^{M/2}}\,\frac{1}{(2+|\xi|)^{M/2}} \le C'\,|x|^{|\alpha|}\,\frac{\sqrt{n}}{N}\,\frac{1}{(1+|x|)^{M/2}}\,\frac{1}{(1+|y|)^{M/2}},$$

since $|y| \le |\xi| + \theta|y-y_m| \le |\xi| + \sqrt{n}/N \le |\xi|+1$ for $N \ge \sqrt{n}$. Inserting the estimate obtained for the integrand in the last displayed integral, we obtain

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

But the second integral in the preceding expression is bounded by

$$\int_{([-N,N]^n)^c} \frac{C'''|x|^{|\alpha|}}{(1+|x-y|)^M} \frac{dy}{(1+|y|)^M} \leq \frac{C'''|x|^{|\alpha|}}{(1+|x|)^{M/2}} \int_{([-N,N]^n)^c} \frac{dy}{(1+|y|)^{M/2}}.$$

Using these estimates it is now easy to see that $\lim_{N\to\infty} \sup_{x\in\mathbf{R}^n} |D_N(x)| = 0$. □

Next we have the following important result regarding distributions with compact support:

Theorem 2.3.21. *If u is in $\mathscr{E}'(\mathbf{R}^n)$, then $\widehat{u}$ is a real analytic function on $\mathbf{R}^n$. In particular, $\widehat{u}$ is a $\mathscr{C}^\infty$ function. Furthermore, $\widehat{u}$ and all of its derivatives have polynomial growth at infinity. Moreover, $\widehat{u}$ has a holomorphic extension on $\mathbf{C}^n$.*

Proof. Given a distribution u with compact support and a polynomial $p(\xi)$, the action of u on the $\mathscr{C}^\infty$ function $\xi \mapsto p(\xi)e^{-2\pi i x\cdot\xi}$ is a well defined function of x, which we denote by $u(p(\cdot)e^{-2\pi i x\cdot(\cdot)})$. Here x is an element of $\mathbf{R}^n$ but the same assertion is valid if $x = (x_1,\dots,x_n) \in \mathbf{R}^n$ is replaced by $z = (z_1,\dots,z_n) \in \mathbf{C}^n$. In this case we define the dot product of ξ and z via $\xi\cdot z = \sum_{k=1}^n \xi_k z_k$.

It is straightforward to verify that the function of $z = (z_1,\dots,z_n)$

$$F(z) = u\big(e^{-2\pi i(\cdot)\cdot z}\big)$$

defined on $\mathbf{C}^n$ is holomorphic, in fact entire. Indeed, the continuity and linearity of u and the fact that $(e^{-2\pi i \xi_j h} - 1)/h \to -2\pi i \xi_j$ in $\mathscr{C}^\infty(\mathbf{R}^n)$ as $h \to 0$, $h \in \mathbf{C}$, 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 $\mathscr{C}^\infty$ function

$$\xi \mapsto (-2\pi i \xi_j)e^{-2\pi i \sum_{j=1}^n \xi_j z_j}.$$

By induction it follows that for all multi-indices α we have

$$\partial_{z_1}^{\alpha_1}\cdots\partial_{z_n}^{\alpha_n} F = u\big((-2\pi i(\cdot))^\alpha e^{-2\pi i \sum_{j=1}^n (\cdot) z_j}\big).$$

Since F is entire, its restriction on $\mathbf{R}^n$, i.e., $F(x_1,\dots,x_n)$, where $x_j = \operatorname{Re} z_j$, is real analytic. Also, an easy calculation using (2.3.4) and Leibniz's rule yield that the restriction of F on $\mathbf{R}^n$ and all of its derivatives have polynomial growth at infinity.

Now for f in $\mathscr{S}(\mathbf{R}^n)$ we have

$$\langle \widehat{u}, f\rangle = \langle u, \widehat{f}\rangle = u\left(\int_{\mathbf{R}^n} f(x)e^{-2\pi i x\cdot\xi}\,dx\right) = \int_{\mathbf{R}^n} f(x)u(e^{-2\pi i x\cdot(\cdot)})\,dx,$$

provided we can justify the passage of u inside the integral. The reason for this is that the Riemann sums of the integral of $f(x)e^{-2\pi i x\cdot\xi}$ over $\mathbf{R}^n$ converge to it in the topology of $\mathscr{C}^\infty$, and thus the linear functional u can be interchanged with the integral. We conclude that the tempered distribution $\widehat{u}$ can be identified with the real analytic function $x \mapsto F(x)$ whose derivatives have polynomial growth at infinity.

To justify the fact concerning the convergence of the Riemann sums, we argue as in the proof of the previous theorem. For each $N = 1,2,\dots$, 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 a multi-index α let

$$D_N(\xi) = \sum_{m=1}^{(2N^2)^n} f(y_m)(-2\pi i y_m)^\alpha e^{-2\pi i y_m\cdot\xi}|Q_m| - \int_{\mathbf{R}^n} f(x)(-2\pi i x)^\alpha e^{-2\pi i x\cdot\xi}\,dx.$$

We must show that for every $M>0$, $\sup_{|\xi|\le M}|D_N(\xi)|$ converges to zero as $N\to\infty$. Setting $g(x) = f(x)(-2\pi i x)^\alpha$, we write

$$D_N(\xi) = \sum_{m=1}^{(2N^2)^n}\int_{Q_m}\left[g(y_m)e^{-2\pi i y_m\cdot\xi} - g(x)e^{-2\pi i x\cdot\xi}\right]dx + \int_{([-N,N]^n)^c} g(x)e^{-2\pi i x\cdot\xi}\,dx.$$

Using the mean value theorem, we bound the absolute value of the expression inside the square brackets by

$$\big(|\nabla g(z_m)| + 2\pi|\xi|\,|g(z_m)|\big)\frac{\sqrt{n}}{N} \le \frac{C_K\,(1+|\xi|)}{(1+|z_m|)^K}\,\frac{\sqrt{n}}{N},$$

for some point z_m in the cube Q_m. Since

$$\sum_{m=1}^{(2N^2)^n}\int_{Q_m}\frac{C_K\,(1+|\xi|)}{(1+|z_m|)^K}\,dx \le C_K'(1+M) < \infty$$

for $|\xi|\le M$, it follows that $\sup_{|\xi|\le M}|D_N(\xi)|\to 0$ as $N\to\infty$. □

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

Proposition 2.3.22. *Given u, v in $\mathscr{S}'(\mathbf{R}^n)$, $f_j, f\in\mathscr{S}$, $y\in\mathbf{R}^n$, b a complex scalar, α a multi-index, and $a>0$, we have*

(1) $\widehat{u+v} = \widehat{u}+\widehat{v}$,

(2) $\widehat{bu} = b\widehat{u}$,

(3) *If $f_j\to f$ in $\mathscr{S}$, then $\widehat{f_j}\to\widehat{f}$ in $\mathscr{S}$ and if $u_j\to u$ in $\mathscr{S}'$, then $\widehat{u}_j\to\widehat{u}$ in $\mathscr{S}'$,*

(4) $(\widetilde{u})\widehat{\ } = (\widehat{u})\widetilde{\ }$,

(5) $(\tau^y u)\widehat{\ } = e^{-2\pi i y\cdot\xi}\widehat{u}$,

(6) $(e^{2\pi i x\cdot y}u)\widehat{\ } = \tau^y\widehat{u}$,

(7) $(\delta^a u)\widehat{\ } = (\widehat{u})_a = a^{-n}\delta^{a^{-1}}\widehat{u}$,

(8) $(\partial^\alpha u)\widehat{\ } = (2\pi i\xi)^\alpha\widehat{u}$,

(9) $\partial^\alpha\widehat{u} = ((-2\pi i x)^\alpha u)\widehat{\ }$,

(10) $(\widehat{u})^\vee = u$,

(11) $\widehat{f * u} = \widehat{f}\,\widehat{u}$,

(12) $\widehat{fu} = \widehat{f} * \widehat{u}$,

(13) ***(Leibniz's rule)*** $\partial_j^m(fu) = \sum_{k=0}^m \binom{m}{k}(\partial_j^k f)(\partial_j^{m-k} u)$, $m \in \mathbf{Z}^+$,

(14) ***(Leibniz's rule)*** $\partial^\alpha(fu) = \sum_{\gamma_1=0}^{\alpha_1} \cdots \sum_{\gamma_n=0}^{\alpha_n} \binom{\alpha_1}{\gamma_1} \cdots \binom{\alpha_n}{\gamma_n}(\partial^\gamma f)(\partial^{\alpha-\gamma} u)$,

(15) *If u_k, $u \in L^p(\mathbf{R}^n)$ and $u_k \to u$ in L^p $(1 \le p \le \infty)$, then $u_k \to u$ in $\mathscr{S}'(\mathbf{R}^n)$. Therefore, convergence in $\mathscr{S}$ implies convergence in L^p, which in turn implies convergence in $\mathscr{S}'(\mathbf{R}^n)$.*

Proof. All the statements can be proved easily using duality and the corresponding statements for Schwartz functions. □

We continue with an application of Theorem 2.3.21.

Proposition 2.3.23. *Given $u \in \mathscr{S}'(\mathbf{R}^n)$, there exists a sequence of $\mathscr{C}_0^\infty$ functions f_k such that $f_k \to u$ in the sense of tempered distributions; in particular, $\mathscr{C}_0^\infty(\mathbf{R}^n)$ is dense in $\mathscr{S}'(\mathbf{R}^n)$.*

Proof. Fix a function in $\mathscr{C}_0^\infty(\mathbf{R}^n)$ with $\varphi(x) = 1$ in a neighborhood of the origin. Let $\varphi_k(x) = \delta^{1/k}(\varphi)(x) = \varphi(x/k)$. It follows from Exercise 2.3.5 (b) that for $u \in \mathscr{S}'(\mathbf{R}^n)$, $\varphi_k u \to u$ in $\mathscr{S}'$. By Proposition 2.3.22 (3), we have that the map $u \mapsto (\varphi_k \widehat{u})^\vee$ is continuous on $\mathscr{S}'(\mathbf{R}^n)$. Now Theorem 2.3.21 gives that $(\varphi_k \widehat{u})^\vee$ is a $\mathscr{C}^\infty$ function and therefore $\varphi_j(\varphi_k \widehat{u})^\vee$ is in $\mathscr{C}_0^\infty(\mathbf{R}^n)$. As observed, $\varphi_j(\varphi_k \widehat{u})^\vee \to (\varphi_k \widehat{u})^\vee$ in $\mathscr{S}'$ when k is fixed and $j \to \infty$. Exercise 2.3.5 (c) gives that the diagonal sequence $\varphi_k(\varphi_k f)^\wedge$ converges to $\widehat{f}$ in $\mathscr{S}$ as $k \to \infty$ for all $f \in \mathscr{S}$. Using duality and Exercise 2.2.2, we conclude that the sequence of $\mathscr{C}_0^\infty$ functions $\varphi_k(\varphi_k \widehat{u})^\vee$ converges to u in $\mathscr{S}'$ as $k \to \infty$. □

Exercises

2.3.1. Show that a positive measure μ that satisfies

$$\int_{\mathbf{R}^n} \frac{d\mu(x)}{(1+|x|)^k} < +\infty,$$

for some $k > 0$, can be identified with a tempered distribution. Show that if we think of Lebesgue measure as a tempered distribution, then it coincides with the constant function 1 also interpreted as a tempered distribution.

2.3.2. Let $\varphi, f \in \mathscr{S}(\mathbf{R}^n)$, and for $\varepsilon > 0$ let $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(\varepsilon^{-1}x)$. Prove that $\varphi_\varepsilon * f \to bf$ in $\mathscr{S}$, where b is the integral of φ.

2.3.3. Prove that for all $a>0$, $u\in\mathscr{S}'(\mathbf{R}^n)$, and $f\in\mathscr{S}(\mathbf{R}^n)$ we have

$$(\delta^a f)*(\delta^a u)=a^{-n}\delta^a(f*u)\,.$$

2.3.4. (a) Prove that the derivative of $\chi_{[a,b]}$ is $\delta_a-\delta_b$.
(b) Compute $\partial_j\chi_{B(0,1)}$ on $\mathbf{R}^2$.
(c) Compute the Fourier transforms of the locally integrable functions $\sin x$ and $\cos x$.
(d) Prove that the derivative of the distribution $\log|x|\in\mathscr{S}'(\mathbf{R})$ is the distribution

$$u(\varphi)=\lim_{\varepsilon\to 0}\int_{\varepsilon\le|x|}\varphi(x)\frac{dx}{x}\,.$$

2.3.5. Let $f\in\mathscr{S}(\mathbf{R}^n)$ and let $\varphi\in\mathscr{C}_0^\infty$ be identically equal to 1 in a neighborhood of the origin. Define $\varphi_k(x)=\varphi(x/k)$ as in the proof of Proposition 2.3.23.
(a) Prove that $(\tau^{-he_j}f-f)/h\to\partial_j f$ in $\mathscr{S}$ as $h\to 0$.
(b) Prove that $\varphi_k f\to f$ in $\mathscr{S}$ as $k\to\infty$.
(c) Prove that the sequence $\varphi_k(\varphi_k f)^\wedge$ converges to $\widehat{f}$ in $\mathscr{S}$ as $k\to\infty$.

2.3.6. Use Theorem 2.3.21 to show that there does not exist a nonzero $\mathscr{C}_0^\infty$ function whose Fourier transform is also a $\mathscr{C}_0^\infty$ function.

2.3.7. Let $f\in L^p(\mathbf{R}^n)$ for some $1\le p<\infty$. Show that the sequence of functions

$$g_N(\xi)=\int_{B(0,N)}f(x)e^{-2\pi i x\cdot\xi}\,dx$$

converges to $\widehat{f}$ in $\mathscr{S}'$.

2.3.8. Let $(c_k)_{k\in\mathbf{Z}^n}$ be a sequence that satisfies $|c_k|\le A(1+|k|)^M$ for all k and some fixed M and $A>0$. Let δ_k denote Dirac mass at the integer k. Show that the sequence of distributions

$$\sum_{|k|\le N}c_k\delta_k$$

converges to some tempered distribution u in $\mathscr{S}'(\mathbf{R}^n)$ as $N\to\infty$. Also show that $\widehat{u}$ is the $\mathscr{S}'$ limit of the sequence of functions

$$h_N(\xi)=\sum_{|k|\le N}c_k e^{-2\pi i\xi\cdot k}\,.$$

2.3.9. A distribution in $\mathscr{S}'(\mathbf{R}^n)$ is called *homogeneous of degree* $\gamma\in\mathbf{C}$ if for all $\lambda>0$ and for all $\varphi\in\mathscr{S}(\mathbf{R}^n)$ we have

$$\langle u,\delta^\lambda\varphi\rangle=\lambda^{-n-\gamma}\langle u,\varphi\rangle\,.$$

(a) Prove that this definition agrees with the usual definition for functions.
(b) Show that δ_0 is homogeneous of degree $-n$.
(c) Prove that if u is homogeneous of degree γ, then $\partial^\alpha u$ is homogeneous of degree $\gamma-|\alpha|$.

(d) Show that u is homogeneous of degree γ if and only if $\widehat{u}$ is homogeneous of degree $-n-\gamma$.

2.3.10. (a) Show that the functions e^{inx} and e^{-inx} converge to zero in $\mathscr{S}'$ and $\mathscr{D}'$ as $n \to \infty$. Conclude that multiplication of distributions is not a continuous operation even when it is defined.
(b) What is the limit of $\sqrt{n}(1+n|x|^2)^{-1}$ in $\mathscr{D}'(\mathbf{R})$ as $n \to \infty$?

2.3.11. (*S. Bernstein*) Let f be a bounded function on $\mathbf{R}^n$ with $\widehat{f}$ supported in the ball $B(0,R)$. Prove that for all multi-indices α there exist constants $C_{\alpha,n}$ (depending only on α and on the dimension n) such that

$$\left\|\partial^\alpha f\right\|_{L^\infty} \le C_{\alpha,n} R^{|\alpha|} \left\|f\right\|_{L^\infty}.$$

$\big[$*Hint:* Write $f = f * h_{1/R}$, where h is a Schwartz function h in $\mathbf{R}^n$ whose Fourier transform is equal to one on the ball $B(0,1)$ and vanishes outside the ball $B(0,2)$.$\big]$

2.3.12. Let $\widehat{\Phi}$ be a $\mathscr{C}_0^\infty$ function that is equal to 1 in $B(0,1)$ and let $\widehat{\Theta}$ be a $\mathscr{C}^\infty$ function that is equal to 1 in a neighborhood of infinity and equal to zero in a neighborhood of the origin. Prove the following.
(a) For all u in $\mathscr{S}'(\mathbf{R}^n)$ we have

$$\Big(\widehat{\Phi}(\xi/2^N)\widehat{u}\Big)^\vee \to u \quad \text{in } \mathscr{S}'(\mathbf{R}^n) \text{ as } N \to \infty.$$

(b) For all u in $\mathscr{S}'(\mathbf{R}^n)$ we have

$$\Big(\widehat{\Theta}(\xi/2^N)\widehat{u}\Big)^\vee \to 0 \quad \text{in } \mathscr{S}'(\mathbf{R}^n) \text{ as } N \to \infty.$$

2.3.13. Prove that there exists a function in L^p for $2 < p < \infty$ whose distributional Fourier transform is not a locally integrable function.
$\big[$*Hint:* Assume the converse. Then for all $f \in L^p(\mathbf{R}^n)$, $\widehat{f}$ is locally integrable and hence the map $f \mapsto \widehat{f}$ is a well defined linear operator from $L^p(\mathbf{R}^n)$ to $L^1(B(0,M))$ for all $M > 0$ (i.e. $\|\widehat{f}\|_{L^1(B(0,M))} < \infty$ for all $f \in L^p(\mathbf{R}^n)$). Use the closed graph theorem to deduce that $\|\widehat{f}\|_{L^1(B(0,M))} \le C_M \|f\|_{L^p(\mathbf{R}^n)}$ for some $C_M < \infty$. To violate this inequality whenever $p > 2$, take $f_N(x) = (1+iN)^{-n/2} e^{-\pi(1+iN)^{-1}|x|^2}$ and let $N \to \infty$, noting that $\widehat{f_N}(\xi) = e^{-\pi|\xi|^2(1+iN)}$.$\big]$

2.4 More About Distributions and the Fourier Transform

In this section we discuss further properties of distributions and Fourier transforms and bring up certain connections that arise between harmonic analysis and partial differential equations.

2.4.1 Distributions Supported at a Point

We begin with the following characterization of distributions supported at a single point.

Proposition 2.4.1. *If $u \in \mathscr{S}'(\mathbf{R}^n)$ is supported in the singleton $\{x_0\}$, then there exists an integer k and complex numbers a_α such that*

$$u = \sum_{|\alpha| \le k} a_\alpha \partial^\alpha \delta_{x_0}.$$

Proof. Without loss of generality we may assume that $x_0 = 0$. By (2.3.3) we have that for some C, m, and k,

$$\big|\big\langle u, f\big\rangle\big| \le C \sum_{\substack{|\alpha|\le m \\ |\beta|\le k}} \sup_{x\in\mathbf{R}^n} |x^\alpha(\partial^\beta f)(x)| \qquad \text{for all } f \in \mathscr{S}(\mathbf{R}^n).$$

We now prove that if $\varphi \in \mathscr{S}$ satisfies

$$(\partial^\alpha \varphi)(0) = 0 \qquad \text{for all } |\alpha| \le k, \tag{2.4.1}$$

then $\big\langle u, \varphi\big\rangle = 0$. To see this, fix a φ satisfying (2.4.1) and let $\zeta(x)$ be a smooth function on $\mathbf{R}^n$ that is equal to 1 when $|x| \ge 2$ and equal to zero for $|x| \le 1$. Let $\zeta^\varepsilon(x) = \zeta(x/\varepsilon)$. Then, using (2.4.1) 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 $|\alpha| \le m$ and $|\beta| \le k$. Then

$$\big|\big\langle u, \varphi\big\rangle\big| \le \big|\big\langle u, \zeta^\varepsilon\varphi\big\rangle\big| + \big|\big\langle u, \zeta^\varepsilon\varphi - \varphi\big\rangle\big| \le 0 + C \sum_{\substack{|\alpha|\le m \\ |\beta|\le k}} \rho_{\alpha,\beta}(\zeta^\varepsilon\varphi - \varphi) \to 0$$

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

Now let $f \in \mathscr{S}(\mathbf{R}^n)$. Let η be a $\mathscr{C}_0^\infty$ function on $\mathbf{R}^n$ that is equal to 1 in a neighborhood of the origin. Write

$$f(x) = \eta(x)\Big(\sum_{|\alpha|\le k} \frac{(\partial^\alpha f)(0)}{\alpha!} x^\alpha + h(x)\Big) + (1-\eta(x))f(x), \tag{2.4.2}$$

where $h(x) = O(x^{k+1})$ as $|x| \to 0$. Then ηh satisfies (2.4.1) and hence $\big\langle u, \eta h\big\rangle = 0$ by the claim. Also,

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

by our hypothesis. Applying u to both sides of (2.4.2), we obtain

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

with $a_\alpha = (-1)^{|\alpha|} u(x^\alpha \eta(x))/\alpha!$. This proves the proposition. □

An immediate consequence is the following result.

Corollary 2.4.2. *Let $u \in \mathscr{S}'(\mathbf{R}^n)$. If $\widehat{u}$ is supported in the singleton $\{\xi_0\}$, then u is a finite linear combination of functions $(-2\pi i\xi)^\alpha e^{2\pi i\xi\cdot\xi_0}$, where α is a multi-index. In particular, if $\widehat{u}$ is supported at the origin, then u is a polynomial.*

Proof. Proposition 2.4.1 gives that $\widehat{u}$ is a linear combination of derivatives of Dirac masses at ξ_0. Then Proposition 2.3.22 (8) yields the required conclusion. □

2.4.2 The Laplacian

The *Laplacian* Δ is a partial differential operator acting on tempered distributions on $\mathbf{R}^n$ as follows:

$$\Delta(u) = \sum_{j=1}^{n} \partial_j^2 u .$$

Solutions of Laplace's equation $\Delta(u) = 0$ are called *harmonic* distributions. We have the following:

Corollary 2.4.3. *Let $u \in \mathscr{S}'(\mathbf{R}^n)$ satisfy $\Delta(u) = 0$. Then u is a polynomial.*

Proof. Taking Fourier transforms, we obtain that $\widehat{\Delta(u)} = 0$. Therefore,

$$-4\pi^2|\xi|^2\widehat{u} = 0 \qquad \text{in } \mathscr{S}'.$$

This implies that $\widehat{u}$ is supported at the origin, and by Corollary 2.4.2 it follows that u must be polynomial. □

Liouville's classical theorem that every bounded harmonic function must be constant is a consequence of Corollary 2.4.3. See Exercise 2.4.2.

Next we would like to compute the fundamental solutions of Laplace's equation in $\mathbf{R}^n$. A distribution is called a *fundamental solution* of a partial differential operator L if we have $L(u) = \delta_0$. The following result gives the fundamental solution of the Laplacian.

Proposition 2.4.4. *For $n \geq 3$ we have*

$$\Delta(|x|^{2-n}) = -(n-2)\frac{2\pi^{n/2}}{\Gamma(n/2)}\delta_0 , \tag{2.4.3}$$

while for $n = 2$,

$$\Delta(\log|x|) = 2\pi\delta_0 . \tag{2.4.4}$$

Proof. We use Green's identity

$$\int_\Omega \big(v\Delta(u) - u\Delta(v)\big)\,dx = \int_{\partial\Omega} \left(v\frac{\partial u}{\partial \nu} - u\frac{\partial v}{\partial \nu}\right) ds ,$$

where Ω is an open set in $\mathbf{R}^n$ with smooth boundary and $\partial v/\partial \nu$ denotes the derivative of v with respect to the outer unit normal vector. Take $\Omega = \mathbf{R}^n \setminus \overline{B(0,\varepsilon)}$, $v = |x|^{2-n}$, and $u = f$ a $\mathscr{C}_0^\infty(\mathbf{R}^n)$ function in the previous identity. The normal derivative of $f(r\theta)$ is the derivative with respect to the radial variable r. Observe that $\Delta(|x|^{2-n}) = 0$ for $x \neq 0$. We obtain

$$\int_{|x|>\varepsilon} \Delta(f)(x)|x|^{2-n}\,dx = -\int\limits_{|\theta|=\varepsilon} \left(\varepsilon^{2-n}\frac{\partial f}{\partial r} - f(r\theta)\frac{\partial r^{2-n}}{\partial r}\right)d\theta, \tag{2.4.5}$$

where $d\theta$ denotes surface measure on the sphere $|\theta| = \varepsilon$. Now observe two things: first, that for some $C = C(f)$ we have

$$\left|\int_{|\theta|=\varepsilon} \frac{\partial f}{\partial r}\,d\theta\right| \le C\varepsilon^{n-1};$$

second, that

$$\int_{|\theta|=\varepsilon} f(r\theta)\varepsilon^{1-n}\,d\theta \to \omega_{n-1} f(0)$$

as $\varepsilon \to 0$. Letting $\varepsilon \to 0$ in (2.4.5), we obtain that

$$\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} \Delta(f)(x)|x|^{2-n}\,dx = -(n-2)\omega_{n-1} f(0),$$

which implies (2.4.3) in view of the formula for ω_{n-1} given in Appendix A.3.

The proof of (2.4.4) is identical. The only difference is that the quantity $\partial r^{2-n}/\partial r$ in (2.4.5) is replaced by $\partial \log r/\partial r$. $\square$

2.4.3 Homogeneous Distributions

The fundamental solutions of the Laplacian are locally integrable functions on $\mathbf{R}^n$ and also homogeneous of degree $2-n$ when $n \ge 3$. Since homogeneous distributions often arise in applications, it is desirable to pursue their study. Here we do not undertake such a study in depth, but we discuss a few important examples.

Our first goal is to understand the action of the distribution $|t|^z$ on $\mathbf{R}^n$ when $\operatorname{Re} z \le -n$. Let us consider first the case $n = 1$. The tempered distribution

$$\langle w_z, \varphi\rangle = \int_{-1}^{1} |t|^z \varphi(t)\,dt$$

is well-defined when $\operatorname{Re} z > -1$. But we can extend the definition for all z with $\operatorname{Re} z > -3$ and $z \neq -1$ by rewriting it as

$$\langle w_z, \varphi\rangle = \int_{-1}^{1} |t|^z \big(\varphi(t) - \varphi(0) - t\varphi'(0)\big)\,dt + \frac{2}{z+1}\varphi(0), \tag{2.4.6}$$

and noting that for all $\varphi \in \mathscr{S}(\mathbf{R})$ we have

$$\big|\langle w_z, \varphi\rangle\big| \le \frac{1}{z+3}\|\varphi''\|_{L^\infty} + \frac{2}{z+1}\|\varphi\|_{L^\infty},$$

thus $w_z \in \mathscr{S}'(\mathbf{R})$. Subtracting the Taylor polynomial of degree 3 centered at zero from $\varphi(t)$ instead of the linear one, as in (2.4.6), allows us to extend the definition for $\mathrm{Re}\, z > -5$ and $\mathrm{Re}\, z \notin \{-1, -3\}$. Subtracting higher order Taylor polynomials allows us to extend the definition of w_z for all $z \in \mathbf{C}$ except at the negative odd integers. To be able to include the points $z = -1, -3, -5, -7, \dots$ we need to multiply w_z by an entire function that has simple zeros at all the negative odd integers to be able to eliminate the simple poles at these points. Such a function is $\Gamma\big(\frac{z+1}{2}\big)^{-1}$. This discussion leads to the following definition.

Definition 2.4.5. For $z \in \mathbf{C}$ we define a distribution u_z as follows:

$$\big\langle u_z, f\big\rangle = \int_{\mathbf{R}^n} \frac{\pi^{\frac{z+n}{2}}}{\Gamma\big(\frac{z+n}{2}\big)} |x|^z f(x)\, dx. \tag{2.4.7}$$

Clearly the u_z's coincide with the locally integrable functions

$$\pi^{\frac{z+n}{2}} \Gamma\big(\tfrac{z+n}{2}\big)^{-1} |x|^z$$

when $\mathrm{Re}\, z > -n$ and the definition makes sense only for that range of z's. It follows from its definition that u_z is a homogeneous distribution of degree z.

We would like to extend the definition of u_z for $z \in \mathbf{C}$. Let $\mathrm{Re}\, z > -n$ first. Fix N to be a positive integer. Given $f \in \mathscr{S}(\mathbf{R}^n)$, write the integral in (2.4.7) as follows:

$$\int_{|x|<1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \left\{ f(x) - \sum_{|\alpha|\le N} \frac{(\partial^\alpha f)(0)}{\alpha!} x^\alpha \right\} |x|^z\, dx$$

$$+ \int_{|x|>1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} f(x)|x|^z\, dx + \int_{|x|<1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \sum_{|\alpha|\le N} \frac{(\partial^\alpha f)(0)}{\alpha!} x^\alpha |x|^z\, dx.$$

The preceding expression is equal to

$$\int_{|x|<1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \left\{ f(x) - \sum_{|\alpha|\le N} \frac{(\partial^\alpha f)(0)}{\alpha!} x^\alpha \right\} |x|^z\, dx$$

$$+ \int_{|x|>1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} f(x)|x|^z\, dx$$

$$+ \sum_{|\alpha|\le N} \frac{(\partial^\alpha f)(0)}{\alpha!} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \int_{r=0}^{1} \int_{\mathbf{S}^{n-1}} (r\theta)^\alpha\, r^{z+n-1}\, dr\, d\theta,$$

where we switched to polar coordinates in the penultimate integral. Now set

$$\begin{aligned} b(n,\alpha,z) &= \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \frac{1}{\alpha!} \left(\int_{\mathbf{S}^{n-1}} \theta^\alpha \, d\theta \right) \int_0^1 r^{|\alpha|+n+z-1} \, dr \\ &= \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \frac{\frac{1}{\alpha!} \int_{\mathbf{S}^{n-1}} \theta^\alpha \, d\theta}{|\alpha|+z+n}, \end{aligned}$$

where $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index. These coefficients are zero when at least one α_j is odd. Consider now the case that all the α_j's are even; then $|\alpha|$ is also even. The function $\Gamma(\frac{z+n}{2})$ has simple poles at the points

$$z = -n, \quad z = -(n+2), \quad z = -(n+4), \qquad \text{and so on;}$$

see Appendix A.5. These poles cancel exactly the poles of the function

$$z \mapsto (|\alpha| + z + n)^{-1}$$

at $z = -n - |\alpha|$ when $|\alpha|$ is an even integer in $[0,N]$. We therefore have

$$\begin{aligned} \langle u_z, f \rangle &= \int_{|x|\geq 1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} f(x)|x|^z \, dx + \sum_{|\alpha|\leq N} b(n,\alpha,z)\,(-1)^{|\alpha|} \langle \partial^\alpha \delta_0, f \rangle \\ &\quad + \int_{|x|<1} \frac{\pi^{\frac{z+n}{2}}}{\Gamma(\frac{z+n}{2})} \left\{ f(x) - \sum_{|\alpha|\leq N} \frac{(\partial^\alpha f)(0)}{\alpha!} x^\alpha \right\} |x|^z \, dx. \end{aligned} \tag{2.4.8}$$

Both integrals converge absolutely when $\operatorname{Re} z > -N-n-1$, since the expression inside the curly brackets above is bounded by a constant multiple of $|x|^{N+1}$, and the resulting function of z in (2.4.8) is a well defined analytic function in the range $\operatorname{Re} z > -N-n-1$.

Since N was arbitrary, $\langle u_z, f \rangle$ has an analytic extension to all of $\mathbf{C}$. Therefore, u_z is a *distribution-valued entire function* of z, i.e., for all $\varphi \in \mathscr{S}(\mathbf{R}^n)$, the function $z \mapsto \langle u_z, \varphi \rangle$ is entire.

Next we would like to calculate the Fourier transform of u_z. We know by Exercise 2.3.9 that $\widehat{u_z}$ is a homogeneous distribution of degree $-n-z$. The choice of constant in the definition of u_z was made to justify the following result:

Theorem 2.4.6. *For all $z \in \mathbf{C}$ we have $\widehat{u_z} = u_{-n-z}$.*

Proof. The idea of the proof is straightforward. First we show that for a certain range of z's we have

$$\int_{\mathbf{R}^n} |\xi|^z \widehat{\varphi}(\xi) \, d\xi = C(n,z) \int_{\mathbf{R}^n} |x|^{-n-z} \varphi(x) \, dx, \tag{2.4.9}$$

for some fixed constant $C(n,z)$ and all $\varphi \in \mathscr{S}(\mathbf{R}^n)$. Next we pick a specific φ to evaluate the constant $C(n,z)$. Then we use analytic continuation to extend the validity of (2.4.9) for all z's. Use polar coordinates by setting $\xi = \rho\varphi$ and $x = r\theta$ in (2.4.9). We have

$$\begin{aligned}
&\int_{\mathbf{R}^n} |\xi|^z \widehat{\varphi}(\xi)\, d\xi \\
&\qquad = \int_0^\infty \rho^{z+n-1} \int_0^\infty \int_{\mathbf{S}^{n-1}} \varphi(r\theta) \left(\int_{\mathbf{S}^{n-1}} e^{-2\pi i r\rho(\theta\cdot\varphi)} d\varphi \right) d\theta\, r^{n-1}\, dr\, d\rho \\
&\qquad = \int_0^\infty \left(\int_0^\infty \sigma_n(r\rho)\rho^{z+n-1}\, d\rho \right) \left(\int_{\mathbf{S}^{n-1}} \varphi(r\theta) d\theta \right) r^{n-1}\, dr \\
&\qquad = C(n,z) \int_0^\infty r^{-z-n} \left(\int_{\mathbf{S}^{n-1}} \varphi(r\theta)\, d\theta \right) r^{n-1}\, dr \\
&\qquad = C(n,z) \int_{\mathbf{R}^n} |x|^{-n-z} \varphi(x)\, dx\,,
\end{aligned}$$

where we set

$$\sigma_n(t) = \int_{\mathbf{S}^{n-1}} e^{-2\pi i t(\theta\cdot\varphi)}\, d\varphi = \int_{\mathbf{S}^{n-1}} e^{-2\pi i t(\varphi_1)}\, d\varphi\,, \tag{2.4.10}$$

$$C(n,z) = \int_0^\infty \sigma_n(t) t^{z+n-1}\, dt\,, \tag{2.4.11}$$

and the second equality in (2.4.10) is a consequence of rotational invariance. It remains to prove that the integral in (2.4.11) converges for some range of z's.

If $n = 1$, then

$$\sigma_1(t) = \int_{\mathbf{S}^0} e^{-2\pi i t\varphi} d\varphi = e^{-2\pi i t} + e^{2\pi i t} = 2\cos(2\pi t)$$

and the integral in (2.4.11) converges conditionally for $-1 < \operatorname{Re} z < 0$.

Let us therefore assume that $n \geq 2$. Since $|\sigma_n(t)| \leq \omega_{n-1}$, the integral converges near zero when $-n < \operatorname{Re} z$. Let us study the behavior of $\sigma_n(t)$ for t large. Using the formula in Appendix D.2 and the definition of Bessel functions in Appendix B.1, we write

$$\sigma_n(t) = \omega_{n-2} \int_{-1}^{1} e^{2\pi i t s} \left(\sqrt{1-s^2}\right)^{n-2} \frac{ds}{\sqrt{1-s^2}} = c_n t^{-\frac{n-2}{2}} J_{\frac{n-2}{2}}(2\pi t),$$

for some constant c_n. Since $n \geq 2$ we have when $n-2 > -1/2$. Then the asymptotics for Bessel functions (Appendix B.7) apply and yield $|\sigma_n(t)| \leq c t^{-(n-1)/2}$ for $t \geq 1$. Splitting the integral in (2.4.11) in $t \leq 1$ and $t \geq 1$ and using the corresponding estimates, we notice that it converges absolutely on $[0,1]$ when $\operatorname{Re} z > -n$ and on $[1,\infty)$ when $\operatorname{Re} z + n - 1 - \frac{n-1}{2} < -1$.

We have now proved that when $-n < \operatorname{Re} z < -\frac{n+1}{2}$ and $n \geq 2$ we have

$$\widehat{u_z} = C(n,z) u_{-n-z}$$

for some constant $C(n,z)$ that we wish to compute. Insert the function $\varphi(x)=e^{-\pi|x|^2}$ in (2.4.9). Example 2.2.9 gives that this function is equal to its Fourier transform. Use polar coordinates to write

$$\omega_{n-1}\int_0^\infty r^{z+n-1}e^{-\pi r^2}\,dr = C(n,z)\,\omega_{n-1}\int_0^\infty r^{-z-n+n-1}e^{-\pi r^2}\,dr\,.$$

Change variables $s=\pi r^2$ and use the definition of the gamma function to obtain that

$$C(n,z)=\frac{\Gamma(\frac{z+n}{2})}{\Gamma(-\frac{z}{2})}\,\frac{\pi^{-\frac{z+n}{2}}}{\pi^{\frac{z}{2}}}\,.$$

It follows that $\widehat{u_z}=u_{-n-z}$ for the range of z's considered.

At this point observe that for every $f\in\mathscr{S}(\mathbf{R}^n)$, the function $z\mapsto\langle\widehat{u_z}-u_{-z-n},f\rangle$ is entire and vanishes for $-n<\operatorname{Re}z<-n+1/2$. Therefore, it must vanish everywhere and the theorem is proved. $\square$

Homogeneous distributions were introduced in Exercise 2.3.9. We already saw that the Dirac mass on $\mathbf{R}^n$ is a homogeneous distribution of degree $-n$. There is another important example of a homogeneous distributions of degree $-n$, which we now discuss.

Let Ω be an integrable function on the sphere $\mathbf{S}^{n-1}$ with integral zero. Define a tempered distribution W_Ω on $\mathbf{R}^n$ by setting

$$\big\langle W_\Omega,f\big\rangle=\lim_{\varepsilon\to0}\int_{|x|\ge\varepsilon}\frac{\Omega(x/|x|)}{|x|^n}\,f(x)\,dx\,. \tag{2.4.12}$$

We check that W_Ω is a well defined tempered distribution on $\mathbf{R}^n$. Indeed, since $\Omega(x/|x|)/|x|^n$ has integral zero over all annuli centered at the origin, we obtain

$$\begin{aligned}\big|\big\langle W_\Omega,\varphi\big\rangle\big| &= \left|\lim_{\varepsilon\to0}\int_{\varepsilon\le|x|\le1}\frac{\Omega(x/|x|)}{|x|^n}(\varphi(x)-\varphi(0))\,dx+\int_{|x|\ge1}\frac{\Omega(x/|x|)}{|x|^n}\varphi(x)\,dx\right|\\ &\le \big\|\nabla\varphi\big\|_{L^\infty}\int_{|x|\le1}\frac{|\Omega(x/|x|)|}{|x|^{n-1}}\,dx+\Big(\sup_{x\in\mathbf{R}^n}|x|\,|\varphi(x)|\Big)\int_{|x|\ge1}\frac{|\Omega(x/|x|)|}{|x|^{n+1}}\,dx\\ &\le C_1\big\|\nabla\varphi\big\|_{L^\infty}\big\|\Omega\big\|_{L^1(\mathbf{S}^{n-1})}+C_2\sum_{|\alpha|\le1}\big\|\varphi(x)x^\alpha\big\|_{L^\infty}\big\|\Omega\big\|_{L^1(\mathbf{S}^{n-1})}\,,\end{aligned}$$

for suitable constants C_1 and C_2 in view of (2.2.2).

One can verify that $W_\Omega\in\mathscr{S}'(\mathbf{R}^n)$ is a homogeneous distribution of degree $-n$ just like the Dirac mass at the origin. It is an interesting fact that all homogeneous distributions on $\mathbf{R}^n$ of degree $-n$ that coincide with a smooth function away from the origin arise in this way. We have the following result.

Proposition 2.4.7. *Suppose that m is a $\mathscr{C}^\infty$ function on $\mathbf{R}^n\setminus\{0\}$ that is homogeneous of degree zero. Then there exist a scalar b and a $\mathscr{C}^\infty$ function Ω on $\mathbf{S}^{n-1}$ with integral zero such that*

$$m^\vee=b\,\delta_0+W_\Omega\,, \tag{2.4.13}$$

where W_Ω denotes the distribution defined in (2.4.12).

To prove this result we need the following proposition, whose proof we postpone until the end of this section.

Proposition 2.4.8. *Suppose that u is a $\mathscr{C}^\infty$ function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree $z \in \mathbf{C}$. Then $\widehat{u}$ is a $\mathscr{C}^\infty$ function on $\mathbf{R}^n \setminus \{0\}$.*

We now prove Proposition 2.4.7 using Proposition 2.4.8.

Proof. Let a be the integral of the smooth function m over $\mathbf{S}^{n-1}$. The function $m-a$ is homogeneous of degree zero and thus locally integrable on $\mathbf{R}^n$; hence it can be thought of as a tempered distribution that we call $\widehat{u}$ (the Fourier transform of a tempered distribution u). Since $\widehat{u}$ is a $\mathscr{C}^\infty$ function on $\mathbf{R}^n \setminus \{0\}$, Proposition 2.4.8 implies that u is also a $\mathscr{C}^\infty$ function on $\mathbf{R}^n \setminus \{0\}$. Let Ω be the restriction of u on $\mathbf{S}^{n-1}$. Then Ω is a well defined $\mathscr{C}^\infty$ function on $\mathbf{S}^{n-1}$. Since u is a homogeneous function of degree $-n$ that coincides with the smooth function Ω on $\mathbf{S}^{n-1}$, it follows that $u(x) = \Omega(x/|x|)/|x|^n$ for x in $\mathbf{R}^n \setminus \{0\}$.

We show that Ω has mean value zero over $\mathbf{S}^{n-1}$. Pick a nonnegative, radial, smooth, and nonzero function ψ on $\mathbf{R}^n$ supported in the annulus $1 < |x| < 2$. Switching to polar coordinates, we write

$$\begin{aligned}\langle u, \psi \rangle &= \int_{\mathbf{R}^n} \frac{\Omega(x/|x|)}{|x|^n} \psi(x)\, dx = c_\psi \int_{\mathbf{S}^{n-1}} \Omega(\theta)\, d\theta,\\ \langle u, \psi \rangle &= \langle \widehat{u}, \widehat{\psi} \rangle = \int_{\mathbf{R}^n} (m(\xi)-a)\widehat{\psi}(\xi)\, d\xi = c'_\psi \int_{\mathbf{S}^{n-1}} \big(m(\theta)-a\big)\, d\theta = 0,\end{aligned}$$

and thus Ω has mean value zero over $\mathbf{S}^{n-1}$ (since $c_\psi \neq 0$).

We can now legitimately define the distribution W_Ω, which coincides with the function $\Omega(x/|x|)/|x|^n$ on $\mathbf{R}^n \setminus \{0\}$. But the distribution u also coincides with this function on $\mathbf{R}^n \setminus \{0\}$. It follows that $u - W_\Omega$ is supported at the origin. Proposition 2.4.1 now gives that $u - W_\Omega$ is a sum of derivatives of Dirac masses. Since both distributions are homogeneous of degree $-n$, it follows that

$$u - W_\Omega = c\delta_0 .$$

But $u = (m-a)^\vee = m^\vee - a\delta_0$, and thus $m^\vee = (c+a)\delta_0 + W_\Omega$. This proves the proposition. □

We now turn to the proof of Proposition 2.4.8.

Proof. Let $u \in \mathscr{S}'$ be homogeneous of degree z and $\mathscr{C}^\infty$ on $\mathbf{R}^n \setminus \{0\}$. We need to show that $\widehat{u}$ is $\mathscr{C}^\infty$ away from the origin. We prove that $\widehat{u}$ is $\mathscr{C}^M$ for all M. Fix $M \in \mathbf{Z}^+$ and let α be any multi-index such that

$$|\alpha| > n + M + \operatorname{Re} z . \tag{2.4.14}$$

Pick a $\mathscr{C}^\infty$ function φ on $\mathbf{R}^n$ that is equal to 1 when $|x| \ge 2$ and equal to zero for $|x| \le 1$. Write $u_0 = (1-\varphi)u$ and $u_\infty = \varphi u$. Then

$$\partial^\alpha u = \partial^\alpha u_0 + \partial^\alpha u_\infty \qquad \text{and thus} \qquad \widehat{\partial^\alpha u} = \widehat{\partial^\alpha u_0} + \widehat{\partial^\alpha u_\infty}\,,$$

where the operations are performed in the sense of distributions. Since u_0 is compactly supported, Theorem 2.3.21 implies that $\widehat{\partial^\alpha u_0}$ is $\mathscr{C}^\infty$. Now Leibniz's rule gives that

$$\partial^\alpha u_\infty = v + \varphi \partial^\alpha u,$$

where v is a smooth function supported in the annulus $1 \le |x| \le 2$. Then $\widehat{v}$ is $\mathscr{C}^\infty$ and we need to show only that $\widehat{\varphi \partial^\alpha u}$ is $\mathscr{C}^M$. The function $\varphi \partial^\alpha u$ is actually $\mathscr{C}^\infty$, and by the homogeneity of $\partial^\alpha u$ (Exercise 2.3.9 (c)) we obtain that $(\partial^\alpha u)(x) = |x|^{-|\alpha|+z}(\partial^\alpha u)(x/|x|)$. Since φ is supported away from zero, it follows that

$$|\varphi(x)(\partial^\alpha u)(x)| \le \frac{C_\alpha}{(1+|x|)^{|\alpha|-\operatorname{Re} z}} \tag{2.4.15}$$

for some $C_\alpha > 0$. It is now straightforward to see that if a function satisfies (2.4.15), then its Fourier transform is $\mathscr{C}^M$ whenever (2.4.14) is satisfied. See Exercise 2.4.1.

We conclude that $\widehat{\partial^\alpha u_\infty}$ is a $\mathscr{C}^M$ function whenever (2.4.14) is satisfied; thus so is $\widehat{\partial^\alpha u}$. Since $\widehat{\partial^\alpha u}(\xi) = (2\pi i \xi)^\alpha \widehat{u}(\xi)$, we deduce smoothness for $\widehat{u}$ away from the origin. Let $\xi \neq 0$. Pick a neighborhood V of ξ such that for η in V we have $\eta_j \neq 0$ for some $j \in \{1,\dots,n\}$. Consider the multi-index $(0,\dots,|\alpha|,\dots,0)$ with $|\alpha|$ in the jth coordinate and zeros elsewhere. Then $(2\pi i \eta_j)^{|\alpha|}\widehat{u}(\eta)$ is a $\mathscr{C}^M$ function on V, and thus so is $\widehat{u}(\eta)$, since we can divide by $\eta_j^{|\alpha|}$. We conclude that $\widehat{u}(\xi)$ is $\mathscr{C}^M$ on $\mathbf{R}^n \setminus \{0\}$. Since M is arbitrary, the conclusion follows. $\square$

We end this section with an example that illustrates the usefulness of some of the ideas discussed in this section.

Example 2.4.9. Let η be a smooth radial function on $\mathbf{R}^n$ that is equal to 1 on the set $|x| \ge 1/2$ and vanishes on the set $|x| \le 1/4$. Fix $z \in \mathbf{C}$ satisfy $0 < \operatorname{Re} z < n$. Let $g = \big(\eta(x)|x|^{-z}\big)^\wedge$ be the distributional Fourier transform of $\eta(x)|x|^{-z}$. We show that g is a function that decays faster than $|\xi|^{-N}$ at infinity (for sufficiently large positive number N) and that

$$g(\xi) - \frac{\pi^{z-\frac{n}{2}}\Gamma\big(\frac{n-z}{2}\big)}{\Gamma\big(\frac{z}{2}\big)}|\xi|^{z-n} \tag{2.4.16}$$

is a $\mathscr{C}^\infty$ function on $\mathbf{R}^n$. This example indicates the interplay between the smoothness of a function and the decay of its Fourier transform. The smoothness of the function $\eta(x)|x|^{-z}$ near zero has as a consequence the rapid decay of g near infinity, while the slow decay of $\eta(x)|x|^{-z}$ at infinity reflects the lack of smoothness of $g(\xi)$ at zero, in view of the moderate blowup $|\xi|^{\operatorname{Re} z - n}$ as $|\xi| \to 0$.

To show that g is a function we write it as $g = (|x|^{-z})^\wedge + \big((\eta(x)-1)|x|^{-z}\big)^\wedge$ and we observe that the first term is a function, since $0 < \operatorname{Re} z < n$. Using Theorem 2.4.6 we write

$$g(\xi) = \frac{\pi^{z-\frac{n}{2}}\Gamma(\frac{n-z}{2})}{\Gamma(\frac{z}{2})}|\xi|^{z-n} + \widehat{\varphi}(\xi),$$

where $\widehat{\varphi}(\xi) = \big((\eta(x)-1)|x|^{-z}\big)^\wedge(\xi)$ is a $\mathscr{C}^\infty$ function, since it is the Fourier transform of a compactly supported integrable function. This proves that g is a function and that the difference in (2.4.16) is $\mathscr{C}^\infty$.

Finally, we assert that every derivative of g satisfies $|\partial^\gamma g(\xi)| \le C_{\gamma,N}|\xi|^{-N}$ for all sufficiently large positive integers N when $\xi \neq 0$. Indeed, fix a multi-index γ and write $\partial^\gamma g(\xi) = (|x|^{-z}\eta(x)(-2\pi i x)^\gamma)^\wedge(\xi)$. It follows that

$$(4\pi^2|\xi|^2)^N|\partial^\gamma g(\xi)| = \big|\big(\Delta^N(|x|^{-z}\eta(x)(-2\pi i x)^\gamma)\big)^\wedge(\xi)\big|$$

for all $N \in \mathbf{Z}^+$, where Δ is the Laplacian in the x variable. Using Leibniz's rule we distribute Δ^N to the product. If a derivative falls on η, we obtain a compactly supported smooth function, hence integrable. If all derivatives fall on $|x|^{-z}x^\gamma$, then we obtain a term that decays like $|x|^{-\operatorname{Re} z+|\gamma|-2N}$ at infinity, which is also integrable if N is sufficiently large. Thus the function $|\xi|^{2N}|\partial^\gamma g(\xi)|$ is equal to the Fourier transform of an L^1 function, hence it is bounded, when $2N > n - \operatorname{Re} z + |\gamma|$.

Exercises

2.4.1. Suppose that a function f satisfies the estimate

$$|f(x)| \le \frac{C}{(1+|x|)^N},$$

for some $C > 0$ and $N > n+1$. Then $\widehat{f}$ is $\mathscr{C}^M$ for all $M \in \mathbf{Z}^+$ with $1 \le M < N-n$.

2.4.2. Use Corollary 2.4.3 to prove Liouville's theorem that every bounded harmonic function on $\mathbf{R}^n$ must be a constant. Derive as a consequence the *fundamental theorem of algebra*, stating that every polynomial on $\mathbf{C}$ must have a complex root.

2.4.3. Prove that e^x is not in $\mathscr{S}'(\mathbf{R})$ but that $e^x e^{ie^x}$ is in $\mathscr{S}'(\mathbf{R})$.

2.4.4. Show that the Schwartz function $x \mapsto \operatorname{sech}(\pi x)$, $x \in \mathbf{R}$, coincides with its Fourier transform.
$\big[$*Hint:* Integrate the function e^{iaz} over the rectangular contour with corners $(-R,0)$, $(R,0)$, $(R,i\pi)$, and $(-R,i\pi)$.$\big]$

2.4.5. ([174]) Construct an uncountable family of linearly independent Schwartz functions f_a such that $|f_a| = |f_b|$ and $|\widehat{f_a}| = |\widehat{f_b}|$ for all f_a and f_b in the family.
$\big[$*Hint:* Let w be a smooth nonzero function whose Fourier transform is supported

in the interval $[-1/2,1/2]$ and let ϕ be a real-valued smooth nonconstant periodic function with period 1. Then take $f_a(x) = w(x)e^{i\phi(x-a)}$ for $a \in \mathbf{R}$.]

2.4.6. Let P_y be the Poisson kernel defined in (2.1.13). Prove that for $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$, the function

$$(x,y) \mapsto (P_y * f)(x)$$

is a harmonic function on $\mathbf{R}^{n+1}_+$. Use the Fourier transform and Exercise 2.2.11 to prove that $(P_{y_1} * P_{y_2})(x) = P_{y_1+y_2}(x)$ for all $x \in \mathbf{R}^n$.

2.4.7. (a) For a fixed $x_0 \in \mathbf{S}^{n-1}$, show that the function

$$v(x;x_0) = \frac{1-|x|^2}{|x-x_0|^n}$$

is harmonic on $\mathbf{R}^n \setminus \{x_0\}$.
(b) For fixed $x_0 \in \mathbf{S}^{n-1}$, prove that the family of functions $\theta \mapsto v(rx_0;\theta)$, $0<r<1$, defined on the sphere satisfies

$$\lim_{r\uparrow 1} \int_{\substack{\theta\in\mathbf{S}^{n-1} \\ |\theta-x_0|>\delta}} v(rx_0;\theta)\,d\theta = 0$$

uniformly in x_0. The function $v(rx_0;\theta)$ is called the *Poisson kernel for the sphere*.
(c) Show that

$$\frac{1}{\omega_{n-1}}(1-|x|^2)\int_{\mathbf{S}^{n-1}} \frac{1}{|x-\theta|^n}\,d\theta = 1$$

for all $|x|<1$.
(d) Let f be a continuous function on $\mathbf{S}^{n-1}$. Prove that the function

$$u(x) = \frac{1}{\omega_{n-1}}(1-|x|^2)\int_{\mathbf{S}^{n-1}} \frac{f(\theta)}{|x-\theta|^n}\,d\theta$$

solves the Dirichlet problem $\Delta(u)=0$ on $|x|<1$ with boundary values $u=f$ on $\mathbf{S}^{n-1}$, in the sense $\lim_{r\uparrow 1} u(rx_0) = f(x_0)$ when $|x_0|=1$.
$\big[$*Hint:* Part (c): Apply the mean value property over spheres to the harmonic function $y \mapsto (1-|x|^2|y|^2)\big||x|y-\frac{x}{|x|}\big|^{-n}$.$\big]$

2.4.8. Fix $n \in \mathbf{Z}^+$ with $n \ge 2$ and a real number λ, $0<\lambda<n$. Also fix $\eta \in \mathbf{S}^n$ and $y \in \mathbf{R}^n$.
(a) Prove that

$$\int_{\mathbf{S}^n} |\xi-\eta|^{-\lambda}\,d\xi = 2^{n-\lambda}\frac{\pi^{\frac{n}{2}}\Gamma(\frac{n-\lambda}{2})}{\Gamma(n-\frac{\lambda}{2})}.$$

(b) Prove that

$$\int_{\mathbf{R}^n} |x-y|^{-\lambda}(1+|x|^2)^{\frac{\lambda}{2}-n}\,dx = \frac{\pi^{\frac{n}{2}}\Gamma(\frac{n-\lambda}{2})}{\Gamma(n-\frac{\lambda}{2})}(1+|y|^2)^{-\frac{\lambda}{2}}.$$

[*Hint:* Part (a): See Appendix D.4 Part (b): Use the stereographic projection in Appendix D.6.]

2.4.9. Prove the following *beta integral identity:*

$$\int_{\mathbf{R}^n} \frac{dt}{|x-t|^{\alpha_1}|y-t|^{\alpha_2}} = \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n-\alpha_1}{2}\right)\Gamma\left(\frac{n-\alpha_2}{2}\right)\Gamma\left(\frac{\alpha_1+\alpha_2-n}{2}\right)}{\Gamma\left(\frac{\alpha_1}{2}\right)\Gamma\left(\frac{\alpha_2}{2}\right)\Gamma\left(n-\frac{\alpha_1+\alpha_2}{2}\right)} |x-y|^{n-\alpha_1-\alpha_2},$$

where $0 < \alpha_1, \alpha_2 < n$, $\alpha_1 + \alpha_2 > n$.
[*Hint:* Reduce to the case $y = 0$, interpret the integral as a convolution, and use Theorem 2.4.6.]

2.4.10. (a) Prove that if a continuous integrable function f on $\mathbf{R}^n$ $(n \geq 2)$ is constant on the spheres $r\mathbf{S}^{n-1}$ for all $r > 0$, then so is its Fourier transform.
(b) If a continuous integrable function on $\mathbf{R}^n$ $(n \geq 3)$ is constant on all $(n-2)$–dimensional spheres orthogonal to $e_1 = (1,0,\dots,0)$, then its Fourier transform has the same property.

2.4.11. ([137]) Suppose that $0 < d_1, d_2, d_3 < n$ satisfy $d_1 + d_2 + d_3 = 2n$. Prove that for any distinct $x, y, z \in \mathbf{R}^n$ we have the identity

$$\int_{\mathbf{R}^n} |x-t|^{-d_2}|y-t|^{-d_3}|z-t|^{-d_1}\,dt$$
$$= \pi^{\frac{n}{2}} \left(\prod_{j=1}^{3} \frac{\Gamma\left(n-\frac{d_j}{2}\right)}{\Gamma\left(\frac{d_j}{2}\right)} \right) |x-y|^{d_1-n}|y-z|^{d_2-n}|z-x|^{d_3-n}.$$

[*Hint:* Reduce matters to the case that $z = 0$ and $y = e_1$. Then take the Fourier transform in x and use that the function $h(t) = |t-e_1|^{-d_3}|t|^{-d_1}$ satisfies $\widehat{h}(\xi) = \widehat{h}(A_\xi^{-2}\xi)$ for all $\xi \neq 0$, where A_ξ is an orthogonal matrix with $A_\xi e_1 = \xi/|\xi|$.]

2.4.12. (a) Integrate the function e^{iz^2} over the contour consisting of the three pieces $P_1 = \{x + i0 : 0 \leq x \leq R\}$, $P_2 = \{Re^{i\theta} : 0 \leq \theta \leq \frac{\pi}{4}\}$, and $P_3 = \{re^{i\frac{\pi}{4}} : 0 \leq r \leq R\}$ (with the proper orientation) to obtain the *Fresnel integral identity*:

$$\lim_{R\to\infty} \int_0^R e^{ix^2}\,dx = \frac{\sqrt{2\pi}}{4}(1+i).$$

(b) Use the result in part (a) to show that the Fourier transform of the function $e^{i\pi|x|^2}$ in $\mathbf{R}^n$ is equal to $e^{i\frac{\pi n}{4}}e^{-i\pi|\xi|^2}$.
[*Hint:* Part (a): On P_2 we have $e^{-R^2\sin(2\theta)} \leq e^{-\frac{4}{\pi}R^2\theta}$, and the integral over P_2 tends to 0. Part (b): Try first $n = 1$.]

2.5 Convolution Operators on L^p Spaces and Multipliers

In this section we study the class of operators that commute with translations. We prove in this section that bounded operators that commute with translations must be of convolution type. Convolution operators arise in many situations, and we would like to know under what circumstances they are bounded between L^p spaces.

2.5.1 Operators That Commute with Translations

Definition 2.5.1. A vector space X of measurable functions on $\mathbf{R}^n$ is called *closed under translations* if for $f \in X$ we have $\tau^z(f) \in X$ for all $z \in \mathbf{R}^n$. Let X and Y be vector spaces of measurable functions on $\mathbf{R}^n$ 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

$$T(\tau^y(f)) = \tau^y(T(f))$$

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

It is straightforward to see that convolution operators commute with translations, i.e., $\tau^y(f * g) = \tau^y(f) * g$ whenever the convolution is defined. One of the goals of this section is to prove the converse: every bounded linear operator that commutes with translations is of convolution type. We have the following:

Theorem 2.5.2. *Let $1 \le p,q \le \infty$ and suppose T is a bounded linear operator from $L^p(\mathbf{R}^n)$ to $L^q(\mathbf{R}^n)$ that commutes with translations. Then there exists a unique tempered distribution w such that*

$$T(f) = f * w \qquad a.e. \text{ for all } f \in \mathscr{S}.$$

A very important point to make is that if $p = \infty$, the restriction of T on $\mathscr{S}$ does not uniquely determine T on the entire L^∞; see Example 2.5.9 and the comments preceding it about this. The theorem is a consequence of the following two results:

Lemma 2.5.3. *Under the hypotheses of Theorem 2.5.2 and for $f \in \mathscr{S}(\mathbf{R}^n)$, the distributional derivatives of $T(f)$ are L^q functions that satisfy*

$$\partial^\alpha\big(T(f)\big) = T(\partial^\alpha f), \qquad \text{for all multi-indices } \alpha. \tag{2.5.1}$$

Lemma 2.5.4. *Let $1 \le q \le \infty$ and let $h \in L^q(\mathbf{R}^n)$. If all distributional derivatives $\partial^\alpha h$ are also in L^q, then h is almost everywhere equal to a continuous function H satisfying*

$$|H(0)| \le C_{n,q} \sum_{|\alpha| \le n+1} \big\|\partial^\alpha h\big\|_{L^q}. \tag{2.5.2}$$

Proof. Assuming Lemmas 2.5.3 and 2.5.4, we prove Theorem 2.5.2.

Given $f \in \mathscr{S}(\mathbf{R}^n)$, by Lemmas 2.5.3 and 2.5.4, there is a continuous function H such that $T(f) = H$ a.e. and such that

$$|H(0)| \le C_{n,q} \sum_{|\alpha| \le n+1} \big\|\partial^\alpha T(f)\big\|_{L^q}$$

holds. Define a linear functional u on $\mathscr{S}$ by setting

$$\big\langle u, f \big\rangle = H(0).$$

This functional is well-defined, for, if there is another continuous function G such that $G = T(f)$ a.e., then $G = H$ a.e. and since both functions are continuous, it follows that $H = G$ everywhere and thus $H(0) = G(0)$.

By (2.5.1), (2.5.2), and the boundedness of T, we have

$$\begin{aligned}
\big|\big\langle u, f \big\rangle\big| &\le C_{n,q} \sum_{|\alpha| \le n+1} \big\|\partial^\alpha T(f)\big\|_{L^q} \\
&\le C_{n,q} \sum_{|\alpha| \le n+1} \big\|T(\partial^\alpha f)\big\|_{L^q} \\
&\le C_{n,q} \big\|T\big\|_{L^p \to L^q} \sum_{|\alpha| \le n+1} \big\|\partial^\alpha f\big\|_{L^p} \\
&\le C'_{n,q} \big\|T\big\|_{L^p \to L^q} \sum_{\substack{|\gamma| \le [\frac{n+1}{p}]+1 \\ |\alpha| \le n+1}} \rho_{\gamma,\alpha}(f),
\end{aligned}$$

where the last estimate uses (2.2.8). This implies that u is in $\mathscr{S}'$. We now set $w = \widetilde{u}$ and we claim that for all $x \in \mathbf{R}^n$ we have

$$\big\langle u, \tau^{-x} f \big\rangle = H(x). \tag{2.5.3}$$

Assuming (2.5.3) we prove that $T(f) = f * w$ for $f \in \mathscr{S}$. To see this, by Theorem 2.3.20 and by the translation invariance of T, for a given $f \in \mathscr{S}(\mathbf{R}^n)$ we have

$$(f * w)(x) = \big\langle \widetilde{u}, \tau^x \widetilde{f} \big\rangle = \big\langle u, \tau^{-x} f \big\rangle = H(x) = T(f)(x),$$

where the last equality holds for almost all x, by the definition of H. Thus $f * w = T(f)$ a.e., as claimed. The uniqueness of w follows from the simple observation that if $f * w = f * w'$ for all $f \in \mathscr{S}(\mathbf{R}^n)$, then $w = w'$.

We now turn to the proof of (2.5.3). Given $f \in \mathscr{S}(\mathbf{R}^n)$ and $x \in \mathbf{R}^n$ and let H_x be the continuous function such that $H_x = T(\tau^{-x} f)$. We show that $H_x(0) = H(x)$. Indeed, we have

$$H_x(y) = T(\tau^{-x} f)(y) = \tau^{-x} T(f)(y) = T(f)(x+y) = H(x+y) = \tau^{-x} H(y),$$

where the equality $T(f)(x+y) = H(x+y)$ holds a.e. in y. Thus the continuous functions H_x and $\tau^{-x}H$ are equal a.e. and thus they must be everywhere equal, in particular, when $y = 0$. This proves that $H_x(0) = H(x)$, which is a restatement of (2.5.3). □

We now return to Lemmas 2.5.3 and 2.5.4. We begin with Lemma 2.5.3.

Proof. Consider first the multi-index $\alpha = (0,\dots,1,\dots,0)$, where 1 is in the jth entry and 0 is elsewhere. Let $e_j = (0,\dots,1,\dots,0)$, where 1 is in the jth entry and zero elsewhere. We have

$$\int_{\mathbf{R}^n} T(f)(y)\frac{\varphi(y+he_j)-\varphi(y)}{h}\,dy = \int_{\mathbf{R}^n} \varphi(y)T\left(\frac{\tau^{he_j}(f)-f}{h}\right)(y)\,dy \tag{2.5.4}$$

since both of these expressions are equal to

$$\int_{\mathbf{R}^n} \varphi(y)\frac{T(f)(y-he_j)-T(f)(y)}{h}\,dy$$

and T commutes with translations. We will let $h \to 0$ in both sides of (2.5.4). We write

$$\frac{\varphi(y+he_j)-\varphi(y)}{h} = \int_0^1 \partial_j\varphi(y+hte_j)\,dt\,,$$

from which it follows that for $|h| < 1/2$ we have

$$\left|\frac{\varphi(y+he_j)-\varphi(y)}{h}\right| \le \int_0^1 \frac{C_M\,dt}{(1+|y+hte_j|)^M} \le \int_0^1 \frac{C_M\,dt}{(1+|y|-\frac{1}{2})^M} \le \frac{C'_M}{(|y|+1)^M}.$$

The integrand on the left-hand side of (2.5.4) is bounded by the integrable function $|T(f)(y)|C'_M(|y|+1)^{-M}$ and converges to $T(f)(y)\,\partial_j\varphi(y)$ as $h \to 0$. The Lebesgue dominated convergence theorem yields that the integral on the left-hand side of (2.5.4) converges to

$$\int_{\mathbf{R}^n} T(f)(y)\,\partial_j\varphi(y)\,dy\,. \tag{2.5.5}$$

Moreover, for a Schwartz function f we have

$$\frac{\tau^{he_j}(f)(y)-f(y)}{h} = \int_0^1 \partial_j f(y+hte_j)\,dt\,,$$

which converges to $\partial_j f(y)$ pointwise as $h \to 0$ and is bounded by $C'_M(1+|y|)^{-M}$ for $|h| < 1/2$ by an argument similar to the preceding one for φ in place of f. Thus

$$\frac{\tau^{he_j}(f)-f}{h} \to \partial_j f \quad \text{in } L^p \text{ as } h \to 0, \tag{2.5.6}$$

by the Lebesgue dominated convergence theorem. The boundedness of T from L^p to L^q yields that

$$T\left(\frac{\tau^{he_j}(f)-f}{h}\right)\to T(\partial_j f) \quad \text{in } L^q \text{ as } h\to 0. \tag{2.5.7}$$

Since $\varphi\in L^{q'}$, by Hölder's inequality, the right-hand side of (2.5.4) converges to

$$\int_{\mathbf{R}^n}\varphi(y)T(\partial_j f)(y)\,dy$$

as $h\to 0$. This limit is equal to (2.5.5) and the required conclusion follows for $\alpha=(0,\dots,0,1,0,\dots,0)$. The general case follows by induction on $|\alpha|$. □

We now prove Lemma 2.5.4.

Proof. Let $R\ge 1$. Fix a $\mathscr{C}_0^\infty$ function φ_R that is equal to 1 in the ball $|x|\le R$ and equal to zero when $|x|\ge 2R$. Since h is in $L^q(\mathbf{R}^n)$, it follows that $\varphi_R h$ is in $L^1(\mathbf{R}^n)$. We show that $\widehat{\varphi_R h}$ is also in L^1. We begin with the inequality

$$1\le C_n(1+|x|)^{-(n+1)}\sum_{|\alpha|\le n+1}|(-2\pi i x)^\alpha|, \tag{2.5.8}$$

which is just a restatement of (2.2.3). Now multiply (2.5.8) by $|\widehat{\varphi_R h}(x)|$ to obtain

$$\begin{aligned}
|\widehat{\varphi_R h}(x)| &\le C_n(1+|x|)^{-(n+1)}\sum_{|\alpha|\le n+1}|(-2\pi i x)^\alpha\widehat{\varphi_R h}(x)|\\
&\le C_n(1+|x|)^{-(n+1)}\sum_{|\alpha|\le n+1}\big\|(\partial^\alpha(\varphi_R h))^\wedge\big\|_{L^\infty}\\
&\le C_n(1+|x|)^{-(n+1)}\sum_{|\alpha|\le n+1}\big\|\partial^\alpha(\varphi_R h)\big\|_{L^1}\\
&\le C_n(2^nR^nv_n)^{1/q'}(1+|x|)^{-(n+1)}\sum_{|\alpha|\le n+1}\big\|\partial^\alpha(\varphi_R h)\big\|_{L^q}\\
&\le C_{n,R}(1+|x|)^{-(n+1)}\sum_{|\alpha|\le n+1}\big\|\partial^\alpha h\big\|_{L^q},
\end{aligned}$$

where we used Leibniz's rule (Proposition 2.3.22 (14)) and the fact that all derivatives of φ_R are pointwise bounded by constants depending on R.

Integrate the previously displayed inequality with respect to x to obtain

$$\big\|\widehat{\varphi_R h}\big\|_{L^1}\le C_{R,n}\sum_{|\alpha|\le n+1}\big\|\partial^\alpha h\big\|_{L^q}<\infty. \tag{2.5.9}$$

Therefore, Fourier inversion holds for $\varphi_R h$ (see Exercise 2.2.6). This implies that $\varphi_R h$ is equal a.e. to a continuous function, namely the inverse Fourier transform of its Fourier transform. Since $\varphi_R=1$ on the ball $B(0,R)$, we conclude that h is a.e. equal to a continuous function in this ball. Since $R>0$ was arbitrary, it follows that

h is a.e. equal to a continuous function on $\mathbf{R}^n$, which we denote by H. Finally, (2.5.2) is a direct consequence of (2.5.9) with $R = 1$, since $|H(0)| \le \|\widehat{\varphi_1 h}\|_{L^1}$. □

2.5.2 The Transpose and the Adjoint of a Linear Operator

We briefly discuss the notions of the transpose and the adjoint of a linear operator. We first recall real and complex inner products. For f, g measurable functions on $\mathbf{R}^n$, we define the *complex inner product*

$$\langle f \,|\, g \rangle = \int_{\mathbf{R}^n} f(x)\overline{g(x)}\, dx\,,$$

whenever the integral converges absolutely. We reserve the notation

$$\langle f, g \rangle = \int_{\mathbf{R}^n} f(x)g(x)\, dx$$

for the *real inner product* on $L^2(\mathbf{R}^n)$ and also for the action of a distribution f on a test function g. (This notation also makes sense when a distribution f coincides with a function.)

Let $1 \le p, q \le \infty$. For a bounded linear operator T from $L^p(X,\mu)$ to $L^q(Y,\nu)$ we denote by T^* its *adjoint operator* defined by

$$\langle T(f) \,|\, g \rangle = \int_Y T(f)\,\overline{g}\, d\nu = \int_X f\,\overline{T^*(g)}\, d\mu = \langle f \,|\, T^*(g) \rangle \tag{2.5.10}$$

for f in $L^p(X,\mu)$ and g in $L^{q'}(Y,\nu)$ (or in a dense subspace of it). We also define the *transpose* of T as the unique operator T^t that satisfies

$$\langle T(f), g \rangle = \int_Y T(f)\, g\, dx = \int_X f\, T^t(g)\, dx = \langle f, T^t(g) \rangle$$

for all $f \in L^p(X,\mu)$ and all $g \in L^{q'}(Y,\nu)$.

If T is an integral operator of the form

$$T(f)(x) = \int_X K(x,y) f(y)\, d\mu(y),$$

then T^* and T^t are also integral operators with kernels $K^*(x,y) = \overline{K(y,x)}$ and $K^t(x,y) = K(y,x)$, respectively. If T has the form $T(f) = (\widehat{f} m)^\vee$, that is, it is given by multiplication on the Fourier transform by a (complex-valued) function $m(\xi)$, then T^* is given by multiplication on the Fourier transform by the function $\overline{m(\xi)}$. Indeed for f, g in $\mathscr{S}(\mathbf{R}^n)$ we have

$$\begin{aligned}\int_{\mathbf{R}^n} f\,\overline{T^*(g)}\,dx &= \int_{\mathbf{R}^n} T(f)\,\overline{g}\,dx\\ &= \int_{\mathbf{R}^n} \widehat{T(f)}\,\overline{\widehat{g}}\,d\xi\\ &= \int_{\mathbf{R}^n} \widehat{f}\,\overline{\overline{m}\widehat{g}}\,d\xi\\ &= \int_{\mathbf{R}^n} f\,\overline{(\overline{m}\widehat{g})^\vee}\,dx\,.\end{aligned}$$

A similar argument (using Theorem 2.2.14 (5)) gives that if T is given by multiplication on the Fourier transform by the function $m(\xi)$, then T^t is given by multiplication on the Fourier transform by the function $m(-\xi)$. Since the complex-valued functions $\overline{m(\xi)}$ and $m(-\xi)$ may be different, the operators T^* and T^t may be different in general. Also, if $m(\xi)$ is real-valued, then T is *self-adjoint* (i.e., $T = T^*$) while if $m(\xi)$ is even, then T is *self-transpose* (i.e., $T = T^t$).

2.5.3 The Spaces $\mathscr{M}^{p,q}(\mathbf{R}^n)$

Definition 2.5.5. Given $1 \le p,q \le \infty$, we denote by $\mathscr{M}^{p,q}(\mathbf{R}^n)$ the set of all bounded linear operators from $L^p(\mathbf{R}^n)$ to $L^q(\mathbf{R}^n)$ that commute with translations.

By Theorem 2.5.2 we have that every T in $\mathscr{M}^{p,q}$ is given by convolution with a tempered distribution. We introduce a norm on $\mathscr{M}^{p,q}$ by setting

$$\|T\|_{\mathscr{M}^{p,q}} = \|T\|_{L^p\to L^q},$$

that is, the norm of T in $\mathscr{M}^{p,q}$ is the operator norm of T as an operator from L^p to L^q. It is a known fact that under this norm, $\mathscr{M}^{p,q}$ is a complete normed space (i.e., a Banach space).

Next we show that when $p > q$ the set $\mathscr{M}^{p,q}$ consists of only one element, namely the zero operator $T = 0$. This means that the only interesting classes of operators arise when $p \le q$.

Theorem 2.5.6. $\mathscr{M}^{p,q} = \{0\}$ *whenever* $1 \le q < p < \infty$.

Proof. Let f be a nonzero $\mathscr{C}_0^\infty$ function and let $h \in \mathbf{R}^n$. We have

$$\big\|\tau^h(T(f)) + T(f)\big\|_{L^q} = \big\|T(\tau^h(f)+f)\big\|_{L^q} \le \big\|T\big\|_{L^p\to L^q}\big\|\tau^h(f)+f\big\|_{L^p}.$$

Now let $|h| \to \infty$ and use Exercise 2.5.1. We conclude that

$$2^{\frac{1}{q}}\big\|T(f)\big\|_{L^q} \le \big\|T\big\|_{L^p\to L^q}2^{\frac{1}{p}}\big\|f\big\|_{L^p},$$

which is impossible if $q < p$ unless T is the zero operator. $\square$

Next we have a theorem concerning the duals of the spaces $\mathscr{M}^{p,q}(\mathbf{R}^n)$.

Theorem 2.5.7. *Let $1 < p \le q < \infty$ and $T \in \mathscr{M}^{p,q}(\mathbf{R}^n)$. Then T can be defined on $L^{q'}(\mathbf{R}^n)$, coinciding with its previous definition on the subspace $L^p(\mathbf{R}^n) \cap L^{q'}(\mathbf{R}^n)$ of $L^p(\mathbf{R}^n)$, so that it maps $L^{q'}(\mathbf{R}^n)$ to $L^{p'}(\mathbf{R}^n)$ with norm*

$$\|T\|_{L^{q'} \to L^{p'}} = \|T\|_{L^p \to L^q}. \tag{2.5.11}$$

In other words, we have the following isometric identification of spaces:

$$\mathscr{M}^{q',p'}(\mathbf{R}^n) = \mathscr{M}^{p,q}(\mathbf{R}^n).$$

Proof. We first observe that if $T : L^p \to L^q$ is given by convolution with $u \in \mathscr{S}'$, then the adjoint operator $T^* : L^{q'} \to L^{p'}$ is given by convolution with $\overline{\widetilde{u}} \in \mathscr{S}'$. Indeed, for $f, g \in \mathscr{S}(\mathbf{R}^n)$ we have

$$\begin{aligned}
\int_{\mathbf{R}^n} f\, \overline{T^*(g)}\, dx &= \int_{\mathbf{R}^n} T(f)\, \overline{g}\, dx \\
&= \int_{\mathbf{R}^n} (f * u)\, \overline{g}\, dx \\
&= \int_{\mathbf{R}^n} f\, (\overline{g} * \widetilde{u})\, dx \\
&= \int_{\mathbf{R}^n} f\, \overline{g * \overline{\widetilde{u}}}\, dx.
\end{aligned}$$

Therefore T^* is given by convolution with $\overline{\widetilde{u}}$ when applied to Schwartz functions.

Next we observe the validity of the identity

$$\overline{f * \overline{\widetilde{u}}} = (\overline{\widetilde{f}} * u)^{\sim}, \qquad\qquad f \in \mathscr{S}. \tag{2.5.12}$$

It remains to show that T (convolution with u) and T^* (convolution with $\overline{\widetilde{u}}$) map $L^{q'}$ to $L^{p'}$ with the same norm. But this easily follows from (2.5.12), which implies that

$$\frac{\|f * \overline{\widetilde{u}}\|_{L^{p'}}}{\|f\|_{L^{q'}}} = \frac{\|\overline{\widetilde{f}} * u\|_{L^{p'}}}{\|\overline{\widetilde{f}}\|_{L^{q'}}},$$

for all nonzero Schwartz functions f. We conclude that

$$\|T^*\|_{L^{q'} \to L^{p'}} = \|T\|_{L^{q'} \to L^{p'}}$$

and therefore

$$\|T\|_{L^p \to L^q} = \|T\|_{L^{q'} \to L^{p'}}.$$

This establishes the claimed assertion. $\square$

We next focus attention on the spaces $\mathscr{M}^{p,q}(\mathbf{R}^n)$ whenever $p = q$. These spaces are of particular interest, since they include the singular integral operators, which we study in Chapter 5.

2.5.4 Characterizations of $\mathscr{M}^{1,1}(\mathbf{R}^n)$ and $\mathscr{M}^{2,2}(\mathbf{R}^n)$

It would be desirable to have a characterization of the spaces $\mathscr{M}^{p,p}$ in terms of properties of the convolving distribution. Unfortunately, this is unknown at present (it is not clear whether it is possible) except for certain cases.

Theorem 2.5.8. *An operator T is in $\mathscr{M}^{1,1}(\mathbf{R}^n)$ if and only if it is given by convolution with a finite Borel (complex-valued) measure. In this case, the norm of the operator is equal to the total variation of the measure.*

Proof. If T is given with convolution with a finite Borel measure μ, then clearly T maps L^1 to itself and $\|T\|_{L^1\to L^1} \le \|\mu\|_{\mathscr{M}}$, where $\|\mu\|_{\mathscr{M}}$ is the total variation of μ.

Conversely, let T be an operator bounded from L^1 to L^1 that commutes with translations. By Theorem 2.5.2, T is given by convolution with a tempered distribution u. Let

$$f_\varepsilon(x) = \varepsilon^{-n} e^{-\pi|x/\varepsilon|^2}.$$

Since the functions f_ε are uniformly bounded in L^1, it follows from the boundedness of T that $f_\varepsilon * u$ are also uniformly bounded in L^1. Since L^1 is naturally embedded in the space of finite Borel measures, which is the dual of the space $\mathscr{C}_{00}$ of continuous functions that tend to zero at infinity, we obtain that the family $f_\varepsilon * u$ lies in a fixed multiple of the unit ball of $\mathscr{C}_{00}^*$. By the Banach–Alaoglu theorem, this is a weak* compact set. Therefore, some subsequence of $f_\varepsilon * u$ converges in the weak* topology to a measure μ. That is, for some $\varepsilon_k \to 0$ and all $g \in \mathscr{C}_{00}(\mathbf{R}^n)$ we have

$$\lim_{k\to\infty} \int_{\mathbf{R}^n} g(x)(f_{\varepsilon_k} * u)(x)\,dx = \int_{\mathbf{R}^n} g(x)\,d\mu(x). \tag{2.5.13}$$

We claim that $u = \mu$. To see this, fix $g \in \mathscr{S}$. Equation (2.5.13) implies that

$$\langle u, \widetilde{f_{\varepsilon_k}} * g\rangle = \langle u, f_{\varepsilon_k} * g\rangle \to \langle \mu, g\rangle$$

as $k\to\infty$. Exercise 2.3.2 gives that $g * f_{\varepsilon_k}$ converges to g in $\mathscr{S}$. Therefore,

$$\langle u, f_{\varepsilon_k} * g\rangle \to \langle u, g\rangle.$$

It follows from (2.5.13) that $\langle u, g\rangle = \langle \mu, g\rangle$, and since g was arbitrary, $u = \mu$.

Next, (2.5.13) implies that for all $g \in \mathscr{C}_{00}$ we have

$$\left|\int_{\mathbf{R}^n} g(x)\,d\mu(x)\right| \le \|g\|_{L^\infty} \sup_k \|f_{\varepsilon_k} * u\|_{L^1} \le \|g\|_{L^\infty}\|T\|_{L^1\to L^1}. \tag{2.5.14}$$

The Riesz representation theorem gives that the norm of the functional

$$g \mapsto \int_{\mathbf{R}^n} g(x)\,d\mu(x)$$

on $\mathscr{C}_{00}$ is exactly $\|\mu\|_{\mathscr{M}}$. It follows from (2.5.14) that $\|T\|_{L^1\to L^1} \ge \|\mu\|_{\mathscr{M}}$. Since the reverse inequality is obvious, we conclude that $\|T\|_{L^1\to L^1} = \|\mu\|_{\mathscr{M}}$. □

Let μ be a finite Borel measure. The operator $h \mapsto h * \mu$ maps $L^p(\mathbf{R}^n)$ to itself for all $1 \le p \le \infty$; hence $\mathscr{M}^{1,1}(\mathbf{R}^n)$ can be identified with a subspace of $\mathscr{M}^{\infty,\infty}(\mathbf{R}^n)$. But there exist bounded linear operators Φ on L^∞ that commute with translations for which there does not exist a finite Borel measure μ such that $\Phi(h) = h * \mu$ for all $h \in L^\infty(\mathbf{R}^n)$. The following example captures such a behavior.

Example 2.5.9. Let $(X, \|\cdot\|_{L^\infty})$ be the space of all complex-valued bounded functions on the real line such that

$$\Phi(f) = \lim_{R \to +\infty} \frac{1}{R} \int_0^R f(t)\, dt$$

exists. Then Φ is a bounded linear functional on X with norm 1 and has a bounded extension $\widetilde{\Phi}$ on L^∞ with norm 1, by the Hahn–Banach theorem. We may view $\widetilde{\Phi}$ as a bounded linear operator from $L^\infty(\mathbf{R})$ to the space of constant functions, which is contained in $L^\infty(\mathbf{R})$. We note that $\widetilde{\Phi}$ commutes with translations, since for all $f \in L^\infty(\mathbf{R})$ and $x \in \mathbf{R}$ we have

$$\widetilde{\Phi}(\tau^x(f)) - \tau^x(\widetilde{\Phi}(f)) = \widetilde{\Phi}(\tau^x(f)) - \widetilde{\Phi}(f) = \widetilde{\Phi}(\tau^x(f) - f) = \Phi(\tau^x(f) - f) = 0,$$

where the last two equalities follow from the fact that for L^∞ functions f the expression $\frac{1}{R}\int_0^R (f(t-x) - f(t))\, dt$ is bounded by $\frac{|x|}{R}\|f\|_{L^\infty}$ when $R > |x|$ and thus tends to zero as $R \to \infty$. If $\Phi(\varphi) = \varphi * u$ for some $u \in \mathscr{S}'(\mathbf{R}^n)$ and all $\varphi \in \mathscr{S}(\mathbf{R}^n)$, since Φ vanishes on $\mathscr{S}$, the uniqueness in Theorem 2.5.2 yields that $u = 0$. Hence, if there existed a finite Borel measure μ such that $\widetilde{\Phi}(h) = h * \mu$ all $h \in L^\infty$, in particular we would have $0 = \Phi(\varphi) = \varphi * \mu$ for all $\varphi \in \mathscr{S}$, hence μ would be the zero measure. But obviously, this is not the case, since Φ is not the zero operator on X.

We now study the case $p = 2$. We have the following theorem:

Theorem 2.5.10. *An operator T is in $\mathscr{M}^{2,2}(\mathbf{R}^n)$ if and only if it is given by convolution with some $u \in \mathscr{S}'$ whose Fourier transform $\widehat{u}$ is an L^∞ function. In this case the norm of $T : L^2 \to L^2$ is equal to $\|\widehat{u}\|_{L^\infty}$.*

Proof. If $\widehat{u} \in L^\infty$, Plancherel's theorem gives

$$\int_{\mathbf{R}^n} |f * u|^2\, dx = \int_{\mathbf{R}^n} |\widehat{f}(\xi)\widehat{u}(\xi)|^2\, d\xi \le \big\|\widehat{u}\big\|_{L^\infty}^2 \big\|\widehat{f}\big\|_{L^2}^2;$$

therefore, $\|T\|_{L^2 \to L^2} \le \|\widehat{u}\|_{L^\infty}$, and hence T is in $\mathscr{M}^{2,2}(\mathbf{R}^n)$.

Now suppose that $T \in \mathscr{M}^{2,2}(\mathbf{R}^n)$ is given by convolution with a tempered distribution u. We show that $\widehat{u}$ is a bounded function. For $R > 0$ let φ_R be a $\mathscr{C}_0^\infty$ function supported inside the ball $B(0, 2R)$ and equal to one on the ball $B(0, R)$. The product of the function φ_R with the distribution $\widehat{u}$ is $\varphi_R \widehat{u} = ((\varphi_R)^\vee * u)\widehat{\ } = T(\varphi_R^\vee)\widehat{\ }$, which is an L^2 function. Since the L^2 function $\varphi_R \widehat{u}$ coincides with the distribution $\widehat{u}$ on the set $B(0, R)$, it follows that $\widehat{u}$ is in $L^2(B(0, R))$ for all $R > 0$ and therefore it is in L^2_{loc}. If $f \in L^\infty(\mathbf{R}^n)$ has compact support, the function $f\widehat{u}$ is in L^2, and therefore Plancherel's theorem and the boundedness of T give

$$\int_{\mathbf{R}^n} |f(x)\widehat{u}(x)|^2\,dx = \int_{\mathbf{R}^n} |T(f^\vee)(x)|^2\,dx \le \|T\|^2_{L^2\to L^2} \int_{\mathbf{R}^n} |f(x)|^2\,dx\,.$$

We conclude that for all bounded functions with compact support f we have

$$\int_{\mathbf{R}^n} \big(\|T\|^2_{L^2\to L^2} - |\widehat{u}(x)|^2\big)|f(x)|^2\,dx \ge 0\,.$$

Taking $f(x_1,\dots,x_n) = (2r)^{-n/2}\prod_{j=1}^n \chi_{[-r,r]}(x_j)$ for $r>0$ and using Corollary 2.1.16, we obtain that $\|T\|^2_{L^2\to L^2} - |\widehat{u}(x)|^2 \ge 0$ for almost all x. Hence $\widehat{u}$ is in L^∞ and $\|\widehat{u}\|_{L^\infty} \le \|T\|_{L^2\to L^2}$. Combining this with the estimate $\|T\|_{L^2\to L^2} \le \|\widehat{u}\|_{L^\infty}$, which holds if $\widehat{u} \in L^\infty$, we deduce that $\|T\|_{L^2\to L^2} = \|\widehat{u}\|_{L^\infty}$. $\square$

2.5.5 The Space of Fourier Multipliers $\mathscr{M}_p(\mathbf{R}^n)$

We have now characterized all convolution operators that map L^2 to L^2. Suppose now that T is in $\mathscr{M}^{p,p}$, where $1<p<2$. As discussed in Theorem 2.5.7, T also maps $L^{p'}$ to $L^{p'}$. Since $p<2<p'$, by Theorem 1.3.4, it follows that T also maps L^2 to L^2. Thus T is given by convolution with a tempered distribution whose Fourier transform is a bounded function.

Definition 2.5.11. Given $1\le p<\infty$, we denote by $\mathscr{M}_p(\mathbf{R}^n)$ the space of all bounded functions m on $\mathbf{R}^n$ such that the operator

$$T_m(f) = (\widehat{f}m)^\vee, \qquad f\in\mathscr{S},$$

is bounded on $L^p(\mathbf{R}^n)$ (or is initially defined in a dense subspace of $L^p(\mathbf{R}^n)$ and has a bounded extension on the whole space). The norm of m in $\mathscr{M}_p(\mathbf{R}^n)$ is defined by

$$\big\|m\big\|_{\mathscr{M}_p} = \big\|T_m\big\|_{L^p\to L^p}\,. \tag{2.5.15}$$

Definition 2.5.11 implies that $m\in\mathscr{M}_p$ if and only if $T_m\in\mathscr{M}^{p,p}$. Elements of the space $\mathscr{M}_p$ are called L^p *multipliers* or L^p *Fourier multipliers*. It follows from Theorem 2.5.10 that $\mathscr{M}_2$, the set of all L^2 multipliers, is L^∞. Theorem 2.5.8 implies that $\mathscr{M}_1(\mathbf{R}^n)$ is the set of the Fourier transforms of finite Borel measures that is usually denoted by $\mathscr{M}(\mathbf{R}^n)$. Theorem 2.5.7 states that a bounded function m is an L^p multiplier if and only if it is an $L^{p'}$ multiplier, and in this case

$$\big\|m\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_{p'}}\,, \qquad 1<p<\infty\,.$$

It is a consequence of Theorem 1.3.4 that the normed spaces $\mathscr{M}_p$ are nested, that is, for $1 \le p \le q \le 2$ we have

$$\mathscr{M}_1 \subseteqq \mathscr{M}_p \subseteqq \mathscr{M}_q \subseteqq \mathscr{M}_2 = L^\infty .$$

Moreover, if $m \in \mathscr{M}_p$ and $1 \le p \le 2 \le p'$, Theorem 1.3.4 gives

$$\big\|T_m\big\|_{L^2 \to L^2} \le \big\|T_m\big\|_{L^p \to L^p}^{\frac{1}{2}} \big\|T_m\big\|_{L^{p'} \to L^{p'}}^{\frac{1}{2}} = \big\|T_m\big\|_{L^p \to L^p}, \tag{2.5.16}$$

since $1/2 = (1/2)/p + (1/2)/p'$. Theorem 1.3.4 also gives that

$$\big\|m\big\|_{\mathscr{M}_p} \le \big\|m\big\|_{\mathscr{M}_q}$$

whenever $1 \le q \le p \le 2$. Thus the $\mathscr{M}_p$'s form an increasing family of spaces as p increases from 1 to 2.

Example 2.5.12. The function $m(\xi) = e^{2\pi i \xi \cdot b}$ is an L^p multiplier for all $b \in \mathbf{R}^n$, since the corresponding operator $T_m(f)(x) = f(x+b)$ is bounded on $L^p(\mathbf{R}^n)$. Clearly $\|m\|_{\mathscr{M}_p} = 1$.

Proposition 2.5.13. *For $1 \le p < \infty$, the normed space $\big(\mathscr{M}_p, \|\cdot\|_{\mathscr{M}_p}\big)$ is a Banach space. Furthermore, $\mathscr{M}_p$ is closed under pointwise multiplication and is a Banach algebra.*

Proof. It suffices to consider the case $1 \le p \le 2$. It is straightforward that if m_1, m_2 are in $\mathscr{M}_p$ and $b \in \mathbf{C}$ then $m_1 + m_2$ and bm_1 are also in $\mathscr{M}_p$. Observe that $m_1 m_2$ is the multiplier that corresponds to the operator $T_{m_1} T_{m_2} = T_{m_1 m_2}$ and thus

$$\big\|m_1 m_2\big\|_{\mathscr{M}_p} = \big\|T_{m_1} T_{m_2}\big\|_{L^p \to L^p} \le \big\|m_1\big\|_{\mathscr{M}_p} \big\|m_2\big\|_{\mathscr{M}_p}.$$

This proves that $\mathscr{M}_p$ is an algebra. To show that $\mathscr{M}_p$ is a complete normed space, consider a Cauchy sequence m_j in $\mathscr{M}_p$. It follows from (2.5.16) that m_j is Cauchy in L^∞, and hence it converges to some bounded function m in the L^∞ norm; moreover all the m_j are a.e. bounded by some constant C uniformly in j. We have to show that $m \in \mathscr{M}_p$. Fix $f \in \mathscr{S}$. We have

$$T_{m_j}(f)(x) = \int_{\mathbf{R}^n} \widehat{f}(\xi) m_j(\xi) e^{2\pi i x \cdot \xi}\, d\xi \to \int_{\mathbf{R}^n} \widehat{f}(\xi) m(\xi) e^{2\pi i x \cdot \xi}\, d\xi = T_m(f)(x)$$

a.e. by the Lebesgue dominated convergence theorem, since $C|\widehat{f}|$ is an integrable upper bound of all integrands on the left in the preceding expression. Since $\{m_j\}_j$ is a Cauchy sequence in $\mathscr{M}_p$, it is bounded in $\mathscr{M}_p$, and thus $\sup_j \|m_j\|_{\mathscr{M}_p} < +\infty$. An application of Fatou's lemma yields that

$$\begin{aligned}
\int_{\mathbf{R}^n} |T_m(f)|^p\, dx &= \int_{\mathbf{R}^n} \liminf_{j\to\infty} |T_{m_j}(f)|^p\, dx \\
&\le \liminf_{j\to\infty} \int_{\mathbf{R}^n} |T_{m_j}(f)|^p\, dx \\
&\le \liminf_{j\to\infty} \big\|m_j\big\|_{\mathscr{M}_p}^p \big\|f\big\|_{L^p}^p ,
\end{aligned}$$

which implies that $m \in \mathscr{M}_p$. This argument shows that if $m_j \in \mathscr{M}_p$ and $m_j \to m$ uniformly, then m is in $\mathscr{M}_p$ and satisfies

$$\|m\|_{\mathscr{M}_p} \le \liminf_{j\to\infty} \|m_j\|_{\mathscr{M}_p} .$$

Apply this inequality to $m_k - m_j$ in place of m_j and $m_k - m$ in place of m, for some fixed k. We obtain

$$\|m_k - m\|_{\mathscr{M}_p} \le \liminf_{j\to\infty} \|m_k - m_j\|_{\mathscr{M}_p} \tag{2.5.17}$$

for each k. Given $\varepsilon > 0$, by the Cauchy criterion, there is an N such that for $j,k > N$ we have $\|m_k - m_j\|_{\mathscr{M}_p} < \varepsilon$. Using (2.5.17) we conclude that $\|m_k - m\|_{\mathscr{M}_p} \le \varepsilon$ when $k > N$, thus m_k converges to m in $\mathscr{M}_p$.

This proves that $\mathscr{M}_p$ is a Banach space. □

The following proposition summarizes some simple properties of multipliers.

Proposition 2.5.14. *For all $m \in \mathscr{M}_p$, $1 \le p < \infty$, $x \in \mathbf{R}^n$, and $h > 0$ we have*

$$\big\|\tau^x(m)\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_p}, \tag{2.5.18}$$

$$\big\|\delta^h(m)\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_p}, \tag{2.5.19}$$

$$\big\|\widetilde{m}\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_p},$$

$$\big\|e^{2\pi i(\,\cdot\,)\cdot x} m\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_p},$$

$$\big\|m\circ A\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_p}, \qquad \textit{A is an orthogonal matrix.}$$

Proof. See Exercise 2.5.2. □

Example 2.5.15. We show that for $-\infty < a < b < \infty$ we have $\|\chi_{[a,b]}\|_{\mathscr{M}_p} = \|\chi_{[0,1]}\|_{\mathscr{M}_p}$. Indeed, using (2.5.18) we obtain that $\|\chi_{[a,b]}\|_{\mathscr{M}_p} = \|\chi_{[0,b-a]}\|_{\mathscr{M}_p}$, and the latter is equal to $\|\chi_{[0,1]}\|_{\mathscr{M}_p}$ in view of (2.5.19). The fact that we have $\|\chi_{[0,1]}\|_{\mathscr{M}_p} < \infty$ for all $1 < p < \infty$ is shown in Chapter 5.

We continue with the following interesting result.

Theorem 2.5.16. *Suppose that $m(\xi,\eta) \in \mathscr{M}_p(\mathbf{R}^{n+m})$, where $1 < p < \infty$. Then for almost every $\xi \in \mathbf{R}^n$ the function $\eta \mapsto m(\xi,\eta)$ is in $\mathscr{M}_p(\mathbf{R}^m)$, with*

$$\big\|m(\xi,\cdot)\big\|_{\mathscr{M}_p(\mathbf{R}^m)} \le \big\|m\big\|_{\mathscr{M}_p(\mathbf{R}^{n+m})}.$$

Proof. Since m lies in $L^\infty(\mathbf{R}^{n+m})$, it follows by Fubini's theorem that for almost all $\xi \in \mathbf{R}^n$, the function $\eta \mapsto m(\xi,\eta)$ lies in $L^\infty(\mathbf{R}^m)$ and

$$\big\|m(\xi,\,\cdot\,)\big\|_{L^\infty(\mathbf{R}^m)} \le \big\|m\big\|_{L^\infty(\mathbf{R}^{n+m})}. \tag{2.5.20}$$

Fix f_1, g_1 in $\mathscr{S}(\mathbf{R}^n)$ and f_2, g_2 in $\mathscr{S}(\mathbf{R}^m)$. Define the functions $(f_1 \otimes f_2)(x,y) = f_1(x)f_2(y)$ when $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$. For all ξ for which (2.5.20) is satisfied define

$$M(\xi) = \int_{\mathbf{R}^m} \big(m(\xi,\cdot\,)\widehat{f_2}\big)^\vee(y)\, g_2(y)\, dy = \int_{\mathbf{R}^m} m(\xi,\eta)\widehat{f_2}(\eta) g_2{}^\vee(\eta)\, d\eta$$

and observe that

$$\begin{aligned}
\left|\int_{\mathbf{R}^n} \big(M(\cdot\,)\widehat{f_1}\big)^\vee(x)\, g_1(x)\, dx\right| &= \left|\int_{\mathbf{R}^n} M(\xi)\widehat{f_1}(\xi) g_1{}^\vee(\xi)\, d\xi\right| \\
&= \left|\iint_{\mathbf{R}^{n+m}} m(\xi,\eta)\widehat{f_1 \otimes f_2}(\xi,\eta)(g_1 \otimes g_2)^\vee(\xi,\eta)\, d\xi\, d\eta\right| \\
&= \left|\iint_{\mathbf{R}^{n+m}} (m\,\widehat{f_1 \otimes f_2})^\vee(x,y)(g_1 \otimes g_2)(x,y)\, dx\, dy\right| \\
&\le \big\|m\big\|_{\mathscr{M}_p(\mathbf{R}^{n+m})} \big\|f_1\big\|_{L^p} \big\|f_2\big\|_{L^p} \big\|g_1\big\|_{L^{p'}} \big\|g_2\big\|_{L^{p'}}.
\end{aligned}$$

In view of the identity

$$\big\|(M(\cdot\,)\widehat{f_1})^\vee\big\|_{L^p} = \sup_{\|g_1\|_{L^{p'}} \le 1} \left|\int_{\mathbf{R}^n} \big(M(\cdot\,)\widehat{f_1}\big)^\vee(x)\, g_1(x)\, dx\right|,$$

it follows that, for the ξ that satisfy (2.5.20), $M(\xi)$ lies in $\mathscr{M}_p(\mathbf{R}^n)$ with

$$\big\|M\big\|_{\mathscr{M}_p(\mathbf{R}^n)} \le \big\|m\big\|_{\mathscr{M}_p(\mathbf{R}^{n+m})} \big\|f_2\big\|_{L^p} \big\|g_2\big\|_{L^{p'}}.$$

Since $\|M\|_{L^\infty} \le \|M\|_{\mathscr{M}_p}$ for almost all $\xi \in \mathbf{R}^n$, we obtain

$$\left|\int_{\mathbf{R}^m} (m(\xi,\cdot)\widehat{f_2})^\vee(y)\, g_2(y)\, dy\right| = |M(\xi)| \le \big\|m\big\|_{\mathscr{M}_p(\mathbf{R}^{n+m})} \big\|f_2\big\|_{L^p} \big\|g_2\big\|_{L^{p'}}, \quad (2.5.21)$$

which of course implies the required conclusion, by taking the supremum over all g_2 in $L^{p'}$ with norm at most 1. □

Example 2.5.17. (The cone multiplier) On $\mathbf{R}^{n+1}$ define the function

$$m_\lambda(\xi_1,\dots,\xi_{n+1}) = \left(1 - \frac{\xi_1^2 + \cdots + \xi_n^2}{\xi_{n+1}^2}\right)_+^\lambda, \qquad \lambda > 0,$$

where the plus sign indicates that $m_\lambda = 0$ if the expression inside the parentheses is negative. The multiplier m_λ is called the *cone multiplier with parameter* λ. If m_λ is in $\mathscr{M}_p(\mathbf{R}^{n+1})$, then the function $b_\lambda(\xi) = (1 - |\xi|^2)_+^\lambda$ defined on $\mathbf{R}^n$ is in $\mathscr{M}_p(\mathbf{R}^n)$. Indeed, by Theorem 2.5.16 we have that for some $\xi_{n+1} = h$, $b_\lambda(\xi_1/h,\dots,\xi_n/h)$ is in $\mathscr{M}_p(\mathbf{R}^n)$ and hence so is b_λ by property (2.5.19).

Exercises

2.5.1. Prove that if $f \in L^q(\mathbf{R}^n)$ and $0 < q < \infty$, then

$$\big\|\tau^h(f) + f\big\|_{L^q} \to 2^{1/q}\big\|f\big\|_{L^q} \qquad \text{as } |h| \to \infty.$$

2.5.2. Prove Proposition 2.5.14. Also prove that if $\delta_j^{h_j}$ is a dilation operator in the jth variable (for instance $\delta_1^{h_1} f(x) = f(h_1x_1, x_2, \dots, x_n)$), then

$$\big\|\delta_1^{h_1} \cdots \delta_n^{h_n} m\big\|_{\mathscr{M}_p} = \big\|m\big\|_{\mathscr{M}_p}.$$

2.5.3. Let $m \in \mathscr{M}_p(\mathbf{R}^n)$ where $1 \le p < \infty$.
(a) If ψ is a function on $\mathbf{R}^n$ whose inverse Fourier transform is an integrable function, then prove that

$$\big\|\psi m\big\|_{\mathscr{M}_p} \le \big\|\psi^\vee\big\|_{L^1}\big\|m\big\|_{\mathscr{M}_p}.$$

(b) If ψ is in $L^1(\mathbf{R}^n)$, then prove that

$$\big\|\psi * m\big\|_{\mathscr{M}_p} \le \big\|\psi\big\|_{L^1}\big\|m\big\|_{\mathscr{M}_p}.$$

2.5.4. Fix a multi-index γ.
(a) Prove that the map $T(f) = f * \partial^\gamma \delta_0$ maps $\mathscr{S}$ continuously into $\mathscr{S}$.
(b) Prove that when $1/p - 1/q \neq |\gamma|/n$, T does not extend to an element of the space $\mathscr{M}^{p,q}$.

2.5.5. Let $K_\gamma(x) = |x|^{-n+\gamma}$, where $0 < \gamma < n$. Use Theorem 1.4.25 to show that the operator

$$T_\gamma(f) = f * K_\gamma, \qquad f \in \mathscr{S},$$

extends to a bounded operator in $\mathscr{M}^{p,q}(\mathbf{R}^n)$, where $1/p - 1/q = \gamma/n$, $1 < p < q < \infty$. This provides an example of a nontrivial operator in $\mathscr{M}^{p,q}(\mathbf{R}^n)$ when $p < q$.

2.5.6. (a) Use the ideas of the proof of Proposition 2.5.13 to show that if $m_j \in \mathscr{M}_p$, $1 \le p < \infty$, $\|m_j\|_{\mathscr{M}_p} \le C$ for all $j = 1, 2, \dots$, and $m_j \to m$ a.e., then $m \in \mathscr{M}_p$ and

$$\big\|m\big\|_{\mathscr{M}_p(\mathbf{R}^n)} \le \liminf_{j\to\infty} \big\|m_j\big\|_{\mathscr{M}_p(\mathbf{R}^n)} \le C.$$

(b) Prove that if $m \in \mathscr{M}_p$, $1 \le p < \infty$, and the limit $m_0(\xi) = \lim\limits_{R\to\infty} m(\xi/R)$ exists for all $\xi \in \mathbf{R}^n$, then m_0 is a radial function in $\mathscr{M}_p(\mathbf{R}^n)$ and satisfies $\|m_0\|_{\mathscr{M}_p} \le \|m\|_{\mathscr{M}_p}$.
(c) If $m \in \mathscr{M}_p(\mathbf{R})$ has left and right limits at the origin, then prove that

$$\big\|m\big\|_{\mathscr{M}_p(\mathbf{R})} \ge \max(|m(0+)|, |m(0-)|).$$

(d) Suppose that for some $1 \le p < \infty$, $m_t \in \mathscr{M}_p(\mathbf{R}^n)$ for all $0 < t < \infty$. Prove that

$$\int_0^\infty \|m_t\|_{\mathscr{M}_p(\mathbf{R}^n)} \frac{dt}{t} < \infty \implies m(\xi) = \int_0^\infty m_t(\xi) \frac{dt}{t} \in \mathscr{M}_p .$$

2.5.7. Let $1 \le p < \infty$ and suppose that $m \in \mathscr{M}_p(\mathbf{R}^n)$ satisfies $|m(\xi)| \ge c(1+|\xi|)^{-N}$ for some $c, N > 0$. Prove that the operator $T(f) = (\widehat{f} m^{-1})^\vee$ satisfies $\|T(f)\|_{L^p} \ge c_p \|f\|_{L^p}$ for all $f \in \mathscr{S}(\mathbf{R}^n)$, where $c_p = \|m\|_{\mathscr{M}_p}^{-1}$.

2.5.8. (a) Prove that if $m \in L^\infty(\mathbf{R}^n)$ satisfies $m^\vee \ge 0$, then for all $1 \le p < \infty$ we have

$$\|m\|_{\mathscr{M}_p} = \|m^\vee\|_{L^1} .$$

(b) (*L. Colzani and E. Laeng*) On the real line let

$$m_1(\xi) = \begin{cases} -1 & \text{for } \xi > 0 \\ 1 & \text{for } \xi < 0, \end{cases} \qquad m_2(\xi) = \begin{cases} \min(\xi - 1, 0) & \text{for } \xi > 0 \\ \max(\xi + 1, 0) & \text{for } \xi < 0. \end{cases}$$

Prove that

$$\|m_1\|_{\mathscr{M}_p} = \|m_2\|_{\mathscr{M}_p}$$

for all $1 < p < \infty$.
$\big[$*Hint:* Part (a): Use Exercise 1.2.9. Part (b): Use part (a) to show that $\|m_2 m_1^{-1}\|_{\mathscr{M}_p} = 1$. Deduce that $\|m_2\|_{\mathscr{M}_p} \le \|m_1\|_{\mathscr{M}_p}$. For the converse use Exercise 2.5.6 (c).$\big]$

2.5.9. ([94]) Let $1 < p < \infty$ and $0 < A < \infty$. Prove that the following are equivalent:
(a) The operator $f \mapsto \sum_{m \in \mathbf{Z}^n} a_m f(x-m)$ is bounded on $L^p(\mathbf{R}^n)$ with norm A.
(b) The $\mathscr{M}_p$ norm of the function $\sum_{m \in \mathbf{Z}^n} a_m e^{-2\pi i m \cdot x}$ is exactly A.
(c) The operator given by convolution with the sequence $\{a_m\}$ is bounded on $\ell^p(\mathbf{Z}^n)$ with norm A.

2.5.10. ([177]) Let $m(\xi)$ in $\mathscr{M}_p(\mathbf{R}^n)$ be supported in $[0,1]^n$. Then the periodic extension of m in $\mathbf{R}^n$,

$$M(\xi) = \sum_{k \in \mathbf{Z}^n} m(\xi - k),$$

is also in $\mathscr{M}_p(\mathbf{R}^n)$.

2.5.11. Suppose that u is a $\mathscr{C}^\infty$ function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree $-n + i\tau$, $\tau \in \mathbf{R}$. Prove that the operator given by convolution with u maps $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n)$.

2.5.12. ([142]) Let $m_1 \in L^r(\mathbf{R}^n)$ and $m_2 \in L^{r'}(\mathbf{R}^n)$ for some $2 \le r \le \infty$. Prove that $m_1 * m_2 \in \mathscr{M}_p(\mathbf{R}^n)$ when $\frac{1}{p} - \frac{1}{2} = \frac{1}{r}$ and $1 \le p \le 2$.
$\big[$*Hint:* Prove that the trilinear operator $(m_1, m_2, f) \mapsto \big((m_1 * m_2)\widehat{f}\,\big)^\vee$ is bounded from $L^2 \times L^2 \times L^1 \to L^1$ and $L^\infty \times L^1 \times L^2 \to L^2$. Apply trilinear complex interpolation (Corollary 7.2.11 in [131]) to deduce the required conclusion for $1 \le p \le 2$.$\big]$

2.5.13. Show that the function $e^{i|\xi|^2}$ is an L^p Fourier multiplier on $\mathbf{R}^n$ if and only if $p=2$.
$\big[$*Hint:* By Exercise 2.4.12 the inverse Fourier transform of $e^{i|\xi|^2}$ is in L^∞, thus the operator $f \mapsto \big(\widehat{f}(\xi)e^{i\pi|\xi|^2}\big)^\vee$ maps L^1 to L^∞. Since this operator also maps L^2 to L^2, it should map L^p to $L^{p'}$ for all $1 \le p \le 2$.$\big]$

2.6 Oscillatory Integrals

Oscillatory integrals have played an important role in harmonic analysis from its outset. The Fourier transform is the prototype of oscillatory integrals and provides the simplest example of a nontrivial phase, a linear function of the variable of integration. More complicated phases naturally appear in the subject; for instance, Bessel functions provide examples of oscillatory integrals in which the phase is a sinusoidal function.

In this section we take a quick look at oscillatory integrals. We mostly concentrate on one-dimensional results, which already require some significant analysis. We examine only a very simple higher-dimensional situation. Our analysis here is far from adequate.

Definition 2.6.1. An *oscillatory integral* is an expression of the form

$$I(\lambda) = \int_{\mathbf{R}^n} e^{i\lambda\varphi(x)}\,\psi(x)\,dx\,, \tag{2.6.1}$$

where λ is a positive real number, φ is a real-valued function on $\mathbf{R}^n$ called the *phase*, and ψ is a complex-valued and smooth integrable function on $\mathbf{R}^n$, which is often taken to have compact support.

2.6.1 Phases with No Critical Points

We begin by studying the simplest possible one-dimensional case. Suppose that φ and ψ are smooth functions on the real line such that $\operatorname{supp}\psi$ is a closed interval and

$$\varphi'(x) \neq 0 \qquad \text{for all } x \in \operatorname{supp}\psi.$$

Since φ' has no zeros, it must be either strictly positive or strictly negative everywhere on the support of ψ. It follows that φ is monotonic on the support of ψ and we are allowed to change variables

$$u = \varphi(x)$$

in (2.6.1). Then $dx = (\varphi'(x))^{-1}du = (\varphi^{-1})'(u)\,du$, where φ^{-1} is the inverse function of φ. We transform the integral in (2.6.1) into

$$\int_{\mathbf{R}} e^{i\lambda u}\psi(\varphi^{-1}(u))(\varphi^{-1})'(u)\,du \tag{2.6.2}$$

and we note that the function $\theta(u) = \psi(\varphi^{-1}(u))(\varphi^{-1})'(u)$ is smooth and has compact support on $\mathbf{R}$. We therefore interpret the integral in (2.6.1) as $\widehat{\theta}(-\lambda/2\pi)$, where $\widehat{\theta}$ is the Fourier transform of θ. Since θ is a smooth function with compact support, it follows that the integral in (2.6.2) has rapid decay as $\lambda \to \infty$.

A quick way to see that the expression $\widehat{\theta}(-\lambda/2\pi)$ has decay of order λ^{-N} for all $N > 0$ as λ tends to ∞ is the following. Write

$$e^{i\lambda u} = \frac{1}{(i\lambda)^N}\frac{d^N}{du^N}(e^{i\lambda u})$$

and integrate by parts N times to express the integral in (2.6.2) as

$$\frac{(-1)^N}{(i\lambda)^N}\int_{\mathbf{R}} e^{i\lambda u}\frac{d^N\theta(u)}{du^N}\,du\,,$$

from which the assertion follows. Hence

$$|I(\lambda)| = |\widehat{\theta}(-\lambda/2\pi)| \le C_N\lambda^{-N}\,, \tag{2.6.3}$$

where $C_N = \|\theta^{(N)}\|_{L^1}$, which depends on derivatives of φ and ψ.

We now turn to a higher-dimensional analogue of this situation.

Definition 2.6.2. We say that a point x_0 is a *critical point* of a phase function φ if

$$\nabla\varphi(x_0) = \big(\partial_1\varphi(x_0),\dots,\partial_n\varphi(x_0)\big) = 0\,.$$

Example 2.6.3. Let $\xi \in \mathbf{R}^n \setminus \{0\}$. Then the phase functions $\varphi_1(x) = x\cdot\xi$, $\varphi_2(x) = e^{x\cdot\xi}$ have no critical points, while the phase function $\varphi_3(x) = |x|^2 - x\cdot\xi$ has one critical point at $x_0 = \frac{1}{2}\xi$.

The next result concerns the behavior of oscillatory integrals whose phase functions have no critical points.

Proposition 2.6.4. *Suppose that ψ is a compactly supported smooth function on $\mathbf{R}^n$ and that φ is a real-valued $\mathscr{C}^\infty$ function on $\mathbf{R}^n$ that has no critical points on the support of ψ. Then the oscillatory integral*

$$I(\lambda) = \int_{\mathbf{R}^n} e^{i\lambda\varphi(x)}\psi(x)\,dx \tag{2.6.4}$$

obeys a bound of the form $|I(\lambda)| \le C_N\lambda^{-N}$ for all $\lambda \ge 1$ and all $N > 0$, where C_N depends on N and on φ and ψ.

Proof. Since the case $n = 1$ has already been discussed, we concentrate on dimensions $n \geq 2$. For each y in the support of ψ there is a unit vector θ_y such that

$$\theta_y \cdot \nabla\varphi(y) = |\nabla\varphi(y)|.$$

By the continuity of $\nabla\varphi$ there is a small neighborhood $B(y, r_y)$ of y such that for all $x \in B(y, r_y)$ we have

$$\theta_y \cdot \nabla\varphi(x) \geq \frac{1}{2}|\nabla\varphi(y)| > 0.$$

Cover the support of ψ by a finite number of balls $B(y_j, r_{y_j})$, $j = 1, \dots, m$, and pick $c = \min_j \frac{1}{2}|\nabla\varphi(y_j)|$; we have

$$\theta_{y_j} \cdot \nabla\varphi(x) \geq c > 0 \tag{2.6.5}$$

for all $x \in B(y_j, r_{y_j})$ and $j = 1, \dots, m$.

Next we find a smooth partition of unity of $\mathbf{R}^n$ such that each member ζ_k of the partition is supported in some ball $B(y_j, r_{y_j})$ or lies outside the support of ψ. We therefore write

$$I(\lambda) = \sum_k \int_{\mathbf{R}^n} e^{i\lambda\varphi(x)}\psi(x)\zeta_k(x)\,dx, \tag{2.6.6}$$

where the sum contains only a finite number of indices, since only a finite number of the ζ_k's meet the support of ψ. It suffices to show that every term in the sum in (2.6.6) has rapid decay in λ as $\lambda \to \infty$.

To this end, we fix a k and we pick a j such that the support of $\psi\zeta_k$ is contained in some ball $B(y_j, r_{y_j})$. We find unit vectors $\theta_{y_j,2}, \dots, \theta_{y_j,n}$, such that the system $\{\theta_{y_j}, \theta_{y_j,2}, \dots, \theta_{y_j,n}\}$ is an orthonormal basis of $\mathbf{R}^n$. Let e_j be the unit (column) vector on $\mathbf{R}^n$ whose jth coordinate is one and whose remaining coordinates are zero. We find an orthogonal matrix R such that $R^t e_1 = \theta_{y_j}$ and we introduce the change of variables $u = y_j + R(x - y_j)$ in the integral

$$I_k(\lambda) = \int_{\mathbf{R}^n} e^{i\lambda\varphi(x)}\psi(x)\zeta_k(x)\,dx.$$

The map $x \mapsto u = (u_1, \dots, u_n)$ is a rotation that fixes y_j and preserves the ball $B(y_j, r_{y_j})$. Defining $\varphi(x) = \varphi^o(u)$, $\psi(x) = \psi^o(u)$, $\zeta_k(x) = \zeta_k^o(u)$, under this new coordinate system we write

$$I_k(\lambda) = \int_K \left\{ \int_{\mathbf{R}} e^{i\lambda\varphi^o(u)}\psi^o(u_1, \dots, u_n)\zeta_k^o(u_1, \dots, u_n)\,du_1 \right\} du_2 \cdots du_n, \tag{2.6.7}$$

where K is a compact subset of $\mathbf{R}^{n-1}$. Since R is an orthogonal matrix, $R^{-1} = R^t$, and the change of variables $x = y_j + R^t(u - y_j)$ implies that

$$\frac{\partial x}{\partial u_1} = \text{first column of } R^t = \text{first row of } R = R^t e_1 = \theta_{y_j}.$$

Thus for all $x \in B(y_j, r_j)$ we have

$$\frac{\partial \varphi^o(u)}{\partial u_1} = \frac{\partial \varphi(y_j + R^t(u - y_j))}{\partial u_1} = \nabla \varphi(x) \cdot \frac{\partial x}{\partial u_1} = \nabla \varphi(x) \cdot \theta_{y_j} \geq c > 0$$

in view of condition (2.6.5). This lower estimate is valid for all $u \in B(y_j, r_{y_j})$, and therefore the inner integral inside the curly brackets in (2.6.7) is at most $C_N \lambda^{-N}$ by estimate (2.6.3). Integrating over K results in the same conclusion for $I(\lambda)$ defined in (2.6.4). □

2.6.2 Sublevel Set Estimates and the Van der Corput Lemma

We discuss a sharp decay estimate for one-dimensional oscillatory integrals. This estimate is obtained as a consequence of delicate size estimates for the Lebesgue measures of the sublevel sets $\{|u| \leq \alpha\}$ for a function u. In what follows, $u^{(k)}$ denotes the kth derivative of a function $u(t)$ defined on $\mathbf{R}$, and $\mathscr{C}^k$ the space of all functions whose kth derivative exists and is continuous.

Lemma 2.6.5. *Let $k \geq 1$ and suppose that $a_0, \dots, a_k$ are distinct real numbers. Let $a = \min(a_j)$ and $b = \max(a_j)$ and let f be a real-valued $\mathscr{C}^{k-1}$ function on $[a,b]$ that is $\mathscr{C}^k$ on (a,b). Then there exists a point y in (a,b) such that*

$$\sum_{m=0}^{k} c_m f(a_m) = f^{(k)}(y)\,,$$

where $c_m = (-1)^k k! \prod_{\substack{\ell=0 \\ \ell \neq m}}^{k} (a_\ell - a_m)^{-1}$.

Proof. Suppose we could find a polynomial $p_k(x) = \sum_{j=0}^{k} b_j x^j$ such that the function

$$\varphi(x) = f(x) - p_k(x) \tag{2.6.8}$$

satisfies $\varphi(a_m) = 0$ for all $0 \leq m \leq k$. Since the a_j are distinct, we apply Rolle's theorem k times to find a point y in (a,b) such that

$$f^{(k)}(y) = k!\, b_k\,.$$

The existence of a polynomial p_k such that (2.6.8) is satisfied is equivalent to the existence of a solution to the matrix equation

$$\begin{pmatrix} a_0^k & a_0^{k-1} & \dots & a_0 & 1 \\ a_1^k & a_1^{k-1} & \dots & a_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k-1}^k & a_{k-1}^{k-1} & \dots & a_{k-1} & 1 \\ a_k^k & a_k^{k-1} & \dots & a_k & 1 \end{pmatrix} \begin{pmatrix} b_k \\ b_{k-1} \\ \vdots \\ b_1 \\ b_0 \end{pmatrix} = \begin{pmatrix} f(a_0) \\ f(a_1) \\ \vdots \\ f(a_{k-1}) \\ f(a_k) \end{pmatrix}.$$

The determinant of the square matrix on the left is called the *Vandermonde determinant* and is equal to

$$\prod_{\ell=0}^{k-1} \prod_{j=\ell+1}^{k} (a_\ell - a_j) \neq 0.$$

Since the a_j are distinct, it follows that the system has a unique solution. Using Cramer's rule, we solve this system to obtain

$$\begin{aligned} b_k &= \sum_{m=0}^{k} (-1)^m f(a_m) \frac{\prod\limits_{\substack{\ell=0 \\ \ell\neq m}}^{k-1} \prod\limits_{\substack{j=\ell+1 \\ j\neq m}}^{k} (a_\ell - a_j)}{\prod\limits_{\ell=0}^{k-1} \prod\limits_{j=\ell+1}^{k} (a_\ell - a_j)} \\ &= \sum_{m=0}^{k} (-1)^m f(a_m) \prod_{\substack{\ell=0 \\ \ell\neq m}}^{k} (a_\ell - a_m)^{-1} (-1)^{k-m}. \end{aligned}$$

The required conclusion now follows with c_m as claimed. □

Lemma 2.6.6. *Let E be a measurable subset of $\mathbf{R}$ with finite nonzero Lebesgue measure and let $k \in \mathbf{Z}^+$. Then there exist $a_0, \dots, a_k$ in E such that for all $\ell = 0, 1, \dots, k$ we have*

$$\prod_{\substack{j=0 \\ j\neq \ell}}^{k} |a_j - a_\ell| \geq (|E|/2e)^k. \tag{2.6.9}$$

Proof. Given a measurable set E with finite measure, pick a compact subset E' of E such that $|E \setminus E'| < \delta$, for some $\delta > 0$. For $x \in \mathbf{R}$ define $T(x) = |(-\infty, x) \cap E'|$. Then T enjoys the distance-decreasing property

$$|T(x) - T(y)| \leq |x - y|$$

for all $x, y \in E'$; consequently, by the intermediate value theorem, T is a surjective map from E' to $[0, |E'|]$. Let a_j be points in E' such that $T(a_j) = \frac{j}{k}|E'|$ for $j = 0, 1, \dots, k$. For k an even integer, we have

$$\prod_{\substack{j=0 \\ j\neq \ell}}^{k} |a_j - a_\ell| \geq \prod_{\substack{j=0 \\ j\neq \ell}}^{k} \left| \frac{j}{k}|E'| - \frac{\ell}{k}|E'| \right| \geq \prod_{\substack{j=0 \\ j\neq \frac{k}{2}}}^{k} \left| \frac{j}{k} - \frac{1}{2} \right| |E'|^k = \prod_{r=0}^{\frac{k}{2}-1} \left(\frac{r - \frac{k}{2}}{k} \right)^2 |E'|^k,$$

and it is easily shown that $\big((k/2)!\big)^2 k^{-k} \geq (2e)^{-k}$.

For k an odd integer we have

$$\prod_{\substack{j=0\\ j\neq\ell}}^{k}|a_j-a_\ell|\geq\prod_{\substack{j=0\\ j\neq\ell}}^{k}\left|\frac{j}{k}|E'|-\frac{\ell}{k}|E'|\right|\geq\prod_{\substack{j=0\\ j\neq\frac{k+1}{2}}}^{k}\left|\frac{j}{k}-\frac{k+1}{2k}\right||E'|^k,$$

while the last product is at least

$$\left\{\frac{1}{k}\cdot\frac{2}{k}\cdots\frac{\frac{k-1}{2}}{k}\right\}^2\frac{k+1}{2k}\geq(2e)^{-k}.$$

We have therefore proved (2.6.9) with E' replacing E. Since $|E\setminus E'|<\delta$ and $\delta>0$ is arbitrarily small, the required conclusion follows. $\square$

The following is the main result of this section.

Proposition 2.6.7. *(a) Let u be a real-valued $\mathscr{C}^k$ function, $k\in\mathbf{Z}^+$, that satisfies $u^{(k)}(t)\geq 1$ for all $t\in\mathbf{R}$. Then the following estimate is valid for all $\alpha>0$:*

$$\left|\left\{t\in\mathbf{R}:\ |u(t)|\leq\alpha\right\}\right|\leq(2e)((k+1)!)^{\frac{1}{k}}\alpha^{\frac{1}{k}}.\tag{2.6.10}$$

(b) Let $-\infty<a<b<\infty$. For all $k\geq 2$, for every real-valued $\mathscr{C}^k$ function u on the line that satisfies $u^{(k)}(t)\geq 1$ for all $t\in[a,b]$, and every $\lambda\in\mathbf{R}\setminus\{0\}$ we have:

$$\left|\int_a^b e^{i\lambda u(t)}\,dt\right|\leq 12k|\lambda|^{-\frac{1}{k}}.\tag{2.6.11}$$

(c) If $k=1$, $u'(t)$ is monotonic on (a,b), and $u'(t)\geq 1$ for all $t\in(a,b)$, then for all nonzero real numbers λ we have

$$\left|\int_a^b e^{i\lambda u(t)}\,dt\right|\leq 3|\lambda|^{-1}.\tag{2.6.12}$$

Proof. Part (a): Let $E=\{t\in\mathbf{R}:\ |u(t)|\leq\alpha\}$. If $|E|$ is nonzero, then by Lemma 2.6.6 there exist $a_0,a_1,\dots,a_k$ in E such that for all ℓ we have

$$|E|^k\leq(2e)^k\prod_{\substack{j=0\\ j\neq\ell}}^{k}|a_j-a_\ell|.\tag{2.6.13}$$

Lemma 2.6.5 implies that there exists $y\in\left(\min a_j,\max a_j\right)$ such that

$$u^{(k)}(y)=(-1)^k k!\sum_{m=0}^{k}u(a_m)\prod_{\substack{\ell=0\\ \ell\neq m}}^{k}(a_\ell-a_m)^{-1}.\tag{2.6.14}$$

Using (2.6.13), we obtain that the expression on the right in (2.6.14) is in absolute value at most

$$(k+1)!\max_{0\le j\le k}|u(a_j)|\,(2e)^k\,|E|^{-k}\le (k+1)!\,\alpha\,(2e)^k\,|E|^{-k},$$

since $a_j\in E$. The bound $u^{(k)}(t)\ge 1$ now implies

$$|E|^k\le (k+1)!\,(2e)^k\,\alpha$$

as claimed. This proves (2.6.10).

Part (b): We now take $k\ge 2$ and we split the interval (a,b) in (2.6.11) into the sets

$$\begin{aligned} R_1 &= \{t\in(a,b):\ |u'(t)|\le\beta\},\\ R_2 &= \{t\in(a,b):\ |u'(t)|>\beta\}, \end{aligned}$$

for some parameter β to be chosen momentarily. The function $v=u'$ satisfies $v^{(k-1)}\ge 1$ and $k-1\ge 1$. It follows from part (a) that

$$\left|\int_{R_1} e^{i\lambda u(t)}\,dt\right|\le |R_1|\le 2e\,(k!)^{\frac{1}{k-1}}\,\beta^{\frac{1}{k-1}}\le 6k\,\beta^{\frac{1}{k-1}}\,.$$

To obtain the corresponding estimate over R_2, we note that if $u^{(k)}\ge 1$, then the set $\{|u'|>\beta\}$ is the union of at most $2k-2$ intervals on each of which u' is monotone. Let (c,d) be one of these intervals on which u' is monotone. Then u' has a fixed sign on (c,d) and we have

$$\begin{aligned} \left|\int_c^d e^{i\lambda u(t)}\,dt\right| &= \left|\int_c^d \big(e^{i\lambda u(t)}\big)'\frac{1}{\lambda u'(t)}\,dt\right|\\ &\le \left|\int_c^d e^{i\lambda u(t)}\Big(\frac{1}{\lambda u'(t)}\Big)'\,dt\right|+\frac{1}{|\lambda|}\left|\frac{e^{i\lambda u(d)}}{u'(d)}-\frac{e^{i\lambda u(c)}}{u'(c)}\right|\\ &\le \frac{1}{|\lambda|}\int_c^d\left|\Big(\frac{1}{u'(t)}\Big)'\right|dt+\frac{2}{|\lambda|\beta}\\ &= \frac{1}{|\lambda|}\left|\int_c^d\Big(\frac{1}{u'(t)}\Big)'\,dt\right|+\frac{2}{|\lambda|\beta}\\ &= \frac{1}{|\lambda|}\left|\frac{1}{u'(d)}-\frac{1}{u'(c)}\right|+\frac{2}{|\lambda|\beta}\\ &\le \frac{3}{|\lambda|\beta}, \end{aligned}$$

where we use the monotonicity of $1/u'(t)$ in moving the absolute value from inside the integral to outside. It follows that

$$\left|\int_{R_2} e^{i\lambda u(t)}\,dt\right|\le\frac{6k}{|\lambda|\beta}\,.$$

Choosing $\beta = |\lambda|^{-(k-1)/k}$ to optimize and adding the corresponding estimates for R_1 and R_2, we deduce the claimed estimate (2.6.11).

Part (c): Repeat the argument in part (b) setting $\beta = 1$ and replacing the interval (c,d) by (a,b). □

Corollary 2.6.8. *Let (a,b), $u(t)$, $\lambda > 0$, and k be as in Proposition 2.6.7. Then for any function ψ on (a,b) with an integrable derivative and $k \geq 2$, we have*

$$\left| \int_a^b e^{i\lambda u(t)} \psi(t)\,dt \right| \leq 12k\lambda^{-1/k} \left[|\psi(b)| + \int_a^b |\psi'(s)|\,ds \right].$$

We also have

$$\left| \int_a^b e^{i\lambda u(t)} \psi(t)\,dt \right| \leq 3\lambda^{-1} \left[|\psi(b)| + \int_a^b |\psi'(s)|\,ds \right],$$

when $k = 1$ and u' is monotonic on (a,b).

Proof. Set

$$F(x) = \int_a^x e^{i\lambda u(t)}\,dt$$

and use integration by parts to write

$$\int_a^b e^{i\lambda u(t)} \psi(t)\,dt = F(b)\psi(b) - \int_a^b F(t)\psi'(t)\,dt.$$

The conclusion easily follows. □

Example 2.6.9. The *Bessel function* of order m is defined as

$$J_m(r) = \frac{1}{2\pi} \int_0^{2\pi} e^{ir\sin\theta} e^{-im\theta}\,d\theta.$$

Here we take both r and m to be real numbers, and we suppose that $m > -\frac{1}{2}$; we refer to Appendix B for an introduction to Bessel functions and their basic properties.

We use Corollary 2.6.8 to calculate the decay of the Bessel function $J_m(r)$ as $r \to \infty$. Set

$$\varphi(\theta) = \sin(\theta)$$

and note that $\varphi'(\theta)$ vanishes only at $\theta = \pi/2$ and $3\pi/2$ inside the interval $[0, 2\pi]$ and that $\varphi''(\pi/2) = -1$, while $\varphi''(3\pi/2) = 1$. We now write $1 = \psi_1 + \psi_2 + \psi_3$, where ψ_1 is smooth and compactly supported in a small neighborhood of $\pi/2$, and ψ_2 is smooth and compactly supported in a small neighborhood of $3\pi/2$. For $j = 1,2$, Corollary 2.6.8 yields

$$\left| \int_0^{2\pi} e^{ir\sin(\theta)} \big(\psi_j(\theta) e^{-im\theta}\big)\,d\theta \right| \leq Cmr^{-1/2}$$

for some constant C, while the corresponding integral containing ψ_3 has arbitrary decay in r in view of estimate (2.6.3) (or Proposition 2.6.4 when $n=1$).

Exercises

2.6.1. Suppose that u is a real-valued $\mathscr{C}^k$ function defined on the line that satisfies $|u^{(k)}(t)| \geq c_0 > 0$ for some $k \geq 2$ and all $t \in (a,b)$. Prove that for $\lambda \in \mathbf{R} \setminus \{0\}$ we have

$$\left| \int_a^b e^{i\lambda u(t)}\, dt \right| \leq 12k\,(\lambda c_0)^{-1/k}$$

and that the same conclusion is valid when $k=1$, provided u' is monotonic.

2.6.2. Show that if u' is not monotonic in part (c) of Proposition 2.6.7, then the conclusion may fail.
$\big[$*Hint:* Let $\varphi(t)$ be a real-valued smooth function that is equal to $2t$ on intervals $[2\pi k+\varepsilon_k, 2\pi(k+\frac{1}{2})-\varepsilon_k]$ and equal to t on intervals $[2\pi(k+\frac{1}{2})+\varepsilon_k, 2\pi(k+1)-\varepsilon_k]$, where $0 \leq k \leq N$, for some $N \in \mathbf{Z}^+$. Show that the absolute value of the integral of $e^{i\varphi(t)}$ over the interval $[\varepsilon_0, 2\pi(N+1)-\varepsilon_N]$ tends to infinity as $N \to \infty$.$\big]$

2.6.3. Prove that the dependence on k of the constant in part (b) of Proposition 2.6.7 is indeed linear.
$\big[$*Hint:* Take $u(t) = t^k/k!$ over the interval $(0,k!)$.$\big]$

2.6.4. Follow the steps below to give an alternative proof of part (b) of Proposition 2.6.7. Assume that the statement is known for some $k \geq 2$ and some constant $C(k)$ for all intervals $[a,b]$ and all $\mathscr{C}^k$ functions satisfying $u^{(k)} \geq 1$ on $[a,b]$. Fix a $\mathscr{C}^{k+1}$ function u such that $u^{(k+1)} \geq 1$ on an interval $[a,b]$. Let c be the unique point at which the function $u^{(k)}$ attains its minimum in $[a,b]$.
(a) If $u^{(k)}(c) = 0$, then for all $\delta > 0$ we have $u^{(k)}(t) \geq \delta$ in the complement of the interval $(c-\delta, c+\delta)$ and derive the bound

$$\left| \int_a^b e^{i\lambda u(t)} dt \right| \leq 2C(k)(\lambda\delta)^{-1/k} + 2\delta\,.$$

(b) If $u^{(k)}(c) \neq 0$, then we must have $c \in \{a,b\}$. Obtain the bound

$$\left| \int_a^b e^{i\lambda u(t)} dt \right| \leq C(k)(\lambda\delta)^{-1/k} + \delta\,.$$

(c) Choose a suitable δ to optimize and deduce the validity of the statement for $k+1$ with $C(k+1) = 2C(k)+2$, hence $C(k) = 3\cdot 2^{k-1} + 2^k - 2$, since $C(1)=3$.

2.6.5. (a) Prove that for some constant C and all $\lambda \in \mathbf{R}$ and $\varepsilon \in (0,1)$ we have

$$\left| \int_{\varepsilon \leq |t| \leq 1} e^{i\lambda t} \frac{dt}{t} \right| \leq C\,.$$

(b) Prove that for some $C' < \infty$, all $\lambda \in \mathbf{R}$, $k > 0$, and $\varepsilon \in (0,1)$ we have

$$\left| \int_{\varepsilon \le |t| \le 1} e^{i\lambda t \pm t^k} \frac{dt}{t} \right| \le C' .$$

(c) Show that there is a constant C'' such that for any $0 < \varepsilon < N < \infty$, for all ξ_1, ξ_2 in $\mathbf{R}$, and for all integers $k \ge 2$, we have

$$\left| \int_{\varepsilon \le |s| \le N} e^{i(\xi_1 s + \xi_2 s^k)} \frac{ds}{s} \right| \le C'' .$$

[*Hint:* Part (a): For $|\lambda|$ small use the inequality $|e^{i\lambda t} - 1| \le |\lambda t|$. If $|\lambda|$ is large, split the domains of integration into the regions $|t| \le |\lambda|^{-1}$ and $|t| \ge |\lambda|^{-1}$ and use integration by parts in the second case. Part (b): Write

$$\frac{e^{i(\lambda t \pm t^k)} - 1}{t} = e^{i\lambda t} \frac{e^{\pm i t^k} - 1}{t} + \frac{e^{i\lambda t}}{t}$$

and use part (a). Part (c): When $\xi_1 = \xi_2 = 0$ it is trivial. If $\xi_2 = 0$, $\xi_1 \ne 0$, change variables $t = \xi_1 s$ and then split the domain of integration into the sets $|t| \le 1$ and $|t| \ge 1$. In the interval over the set $|t| \le 1$ apply part (b) and over the set $|t| \ge 1$ use integration by parts. In the case $\xi_2 \ne 0$, change variables $t = |\xi_2|^{1/k} s$ and split the domain of integration into the sets $|t| \ge 1$ and $|t| \le 1$. When $|t| \le 1$ use part (b) and in the case $|t| \ge 1$ use Corollary 2.6.8, noting that $\frac{d^k(\xi_1 |\xi_2|^{-1/k} t \pm t^k)}{dt} = k! \ge 1$.]

2.6.6. (a) Show that for all $a > 0$ and $\lambda > 0$ the following is valid:

$$\left| \int_0^{a\lambda} e^{i\lambda \log t} \, dt \right| \le a .$$

(b) Prove that there is a constant $c > 0$ such that for all $b > \lambda > 10$ we have

$$\left| \int_0^b e^{i\lambda t \log t} \, dt \right| \le \frac{c}{\lambda \log \lambda} .$$

[*Hint:* Part (b): Consider the intervals $(0, \delta)$ and $[\delta, b)$ for some δ. Apply Proposition 2.6.7 with $k = 1$ on one of these intervals and with $k = 2$ on the other. Then choose a suitable δ.]

2.6.7. Show that there is a constant $C < \infty$ such that for all nonintegers $\gamma > 1$ and all $\lambda, b > 1$ we have

$$\left| \int_0^b e^{i\lambda t^\gamma} dt \right| \le \frac{C}{\lambda^\gamma} .$$

[*Hint:* On the interval $(0, \delta)$ apply Proposition 2.6.7 with $k = [\gamma] + 1$ and on the interval (δ, b) with $k = [\gamma]$. Then optimize by choosing $\delta = \lambda^{-1/\gamma}$.]

HISTORICAL NOTES

The one-dimensional maximal function originated in the work of Hardy and Littlewood [146]. Its n-dimensional analogue was introduced by Wiener [375], who used Lemma 2.1.5, a variant of the Vitali covering lemma, to derive its L^p boundedness. One may consult the books of de Guzmán [92], [93] for extensions and other variants of such covering lemmas. The actual covering lemma proved by Vitali [368] says that if a family of closed cubes in $\mathbf{R}^n$ has the property that for every point $x \in A \subseteq \mathbf{R}^n$ there exists a sequence of cubes in the family that tends to x, then it is always possible to extract a sequence of pairwise disjoint cubes E_j from the family such that $|A \setminus \bigcup_j E_j| = 0$. We refer to Saks [310] for details and extensions of this theorem.

The class $L\log L$ was introduced by Zygmund to give a sufficient condition on the local integrability of the Hardy–Littlewood maximal operator. The necessity of this condition was observed by Stein [336]. Stein [341] also showed that the $L^p(\mathbf{R}^n)$ norm of the centered Hardy–Littlewood maximal operator $\mathcal{M}$ is bounded above by some dimension-free constant; see also Stein and Strömberg [345]. Analogous results for maximal operators associated with convex bodies are contained in Bourgain [35], Carbery [51], and Müller [263]. Bourgain [37] showed the the Hardy-Littlewood maximal operator associated with cubes is bounded on $L^p(\mathbf{R}^n)$ with dimension-free bounds when $p > 1$. Aldaz [2] studied the corresponding weak type $(1,1)$ bounds and proved that they grow to infinity with the dimension; the constant was improved by Aubrun [15]. The situation for the uncentered maximal operator M on L^p is different, since given any $1 < p < \infty$ there exists $C_p > 1$ such that $\|M\|_{L^p(\mathbf{R}^n)\to L^p(\mathbf{R}^n)} \geq C_p^n$ (see Exercise 2.1.8 for a value of such a constant C_p and also the article of Grafakos and Montgomery-Smith [136] for a larger value).

The centered maximal function $\mathcal{M}_\mu$ with respect to a general inner regular locally finite positive measure μ on $\mathbf{R}^n$ is bounded on $L^p(\mathbf{R}^n,\mu)$ without the additional hypothesis that the measure is doubling; see Fefferman [117]. The proof of this result requires the following covering lemma, obtained by Besicovitch [27]: Given any family of closed balls whose centers form a bounded subset of $\mathbf{R}^n$, there exists an at most countable subfamily of balls that covers the set of centers and has bounded overlap, i.e., no point in $\mathbf{R}^n$ belongs to more than a finite number (depending on the dimension) of the balls in the subfamily. A similar version of this lemma was obtained independently by Morse [258]. See also Ziemer [385] for an alternative formulation. The uncentered maximal operator M_μ of Exercise 2.1.1 may not be weak type $(1,1)$ if the measure μ is nondoubling, as shown by Sjögren [323]; related positive weak type $(1,1)$ results are contained in the article of Vargas [365]. The precise value of the operator norm of the uncentered Hardy–Littlewood maximal function on $L^p(\mathbf{R})$ was shown by Grafakos and Montgomery-Smith [136] to be the unique positive solution of the equation $(p-1)x^p - px^{p-1} - 1 = 0$. This constant raised to the power n is the operator norm of the strong maximal function M_s on $L^p(\mathbf{R}^n)$ for $1 < p \leq \infty$. The best weak type $(1,1)$ constant for the centered Hardy–Littlewood maximal operator was shown by Melas [248] to be the largest root of the quadratic equation $12x^2 - 22x + 5 = 0$. The strong maximal operator M_s is not weak type $(1,1)$, but it satisfies the substitute inequality $d_{M_s(f)}(\alpha) \leq C\int_{\mathbf{R}^n} \frac{|f(x)|}{\alpha}(1+\log^+\frac{|f(x)|}{\alpha})^{n-1}\,dx$. This result is due to Jessen, Marcinkiewicz, and Zygmund [176], but a geometric proof of it was obtained by Córdoba and Fefferman [73].

The basic facts about the Fourier transform go back to Fourier [119]. The theory of distributions was developed by Schwartz [314], [315]. For a concise introduction to the theory of distributions we refer to Hörmander [160] and Yosida [382]. Homogeneous distributions were considered by Riesz [295] in the study of the Cauchy problem in partial differential equations, although some earlier accounts are found in the work of Hadamard. They were later systematically studied by Gelfand and Šilov [126], [127]. References on the uncertainty principle include the articles of Fefferman [114] and Folland and Sitaram [118]. The best possible constant B_p in the Hausdorff–Young inequality $\|\widehat{f}\|_{L^{p'}(\mathbf{R}^n)} \leq B_p\|f\|_{L^p(\mathbf{R}^n)}$ when $1 \leq p \leq 2$ was shown by Beckner [21] to be $B_p = (p^{1/p}(p')^{-1/p'})^{n/2}$. This best constant was previously obtained by Babenko [16] in the case when p' is an even integer.

A nice treatise of the spaces $\mathscr{M}^{p,q}$ is found in Hörmander [159]. This reference also contains Theorem 2.5.6, which is due to him. Theorem 2.5.16 is due to de Leeuw [94], but the proof presented here is taken from Jodeit [178]. De Leeuw's result in Exercise 2.5.9 says that periodic elements of $\mathscr{M}_p(\mathbf{R}^n)$ can be isometrically identified with elements of $\mathscr{M}(\mathbf{T}^n)$, the latter being the space of all multipliers on $\ell^p(\mathbf{Z}^n)$. The hint in Exercise 2.5.13 was suggested by M. Peloso.

Parts (b) and (c) of Proposition 2.6.7 are due to van der Corput [364] and are referred to in the literature as van der Corput's lemma. The refinement in part (a) was subsequently obtained by Arhipov, Karachuba, and Čubarikov [8]. The treatment of these results in the text is based on the article of Carbery, Christ, and Wright [53], which also investigates higher-dimensional analogues of the theory. Precise asymptotics can be obtained for a variety of oscillatory integrals via the method of stationary phase; see Hörmander [160]. References on oscillatory integrals include the books of Titchmarsh [362], Erdélyi [107], Zygmund [388], [389], Stein [344], and Sogge [328]. The latter provides a treatment of Fourier integral operators.

http://www.springer.com/978-1-4939-1193-6

Classical Fourier Analysis
Grafakos, L.
2014, XVII, 638 p. 14 illus., 2 illus. in color., Hardcover
ISBN: 978-1-4939-1193-6