8

Elements of Fourier Analysis

It is easy to say that Fourier analysis, or harmonic analysis, originated in the work of Euler, Fourier, and others on trigonometric series; it is much harder to describe succinctly what the subject comprises today, for it is a meeting ground for ideas from many parts of analysis and has applications in such diverse areas as partial differential equations and algebraic number theory. Two of the central ingredients of harmonic analysis, however, are convolution operators and the Fourier transform, which we study in this chapter.

8.1 PRELIMINARIES

We begin by making some notational conventions. Throughout this chapter we shall be working on $\mathbb{R}^n$, and n will always refer to the dimension. In any measure-theoretic considerations we always have Lebesgue measure in mind unless we specify otherwise. Thus, if E is a measurable set in $\mathbb{R}^n$, we shall denote $L^p(E, m)$ by $L^p(E)$. If U is open in $\mathbb{R}^n$ and $k \in \mathbb{N}$, we denote by $C^k(U)$ the space of all functions on U whose partial derivatives of order $\leq k$ all exist and are continuous, and we set $C^\infty(U) = \bigcap_1^\infty C^k(U)$. Furthermore, for any $E \subset \mathbb{R}^n$ we denote by $C_c^\infty(E)$ the space of all C^∞ functions on $\mathbb{R}^n$ whose support is compact and contained in E. If $E = \mathbb{R}^n$ or $U = \mathbb{R}^n$, we shall usually omit it in naming function spaces: thus, $L^p = L^p(\mathbb{R}^n)$, $C^k = C^k(\mathbb{R}^n)$, $C_c^\infty = C_c^\infty(\mathbb{R}^n)$. If $x, y \in \mathbb{R}^n$, we set

$$x \cdot y = \sum_1^n x_j y_j, \qquad |x| = \sqrt{x \cdot x}.$$

It will be convenient to have a compact notation for partial derivatives. We shall write

$$\partial_j = \frac{\partial}{\partial x_j},$$

and for higher-order derivatives we use multi-index notation. A **multi-index** is an ordered n-tuple of nonnegative integers. If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, we set

$$|\alpha| = \sum_1^n \alpha_j, \qquad \alpha! = \prod_1^n \alpha_j!, \qquad \partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n},$$

and if $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$x^\alpha = \prod_1^n x_j^{\alpha_j}.$$

(The notation $|\alpha| = \sum \alpha_j$ is inconsistent with the notation $|x| = (\sum x_j^2)^{1/2}$, but the meaning will always be clear from the context.) Thus, for example, Taylor's formula for functions $f \in C^k$ reads

$$f(x) = \sum_{|\alpha| \le k} (\partial^\alpha f)(x_0) \frac{(x - x_0)^\alpha}{\alpha!} + R_k(x), \qquad \lim_{x \to x_0} \frac{|R_k(x)|}{|x - x_0|^k} = 0,$$

and the product rule for derivatives becomes

$$\partial^\alpha (fg) = \sum_{\beta + \gamma = \alpha} \frac{\alpha!}{\beta!\gamma!} (\partial^\beta f)(\partial^\gamma g)$$

(Exercise 1).

We shall often avail ourselves of the sloppy but handy device of using the same notation for a function and its value at a point. Thus, "x^α" may be used to denote the function whose value at any point x is x^α.

Two spaces of C^∞ functions on $\mathbb{R}^n$ will be of particular importance for us. The first is the space C_c^∞ of C^∞ functions with compact support. The existence of nonzero functions in C_c^∞ is not quite obvious; the standard construction is based on the fact that the function $\eta(t) = e^{-1/t}\chi_{(0,\infty)}(t)$ is C^∞ even at the origin (Exercise 3). If we set

$$\text{(8.1)} \qquad \psi(x) = \eta(1 - |x|^2) = \begin{cases} \exp[(|x|^2 - 1)^{-1}] & \text{if } |x| < 1, \\ 0 & \text{if } |x| \ge 1, \end{cases}$$

it follows that $\psi \in C^\infty$, and $\text{supp}(\psi)$ is the closed unit ball. In the next section we shall use this single function to manufacture elements of C_c^∞ in great profusion; see Propositions 8.17 and 8.18.

The other space of C^∞ functions we shall need is the **Schwartz space** $\mathcal{S}$ consisting of those C^∞ functions which, together with all their derivatives, vanish at infinity faster than any power of $|x|$. More precisely, for any nonnegative integer N and any multi-index α we define

$$\|f\|_{(N,\alpha)} = \sup_{x\in\mathbb{R}^n} (1+|x|)^N |\partial^\alpha f(x)|;$$

then

$$\mathcal{S} = \{f \in C^\infty : \|f\|_{(N,\alpha)} < \infty \text{ for all } N, \alpha\}.$$

Examples of functions in $\mathcal{S}$ are easy to find: for instance, $f_\alpha(x) = x^\alpha e^{-|x|^2}$ where α is any multi-index. Also, clearly $C_c^\infty \subset \mathcal{S}$.

It is an important observation that if $f \in \mathcal{S}$, then $\partial^\alpha f \in L^p$ for all α and all $p \in [1,\infty]$. Indeed, $|\partial^\alpha f(x)| \leq C_N(1+|x|)^{-N}$ for all N, and $(1+|x|)^{-N} \in L^p$ for $N > n/p$ by Corollary 2.52.

8.2 Proposition. *$\mathcal{S}$ is a Fréchet space with the topology defined by the norms $\|\cdot\|_{(N,\alpha)}$.*

Proof. The only nontrivial point is completeness. If $\{f_k\}$ is a Cauchy sequence in $\mathcal{S}$, then $\|f_j - f_k\|_{(N,\alpha)} \to 0$ for all N, α. In particular, for each α the sequence $\{\partial^\alpha f_k\}$ converges uniformly to a function g_α. Denoting by e_j the vector $(0,\dots,1,\dots,0)$ with the 1 in the jth position, we have

$$f_k(x+te_j) - f_k(x) = \int_0^t \partial_j f_k(x+se_j)\,ds.$$

Letting $k \to \infty$, we obtain

$$g_0(x+te_j) - g_0(x) = \int_0^t g_{e_j}(x+se_j)\,ds.$$

The fundamental theorem of calculus implies that $g_{e_j} = \partial_j g_0$, and an induction on $|\alpha|$ then yields $g_\alpha = \partial^\alpha g_0$ for all α. It is then easy to check that $\|f_k - g_0\|_{(N,\alpha)} \to 0$ for all α. ∎

Another useful characterization of $\mathcal{S}$ is the following.

8.3 Proposition. *If $f \in C^\infty$, then $f \in \mathcal{S}$ iff $x^\beta \partial^\alpha f$ is bounded for all multi-indices α, β iff $\partial^\alpha(x^\beta f)$ is bounded for all multi-indices α, β.*

Proof. Obviously $|x^\beta| \leq (1+|x|)^N$ for $|\beta| \leq N$. On the other hand, $\sum_1^n |x_j|^N$ is strictly positive on the unit sphere $|x| = 1$, so it has a positive minimum δ there. It follows that $\sum_1^n |x_j|^N \geq \delta|x|^N$ for all x since both sides are homogeneous of degree N, and hence

$$(1+|x|)^N \leq 2^N(1+|x|^N) \leq 2^N\Big[1 + \delta^{-1}\sum_1^n |x_j^N|\Big] \leq 2^N\delta^{-1}\sum_{|\beta|\leq N} |x^\beta|.$$

This establishes the first equivalence. The second one follows from the fact that each $\partial^\alpha(x^\beta f)$ is a linear combination of terms of the form $x^\gamma \partial^\delta f$ and vice versa, by the product rule (Exercise 1). ∎

We next investigate the continuity of translations on various function spaces. The following notation for translations will be used throughout this chapter and the next one: If f is a function on $\mathbb{R}^n$ and $y \in \mathbb{R}^n$,

$$\tau_y f(x) = f(x - y).$$

We observe that $\|\tau_y f\|_p = \|f\|_p$ for $1 \leq p \leq \infty$ and that $\|\tau_y f\|_u = \|f\|_u$. A function f is called **uniformly continuous** if $\|\tau_y f - f\|_u \to 0$ as $y \to 0$. (The reader should pause to check that this is equivalent to the usual ϵ-δ definition of uniform continuity.)

8.4 Lemma. *If $f \in C_c(\mathbb{R}^n)$, then f is uniformly continuous.*

Proof. Given $\epsilon > 0$, for each $x \in \operatorname{supp}(f)$ there exists $\delta_x > 0$ such that $|f(x-y)-f(x)| < \frac{1}{2}\epsilon$ if $|y| < \delta_x$. Since $\operatorname{supp}(f)$ is compact, there exist $x_1, \ldots, x_N$ such that the balls of radius $\frac{1}{2}\delta_{x_j}$ about x_j cover $\operatorname{supp}(f)$. If $\delta = \frac{1}{2}\min\{\delta_{x_j}\}$, then, one easily sees that $\|\tau_y f - f\|_u < \epsilon$ whenever $|y| < \delta$. ∎

8.5 Proposition. *If $1 \leq p < \infty$, translation is continuous in the L^p norm; that is, if $f \in L^p$ and $z \in \mathbb{R}^n$, then $\lim_{y\to 0} \|\tau_{y+z} f - \tau_z f\|_p = 0$.*

Proof. Since $\tau_{y+z} = \tau_y \tau_z$, by replacing f by $\tau_z f$ it suffices to assume that $z = 0$. First, if $g \in C_c$, for $|y| \leq 1$ the functions $\tau_y g$ are all supported in a common compact set K, so by Lemma 8.4,

$$\int |\tau_y g - g|^p \leq \|\tau_y g - g\|_u^p\, m(K) \to 0 \text{ as } y \to 0.$$

Now suppose $f \in L^p$. If $\epsilon > 0$, by Proposition 7.9 there exists $g \in C_c$ with $\|g - f\|_p < \epsilon/3$, so

$$\|\tau_y f - f\|_p \leq \|\tau_y(f - g)\|_p + \|\tau_y g - g\|_p + \|g - f\|_p < \tfrac{2}{3}\epsilon + \|\tau_y g - g\|_p,$$

and $\|\tau_y g - g\|_p < \epsilon/3$ if y is sufficiently small. ∎

Proposition 8.5 is false for $p = \infty$, as one should expect since the L^∞ norm is closely related to the uniform norm; see Exercise 4.

Some of our results will concern multiply periodic functions in $\mathbb{R}^n$, and for simplicity we shall take the fundamental period in each variable to be 1. That is, we define a function f on $\mathbb{R}^n$ to be **periodic** if $f(x + k) = f(x)$ for all $x \in \mathbb{R}^n$ and $k \in \mathbb{Z}^n$. Every periodic function is thus completely determined by its values on the unit cube

$$Q = \left[-\tfrac{1}{2}, \tfrac{1}{2}\right)^n.$$

Periodic functions may be regarded as functions on the space $\mathbb{R}^n/\mathbb{Z}^n \cong (\mathbb{R}/\mathbb{Z})^n$ of cosets of $\mathbb{Z}^n$, which we call the n-**dimensional torus** and denote by $\mathbb{T}^n$. (When $n = 1$ we write $\mathbb{T}$ rather than $\mathbb{T}^1$.) $\mathbb{T}^n$ is a compact Hausdorff space; it may be identified with the set of all $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ such that $|z_j| = 1$ for all j, via the map

$$(x_1, \ldots, x_n) \mapsto (e^{2\pi i x_1}, \ldots, e^{2\pi i x_n}).$$

On the other hand, for measure-theoretic purposes we identify $\mathbb{T}^n$ with the unit cube Q, and when we speak of Lebesgue measure on $\mathbb{T}^n$ we mean the measure induced on $\mathbb{T}^n$ by Lebesgue measure on Q. In particular, $m(\mathbb{T}^n) = 1$. Functions on $\mathbb{T}^n$ may be considered as periodic functions on $\mathbb{R}^n$ or as functions on Q; the point of view will be clear from the context when it matters.

Exercises

1. Prove the product rule for partial derivatives as stated in the text. Deduce that

$$\partial^\alpha(x^\beta f) = x^\beta \partial^\alpha f + \sum c_{\gamma\delta} x^\delta \partial^\gamma f, \qquad x^\beta \partial^\alpha f = \partial^\alpha(x^\beta f) + \sum c'_{\gamma\delta} \partial^\gamma(x^\delta f)$$

for some constants $c_{\gamma\delta}$ and $c'_{\gamma\delta}$ with $c_{\gamma\delta} = c'_{\gamma\delta} = 0$ unless $|\gamma| < |\alpha|$ and $|\delta| < |\beta|$.

2. Observe that the binomial theorem can be written as follows:

$$(x_1 + x_2)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha \qquad \big(x = (x_1, x_2),\ \alpha = (\alpha_1, \alpha_2)\big).$$

Prove the following generalizations:

a. The multinomial theorem: If $x \in \mathbb{R}^n$,

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha.$$

b. The n-dimensional binomial theorem: If $x, y \in \mathbb{R}^n$,

$$(x + y)^\alpha = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} x^\beta y^\gamma.$$

3. Let $\eta(t) = e^{-1/t}$ for $t > 0$, $\eta(t) = 0$ for $t \leq 0$.

a. For $k \in \mathbb{N}$ and $t > 0$, $\eta^{(k)}(t) = P_k(1/t)e^{-1/t}$ where P_k is a polynomial of degree $2k$.

b. $\eta^{(k)}(0)$ exists and equals zero for all $k \in \mathbb{N}$.

4. If $f \in L^\infty$ and $\|\tau_y f - f\|_\infty \to 0$ as $y \to 0$, then f agrees a.e. with a uniformly continuous function. (Let $A_r f$ be as in Theorem 3.18. Then $A_r f$ is uniformly continuous for $r > 0$ and uniformly Cauchy as $r \to 0$.)

8.2 CONVOLUTIONS

Let f and g be measurable functions on $\mathbb{R}^n$. The **convolution** of f and g is the function $f * g$ defined by

$$f * g(x) = \int f(x - y) g(y)\, dy$$

for all x such that the integral exists. Various conditions can be imposed on f and g to guarantee that $f * g$ is defined at least almost everywhere. For example, if f is bounded and compactly supported, g can be any locally integrable function; see also Propositions 8.7–8.9 below.

In what follows, we shall need the fact that if f is a measurable function on $\mathbb{R}^n$, then the function $K(x, y) = f(x-y)$ is measurable on $\mathbb{R}^n \times \mathbb{R}^n$. We have $K = f \circ s$ where $s(x, y) = x - y$; since s is continuous, K is Borel measurable if f is Borel measurable. This can always be assumed without affecting the definition of $f * g$, by Proposition 2.12. However, the Lebesgue measurability of K also follows from the Lebesgue measurability of f; see Exercise 5.

The elementary properties of convolutions are summarized in the following proposition.

8.6 Proposition. *Assuming that all integrals in question exist, we have*

a. $f * g = g * f$,

b. $(f * g) * h = f * (g * h)$.

c. For $z \in \mathbb{R}^n$, $\tau_z(f * g) = (\tau_z f) * g = f * (\tau_z g)$.

d. If A is the closure of $\{x+y : x \in \operatorname{supp}(f),\ y \in \operatorname{supp}(g)\}$, *then* $\operatorname{supp}(f*g) \subset A$.

Proof. (a) is proved by the substitution $z = x - y$:

$$f * g(x) = \int f(x-y)g(y)\,dy = \int f(z)g(x-z)\,dz = g * f(x).$$

(b) follows from (a) and Fubini's theorem:

$$\begin{aligned}\int (f * g) * h(x) &= \iint f(y)g(x-z-y)h(z)\,dy\,dz \\ &= \iint f(y)g(x-y-z)h(z)\,dz\,dy = f * (g * h)(x).\end{aligned}$$

As for (c),

$$\tau_z(f * g)(x) = \int f(x-z-y)g(y)\,dy = \int \tau_z f(x-y)g(y)\,dy = (\tau_z f) * g(x),$$

and by (a),

$$\tau_z(f * g) = \tau_z(g * f) = (\tau_z g) * f = f * (\tau_z g).$$

For (d), we observe that if $x \notin A$, then for any $y \in \operatorname{supp}(g)$ we have $x-y \notin \operatorname{supp}(f)$; hence $f(x-y)g(y) = 0$ for all y, so $f * g(x) = 0$. ∎

The following two propositions contain the basic facts about convolutions of L^p functions.

8.7 Young's Inequality. *If $f \in L^1$ and $g \in L^p$ ($1 \le p \le \infty$), then $f * g(x)$ exists for almost every x, $f * g \in L^p$, and $\|f * g\|_p \le \|f\|_1 \|g\|_p$.*

Proof. This is a special case of Theorem 6.18, with $K(x,y) = f(x-y)$. Alternatively, one can use Minkowski's inequality for integrals:

$$\|f * g\|_p = \left\|\int f(y)g(\cdot - y)\,dy\right\|_p \leq \int |f(y)|\,\|\tau_y g\|_p\,dy = \|f\|_1\|g\|_p.$$

∎

8.8 Proposition. *If p and q are conjugate exponents, $f \in L^p$, and $g \in L^q$, then $f * g(x)$ exists for every x, $f * g$ is bounded and uniformly continuous, and $\|f * g\|_u \leq \|f\|_p\|g\|_q$. If $1 < p < \infty$ (so that $1 < q < \infty$ also), then $f * g \in C_0(\mathbb{R}^n)$.*

Proof. The existence of $f * g$ and the estimate for $\|f * g\|_u$ follow immediately from Hölder's inequality. In view of Propositions 8.5 and 8.6, so does the uniform continuity of $f * g$: If $1 \leq p < \infty$,

$$\|\tau_y(f * g) - f * g\|_u = \|(\tau_y f - f) * g\|_\infty \leq \|\tau_y f - f\|_p\|g\|_q \to 0 \text{ as } y \to 0.$$

(If $p = \infty$, interchange the roles of f and g.) Finally, if $1 < p, q < \infty$, choose sequences $\{f_n\}$ and $\{g_n\}$ of functions with compact support such that $\|f_n - f\|_p \to 0$ and $\|g_n - g\|_q \to 0$. By Proposition 8.6d and what we have just proved, $f_n * g_n \in C_c$. But

$$\|f_n * g_n - f * g\|_u \leq \|f_n - f\|_p\|g_n\|_q + \|f\|_p\|g_n - g\|_q \to 0,$$

so $f * g \in C_0$ by Proposition 4.35. ∎

The preceding results are all we shall use, but for the sake of completeness we state also the following generalization.

8.9 Proposition. *Suppose $1 \leq p, q, r \leq \infty$ and $p^{-1} + q^{-1} = r^{-1} + 1$.*

a. **(Young's Inequality, General Form)** *If $f \in L^p$ and $g \in L^q$, then $f * g \in L^r$ and $\|f * g\|_r \leq \|f\|_p\|g\|_q$.*
b. *Suppose also that $p > 1$, $q > 1$, and $r < \infty$. If $f \in L^p$ and $g \in$ weak L^q, then $f * g \in L^r$ and $\|f * g\|_r \leq C_{pq}\|f\|_p[g]_q$ where C_{pq} is independent of f and g.*
c. *Suppose that $p = 1$ and $r = q > 1$. If $f \in L^1$ and $g \in$ weak L^q, then $f * g \in$ weak L^q and $[f * g]_q \leq C_q\|f\|_1$, where C_q is independent of f and g.*

Proof. To prove (a), let q be fixed. The special cases $p = 1$, $r = q$ and $p = q/(q-1)$, $r = \infty$ are Propositions 8.7 and 8.8. The general case then follows from the Riesz-Thorin interpolation theorem. (See also Exercise 6 for a direct proof.) (b) and (c) are special cases of Theorem 6.36. ∎

One of the most important properties of convolution is that, roughly speaking, $f * g$ is at least as smooth as either f or g, because formally we have

$$\partial^\alpha(f * g)(x) = \partial^\alpha \int f(x-y)g(y)\,dy = \int \partial^\alpha f(x-y)g(y)\,dy = (\partial^\alpha f) * g(x),$$

and similarly $\partial^\alpha(f*g) = f*(\partial^\alpha g)$. To make this precise, one needs only to impose conditions on f and g so that differentiation under the integral sign is legitimate. One such result is the following; see also Exercises 7 and 8.

8.10 Proposition. *If $f \in L^1$, $g \in C^k$, and $\partial^\alpha g$ is bounded for $|\alpha| \le k$, then $f*g \in C^k$ and $\partial^\alpha(f*g) = f*(\partial^\alpha g)$ for $|\alpha| \le k$.*

Proof. This is clear from Theorem 2.27. ∎

8.11 Proposition. *If $f, g \in \mathcal{S}$, then $f*g \in \mathcal{S}$.*

Proof. First, $f*g \in C^\infty$ by Proposition 8.10. Since

$$1+|x| \le 1+|x-y|+|y| \le (1+|x-y|)(1+|y|), \tag{8.12}$$

we have

$$\begin{aligned}(1+|x|)^N|\partial^\alpha(f*g)(x)| &\le \int (1+|x-y|)^N|\partial^\alpha f(x-y)|(1+|y|)^N|g(y)|\,dy \\ &\le \|f\|_{(N,\alpha)}\|g\|_{(N+n+1,\alpha)}\int(1+|y|)^{-n-1}\,dy,\end{aligned}$$

which is finite by Corollary 2.52. ∎

Convolutions of functions on the torus $\mathbb{T}^n$ are defined just as for functions on $\mathbb{R}^n$. (If one regards functions on $\mathbb{T}^n$ as periodic functions on $\mathbb{R}^n$, of course, the integration is to be extended over the unit cube rather than $\mathbb{R}^n$.) All of the preceding results remain valid, with the same proofs.

The following theorem underlies many of the important applications of convolutions on $\mathbb{R}^n$. We introduce a bit of notation that will be used frequently hereafter: If ϕ is any function on $\mathbb{R}^n$ and $t > 0$, we set

$$\phi_t(x) = t^{-n}\phi(t^{-1}x). \tag{8.13}$$

We observe that if $\phi \in L^1$, then $\int \phi_t$ is independent of t, by Theorem 2.44:

$$\int \phi_t = \int \phi(t^{-1}x)t^{-n}\,dx = \int \phi(y)\,dy = \int \phi.$$

Moreover, the "mass" of ϕ_t becomes concentrated at the origin as $t \to 0$. (Draw a picture if this isn't clear.)

8.14 Theorem. *Suppose $\phi \in L^1$ and $\int \phi(x)\,dx = a$.*

a. If $f \in L^p$ $(1 \le p < \infty)$, then $f*\phi_t \to af$ in the L^p norm as $t \to 0$.

b. If f is bounded and uniformly continuous, then $f*\phi_t \to af$ uniformly as $t \to 0$.

c. If $f \in L^\infty$ and f is continuous on an open set U, then $f*\phi_t \to af$ uniformly on compact subsets of U as $t \to 0$.

Proof. Setting $y = tz$, we have

$$\begin{aligned} f * \phi_t(x) - af(x) &= \int [f(x-y) - f(x)]\phi_t(y)\,dy \\ &= \int [f(x-tz) - f(x)]\phi(z)\,dz \\ &= \int [\tau_{tz}f(x) - f(x)]\phi(z)\,dz. \end{aligned}$$

Apply Minkowski's inequality for integrals:

$$\|f * \phi_t - af\|_p \leq \int \|\tau_{tz}f - f\|_p |\phi(z)|\,dz.$$

Now, $\|\tau_{tz}f - f\|_p$ is bounded by $2\|f\|_p$ and tends to 0 as $t \to 0$ for each z, by Proposition 8.5. Assertion (a) therefore follows from the dominated convergence theorem.

The proof of (b) is exactly the same, with $\|\cdot\|_p$ replaced by $\|\cdot\|_u$. The estimate for $\|f * \phi_t - af\|_u$ is obvious, and $\|\tau_{tz}f - f\|_u \to 0$ as $t \to 0$ by the uniform continuity of f.

As for (c), given $\epsilon > 0$ let us choose a compact $E \subset \mathbb{R}^n$ such that $\int_{E^c} |\phi| < \epsilon$. Also, let K be a compact subset of U. If t is sufficiently small, then, we will have $x - tz \in U$ for all $x \in K$ and $z \in E$, so from the compactness of K it follows as in Lemma 8.4 that

$$\sup_{x \in K,\ z \in E} |f(x - tz) - f(x)| < \epsilon$$

for small t. But then

$$\begin{aligned} \sup_{x\in K} |f * \phi_t(x) - af(x)| &\leq \sup_{x\in K} \left[\int_E + \int_{E^c}\right] |f(x-tz) - f(x)|\,|\phi(z)|\,dz \\ &\leq \epsilon \int |\phi| + 2\|f\|_\infty \epsilon, \end{aligned}$$

from which (c) follows. ∎

If we impose slightly stronger conditions on ϕ, we can also show that $f * \phi_t \to af$ almost everywhere for $f \in L^p$. The device in the following proof of breaking up an integral into pieces corresponding to the dyadic intervals $[2^k, 2^{k+1}]$ and estimating each piece separately is a standard trick of the trade in Fourier analysis.

8.15 Theorem. *Suppose $|\phi(x)| \leq C(1+|x|)^{-n-\epsilon}$ for some $C, \epsilon > 0$ (which implies that $\phi \in L^1$ by Corollary 2.52), and $\int \phi(x)\,dx = a$. If $f \in L^p$ ($1 \leq p \leq \infty$), then $f * \phi_t(x) \to af(x)$ as $t \to 0$ for every x in the Lebesgue set of f — in particular, for almost every x, and for every x at which f is continuous.*

Proof. If x is in the Lebesgue set of f, for any $\delta > 0$ there exists $\eta > 0$ such that

$$\int_{|y|<r} |f(x-y) - f(x)|\,dy \leq \delta r^n \text{ for } r \leq \eta. \tag{8.16}$$

Let us set

$$I_1 = \int_{|y|<\eta} |f(x-y) - f(x)|\,|\phi_t(y)|\,dy,$$

$$I_2 = \int_{|y|\geq\eta} |f(x-y) - f(x)|\,|\phi_t(y)|\,dy.$$

We claim that I_1 is bounded by $A\delta$ where A is independent of t, whereas $I_2 \to 0$ as $t \to 0$. Since

$$|f * \phi_t(x) - af(x)| \leq I_1 + I_2,$$

we will have

$$\limsup_{t\to 0} |f * \phi_t(x) - af(x)| \leq A\delta,$$

and since δ is arbitrary, this will complete the proof.

To estimate I_1, let K be the integer such that $2^K \leq \eta/t < 2^{K+1}$ if $\eta/t \geq 1$, and $K = 0$ if $\eta/t < 1$. We view the ball $|y| < \eta$ as the union of the annuli $2^{-k}\eta \leq |y| < 2^{1-k}\eta$ $(1 \leq k \leq K)$ and the ball $|y| < 2^{-K}\eta$. On the kth annulus we use the estimate

$$|\phi_t(y)| \leq Ct^{-n}\left|\frac{y}{t}\right|^{-n-\epsilon} \leq Ct^{-n}\left[\frac{2^{-k}\eta}{t}\right]^{-n-\epsilon},$$

and on the ball $|y| < 2^{-K}\eta$ we use the estimate $|\phi_t(y)| \leq Ct^{-n}$. Thus

$$I_1 \leq \sum_1^K Ct^{-n}\left[\frac{2^{-k}\eta}{t}\right]^{-n-\epsilon} \int_{2^{-k}\eta\leq|y|<2^{1-k}\eta} |f(x-y) - f(x)|\,dy$$
$$+Ct^{-n}\int_{|y|<2^{-K}\eta} |f(x-y) - f(x)|\,dy.$$

Therefore, by (8.16) and the fact that $2^K \leq \eta/t < 2^{K+1}$,

$$\begin{aligned}
I_1 &\leq C\delta\sum_1^K (2^{1-k}\eta)^n t^{-n}\left[\frac{2^{-k}\eta}{t}\right]^{-n-\epsilon} + C\delta t^{-n}(2^{-K}\eta)^n \\
&= 2^n C\delta\left[\frac{\eta}{t}\right]^{-\epsilon}\sum_1^K 2^{k\epsilon} + C\delta\left[\frac{2^{-K}\eta}{t}\right]^n \\
&= 2^n C\delta\left[\frac{\eta}{t}\right]^{-\epsilon}\frac{2^{(K+1)\epsilon} - 2^\epsilon}{2^\epsilon - 1} + C\delta\left[\frac{2^{-K}\eta}{t}\right]^n \\
&\leq 2^n C\big[2^\epsilon(2^\epsilon - 1)^{-1} + 1\big]\delta.
\end{aligned}$$

As for I_2, if p' is the conjugate exponent to p and χ is the characteristic function of $\{y : |y| \geq \eta\}$, by Hölder's inequality we have

$$\begin{aligned}
I_2 &\leq \int_{|y|\geq\eta} \big(|f(x-y)| + |f(x)|\big)|\phi_t(y)|\,dy \\
&\leq \|f\|_p\|\chi\phi_t\|_{p'} + |f(x)|\,\|\chi\phi_t\|_1,
\end{aligned}$$

so it suffices to show that for $1 \leq q \leq \infty$, and in particular for $q = 1$ and $q = p'$, $\|\chi\phi_t\|_q \to 0$ as $t \to 0$. If $q = \infty$, this is obvious:

$$\|\chi\phi_t\|_\infty \leq Ct^{-n}\big[1 + (\eta/t)\big]^{-n-\epsilon} = Ct^\epsilon(t+\eta)^{-n-\epsilon} \leq C\eta^{-n-\epsilon}t^\epsilon.$$

If $q < \infty$, by Corollary 2.51 we have

$$\begin{aligned}\|\chi\phi_t\|_q^q &= \int_{|y|\geq\eta} t^{-nq}|\phi(t^{-1}y)|^q\,dy = t^{n(1-q)}\int_{|z|\geq\eta/t}|\phi(z)|^q\,dz \\ &\leq C_1 t^{n(1-q)}\int_{\eta/t}^\infty r^{n-1-(n+\epsilon)q}\,dr = C_2 t^{n(1-q)}\left[\frac{\eta}{t}\right]^{n-(n+\epsilon)q} = C_3 t^{\epsilon q}.\end{aligned}$$

In either case, $\|\chi\phi_t\|_q$ is dominated by t^ϵ, so we are done. ∎

In most of the applications of the preceding two theorems one has $a = 1$, although the case $a = 0$ is also useful. If $a = 1$, $\{\phi_t\}_{t>0}$ is called an **approximate identity**, as it furnishes an approximation to the identity operator on L^p by convolution operators. This construction is useful for approximating L^p functions by functions having specified regularity properties. For example, we have the following two important results:

8.17 Proposition. *C_c^∞ (and hence also $\mathcal{S}$) is dense in L^p ($1 \leq p < \infty$) and in C_0.*

Proof. Given $f \in L^p$ and $\epsilon > 0$, there exists $g \in C_c$ with $\|f - g\|_p < \epsilon/2$, by Proposition 7.9. Let ϕ be a function in C_c^∞ such that $\int \phi = 1$ — for example, take $\phi = (\int \psi)^{-1}\psi$ where ψ is as in (8.1). Then $g * \phi_t \in C_c^\infty$ by Propositions 8.6d and 8.10, and $\|g * \phi_t - g\|_p < \epsilon/2$ for sufficiently small t by Theorem 8.14. The same argument applies if L^p is replaced by C_0, $\|\cdot\|_p$ by $\|\cdot\|_u$, and Proposition 7.9 by Proposition 4.35. ∎

8.18 The C^∞ Urysohn Lemma. *If $K \subset \mathbb{R}^n$ is compact and U is an open set containing K, there exists $f \in C_c^\infty$ such that $0 \leq f \leq 1$, $f = 1$ on K, and $\operatorname{supp}(f) \subset U$.*

Proof. Let $\delta = \rho(K, U^c)$ (the distance from K to U^c, which is positive since K is compact), and let $V = \{x : \rho(x, K) < \delta/3\}$. Choose a nonnegative $\phi \in C_c^\infty$ such that $\int \phi = 1$ and $\phi(x) = 0$ for $|x| \geq \delta/3$ (for example, $(\int \psi)^{-1}\psi_{\delta/3}$ with ψ as in (8.1)), and set $f = \chi_V * \phi$. Then $f \in C_c^\infty$ by Propositions 8.6d and 8.10, and it is easily checked that $0 \leq f \leq 1$, $f = 1$ on K, and $\operatorname{supp}(f) \subset \{x : \rho(x, K) \leq 2\delta/3\} \subset U$. ∎

Exercises

5. If $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is defined by $s(x, y) = x - y$, then $s^{-1}(E)$ is Lebesgue measurable whenever E is Lebesgue measurable. (For $n = 1$, draw a picture of $s^{-1}(E) \subset \mathbb{R}^2$. It should be clear that after rotation through an angle $\pi/4$, $s^{-1}(E)$

becomes $F \times \mathbb{R}$ where $F = \{x : \sqrt{2}\,x \in E\}$, and Theorem 2.44 can be applied. The same idea works in higher dimensions.)

6. Prove Theorem 8.9a by using Exercise 31 in §6.3 to show that

$$|f * g(x)|^r \leq \|f\|_p^{r-p}\|g\|_q^{r-q} \int |f(y)|^p |g(x-y)|^q \, dy.$$

7. If f is locally integrable on $\mathbb{R}^n$ and $g \in C^k$ has compact support, then $f * g \in C^k$.

8. Suppose that $f \in L^p(\mathbb{R})$. If there exists $h \in L^p(\mathbb{R})$ such that

$$\lim_{y \to 0} \left\| y^{-1}(\tau_{-y} f - f) - h \right\|_p = 0,$$

we call h the **(strong) L^p derivative** of f. If $f \in L^p(\mathbb{R}^n)$, L^p partial derivatives of f are defined similarly. Suppose that p and q are conjugate exponents, $f \in L^p$, $g \in L^q$, and the L^p derivative $\partial_j f$ exists. Then $\partial_j(f * g)$ exists (in the ordinary sense) and equals $(\partial_j f) * g$.

9. If $f \in L^p(\mathbb{R})$, the L^p derivative of f (call it h; see Exercise 8) exists iff f is absolutely continuous on every bounded interval (perhaps after modification on a null set) and its pointwise derivative f' is in L^p, in which case $h = f'$ a.e. (For "only if," use Exercise 8: If $g \in C_c$ with $\int g = 1$, then $f * g_t \to f$ and $(f * g_t)' \to h$ as $t \to 0$. For "if," write

$$\frac{f(x+y) - f(x)}{y} - f'(x) = \frac{1}{y} \int_0^y [f'(x+t) - f'(x)] \, dt$$

and use Minkowski's inequality for integrals.)

10. Let ϕ satisfy the hypotheses of Theorem 8.15. If $f \in L^p$ $(1 \leq p \leq \infty)$, define the **ϕ-maximal function** of f to be $M_\phi f(x) = \sup_{t>0} |f * \phi_t(x)|$. (Observe that the Hardy-Littlewood maximal function Hf is $M_\phi |f|$ where ϕ is the characteristic function of the unit ball divided by the volume of the ball.) Show that there is a constant C, independent of f, such that $M_\phi f \leq C \cdot Hf$. (Break up the integral $\int f(x-y)\phi_t(y)\,dy$ as the sum of the integrals over $|y| \leq t$ and over $2^k t < |y| \leq 2^{k+1} t$ $(k = 0, 1, 2, \ldots)$, and estimate ϕ_t on each region.) It follows from Theorem 3.17 that M_ϕ is weak type (1,1), and the proof of Theorem 3.18 can then be adapted to give an alternate demonstration that $f * \phi_t \to (\int \phi) f$ a.e.

11. Young's inequality shows that L^1 is a Banach algebra, the product being convolution.

   a. If $\mathfrak{I}$ is an ideal in the algebra L^1, so is its closure in L^1.

   b. If $f \in L^1$, the smallest closed ideal in L^1 containing f is the smallest closed subspace of L^1 containing all translates of f. (If $g \in C_c$, $f * g(x)$ can be approximated by sums $\sum f(x - y_j) g(y_j) \Delta y_j$. On the other hand, if $\{\phi_t\}$ is an approximate identity, $f * \tau_y(\phi_t) \to \tau_y f$ as $t \to 0$.)

8.3 THE FOURIER TRANSFORM

One of the fundamental principles of harmonic analysis is the exploitation of symmetry. To be more specific, if one is doing analysis on a space on which a group acts, it is a good idea to study functions (or other analytic objects) that transform in simple ways under the group action, and then try to decompose arbitrary functions as sums or integrals of these basic functions.

The spaces we are studying are $\mathbb{R}^n$ and $\mathbb{T}^n$, which are Abelian groups under addition and act on themselves by translation. The building blocks of harmonic analysis on these spaces are the functions that transform under translation by multiplication by a factor of absolute value one, that is, functions f such that for each x there is a number $\phi(x)$ with $|\phi(x)| = 1$ such that $f(y+x) = \phi(x)f(y)$. If f and ϕ have this property, then $f(x) = \phi(x)f(0)$, so f is completely determined by ϕ once $f(0)$ is given; moreover,

$$\phi(x)\phi(y)f(0) = \phi(x)f(y) = f(x+y) = \phi(x+y)f(0),$$

so that (unless $f = 0$) $\phi(x+y) = \phi(x)\phi(y)$. In short, to find all f's that transform as described above, it suffices to find all ϕ's of absolute value one that satisfy the functional equation $\phi(x+y) = \phi(x)\phi(y)$. Upon imposing the natural requirement that ϕ should be measurable, we have a complete solution to this problem.

8.19 Theorem. *If ϕ is a measurable function on $\mathbb{R}^n$ (resp. $\mathbb{T}^n$) such that $\phi(x+y) = \phi(x)\phi(y)$ and $|\phi| = 1$, there exists $\xi \in \mathbb{R}^n$ (resp. $\xi \in \mathbb{T}^n$) such that $\phi(x) = e^{2\pi i\xi\cdot x}$.*

Proof. We first prove this assertion on $\mathbb{R}$. Let $a \in \mathbb{R}$ be such that $\int_0^a \phi(t)\,dt \neq 0$; such an a surely exists, for otherwise the Lebesgue differentiation theorem would imply that $\phi = 0$ a.e. Setting $A = (\int_0^a \phi(t)\,dt)^{-1}$, then, we have

$$\phi(x) = A\int_0^a \phi(x)\phi(t)\,dt = A\int_0^a \phi(x+t)\,dt = A\int_x^{x+a} \phi(t)\,dt.$$

Thus ϕ, being the indefinite integral of a locally integrable function, is continuous; and then, being the integral of a continuous function, it is C^1. Moreover,

$$\phi'(x) = A[\phi(x+a) - \phi(x)] = B\phi(x), \text{ where } B = A[\phi(a) - 1].$$

It follows that $(d/dx)(e^{-Bx}\phi(x)) = 0$, so that $e^{-Bx}\phi(x)$ is constant. Since $\phi(0) = 1$, we have $\phi(x) = e^{Bx}$, and since $|\phi| = 1$, B is purely imaginary, so $B = 2\pi i\xi$ for some $\xi \in \mathbb{R}$. This completes the proof for $\mathbb{R}$; as for $\mathbb{T}$, the ϕ we have been considering will be periodic (with period 1) iff $e^{2\pi i\xi} = 1$ iff $\xi \in \mathbb{Z}$.

The n-dimensional case follows easily, for if $e_1, \ldots, e_n$ is the standard basis for $\mathbb{R}^n$, the functions $\psi_j(t) = \phi(te_j)$ satisfy $\psi_j(t+s) = \psi_j(t)\psi_j(s)$ on $\mathbb{R}$, so that $\psi_j(t) = e^{2\pi i\xi_j t}$, and hence

$$\phi(x) = \phi\Big(\sum_1^n x_je_j\Big) = \prod_1^n \psi_j(x_j) = e^{2\pi i\xi\cdot x}.$$

∎

The idea now is to decompose more or less arbitrary functions on $\mathbb{T}^n$ or $\mathbb{R}^n$ in terms of the exponentials $e^{2\pi i\xi\cdot x}$. In the case of $\mathbb{T}^n$ this works out very simply for L^2 functions:

8.20 Theorem. *Let $E_\kappa(x) = e^{2\pi i\kappa\cdot x}$. Then $\{E_\kappa : \kappa \in \mathbb{Z}^n\}$ is an orthonormal basis of $L^2(\mathbb{T}^n)$.*

Proof. Verification of orthonormality is an easy exercise in calculus; by Fubini's theorem it boils down to the fact that $\int_0^1 e^{2\pi ikt}\,dt$ equals 1 if $k = 0$ and equals 0 otherwise. Next, since $E_\kappa E_\lambda = E_{\kappa+\lambda}$, the set of finite linear combinations of the E_κ's is an algebra. It clearly separates points on $\mathbb{T}^n$; also, $E_0 = 1$ and $\overline{E}_\kappa = E_{-\kappa}$. Since $\mathbb{T}^n$ is compact, the Stone-Weierstrass theorem implies that this algebra is dense in $C(\mathbb{T}^n)$ in the uniform norm and hence in the L^2 norm, and $C(\mathbb{T}^n)$ is itself dense in $L^2(\mathbb{T}^n)$ by Proposition 7.9. It follows that $\{E_\kappa\}$ is a basis. ∎

To restate this result: If $f \in L^2(\mathbb{T}^n)$, we define its **Fourier transform** $\widehat{f}$, a function on $\mathbb{Z}^n$, by

$$\widehat{f}(\kappa) = \langle f, E_\kappa\rangle = \int_{\mathbb{T}^n} f(x)e^{-2\pi i\kappa\cdot x}\,dx.$$

and we call the series

$$\sum_{\kappa\in\mathbb{Z}^n} \widehat{f}(\kappa)E_\kappa$$

the **Fourier series** of f. The term "Fourier transform" is also used to mean the map $f \mapsto \widehat{f}$. Theorem 8.20 then says that the Fourier transform maps $L^2(\mathbb{T}^n)$ onto $l^2(\mathbb{Z}^n)$, that $\|\widehat{f}\|_2 = \|f\|_2$ (Parseval's identity), and that the Fourier series of f converges to f in the L^2 norm. We shall consider the question of pointwise convergence in the next two sections.

Actually, the definition of $\widehat{f}(\kappa)$ makes sense if f is merely in $L^1(\mathbb{T}^n)$, and $|\widehat{f}(\kappa)| \leq \|f\|_1$, so the Fourier transform extends to a norm-decreasing map from $L^1(\mathbb{T}^n)$ to $l^\infty(\mathbb{Z}^n)$. (The Fourier series of an L^1 function may be quite badly behaved, but there are still methods for recovering f from $\widehat{f}$ when $f \in L^1$, as we shall see in the next section.) Interpolating between L^1 and L^2, we have the following result.

8.21 The Hausdorff-Young Inequality. *Suppose that $1 \leq p \leq 2$ and q is the conjugate exponent to p. If $f \in L^p(\mathbb{T}^n)$, then $\widehat{f} \in l^q(\mathbb{Z}^n)$ and $\|\widehat{f}\|_q \leq \|f\|_p$.*

Proof. Since $\|\widehat{f}\|_\infty \leq \|f\|_1$ and $\|\widehat{f}\|_2 = \|f\|_2$ for $f \in L^1$ or $f \in L^2$, the assertion follows from the Riesz-Thorin interpolation theorem. ∎

The situation on $\mathbb{R}^n$ is more delicate. The formal analogue of Theorem 8.20 should be

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

These relations turn out to be valid when suitably interpreted, but some care is needed. In the first place, the integral defining $\widehat{f}(\xi)$ is likely to diverge if $f \in L^2$. However, it certainly converges if $f \in L^1$. We therefore begin by defining the **Fourier transform** of $f \in L^1(\mathbb{R}^n)$ by

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

(We use the notation $\mathcal{F}$ for the Fourier transform only in certain situations where it is needed for clarity.) Clearly $\|\widehat{f}\|_u \le \|f\|_1$, and $\widehat{f}$ is continuous by Theorem 2.27; thus

$$\mathcal{F} : L^1(\mathbb{R}^n) \to BC(\mathbb{R}^n).$$

We summarize the elementary properties of $\mathcal{F}$ in a theorem.

8.22 Theorem. *Suppose $f, g \in L^1(\mathbb{R}^n)$.*

a. $(\tau_y f)\widehat{\ }(\xi) = e^{-2\pi i\xi\cdot y}\widehat{f}(\xi)$ *and* $\tau_\eta(\widehat{f}) = \widehat{h}$ *where* $h(x) = e^{2\pi i\eta\cdot x}f(x)$.

b. If T is an invertible linear transformation of $\mathbb{R}^n$ and $S = (T^*)^{-1}$ is its inverse transpose, then $(f \circ T)\widehat{\ } = |\det T|^{-1}\widehat{f} \circ S$. In particular, if T is a rotation, then $(f \circ T)\widehat{\ } = \widehat{f} \circ T$; and if $Tx = t^{-1}x$ $(t > 0)$, then $(f \circ T)\widehat{\ }(\xi) = t^n\widehat{f}(t\xi)$, so that $(f_t)\widehat{\ }(\xi) = \widehat{f}(t\xi)$ in the notation of (8.13).

c. $(f * g)\widehat{\ } = \widehat{f}\widehat{g}$.

d. If $x^\alpha f \in L^1$ for $|\alpha| \le k$, then $\widehat{f} \in C^k$ and $\partial^\alpha \widehat{f} = [(-2\pi i x)^\alpha f]\widehat{\ }$.

e. If $f \in C^k$, $\partial^\alpha f \in L^1$ for $|\alpha| \le k$, and $\partial^\alpha f \in C_0$ for $|\alpha| \le k-1$, then $(\partial^\alpha f)\widehat{\ }(\xi) = (2\pi i\xi)^\alpha \widehat{f}(\xi)$.

f. (**The Riemann-Lebesgue Lemma**) $\mathcal{F}(L^1(\mathbb{R}^n)) \subset C_0(\mathbb{R}^n)$.

Proof. a. We have

$$(\tau_y f)\widehat{\ }(\xi) = \int f(x-y)e^{-2\pi i\xi\cdot x}\,dx = \int f(x)e^{-2\pi i\xi\cdot(x+y)}\,dx = e^{-2\pi i\xi\cdot y}\widehat{f}(\xi),$$

and similarly for the other formula.

b. By Theorem 2.44,

$$\begin{aligned}(f \circ T)\widehat{\ }(\xi) &= \int f(Tx)e^{-2\pi i\xi\cdot x}\,dx = |\det T|^{-1}\int f(x)e^{-2\pi i\xi\cdot T^{-1}x}\,dx \\ &= |\det T|^{-1}\int f(x)e^{-2\pi iS\xi\cdot x}\,dx = |\det T|^{-1}\widehat{f}(S\xi).\end{aligned}$$

c. By Fubini's theorem,

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

d. By Theorem 2.27 and induction on $|\alpha|$,

$$\partial^\alpha \widehat{f}(\xi) = \partial_\xi^\alpha \int f(x)e^{-2\pi i\xi\cdot x}\,dx = \int f(x)(-2\pi i x)^\alpha e^{-2\pi i\xi\cdot x}\,dx.$$

e. First assume $n = |\alpha| = 1$. Since $f \in C_0$, we can integrate by parts:

$$\begin{aligned}\int f'(x)e^{-2\pi i\xi\cdot x}\,dx &= f(x)e^{-2\pi i\xi\cdot x}\Big|_{-\infty}^{\infty} - \int f(x)(-2\pi i\xi)e^{-2\pi i\xi\cdot x}\,dx \\ &= 2\pi i\xi\widehat{f}(\xi).\end{aligned}$$

The argument for $n > 1$, $|\alpha| = 1$ is the same — to compute $(\partial_j f)\widehat{\ }$, integrate by parts in the jth variable — and the general case follows by induction on $|\alpha|$.

f. By (e), if $f \in C^1 \cap C_c$, then $|\xi|\widehat{f}(\xi)$ is bounded and hence $\widehat{f} \in C_0$. But the set of all such f's is dense in L^1 by Proposition 8.17, and $\widehat{f}_n \to \widehat{f}$ uniformly whenever $f_n \to f$ in L^1. Since C_0 is closed in the uniform norm, the result follows. ∎

Parts (d) and (e) of Theorem 8.22 point to a fundamental property of the Fourier transform: Smoothness properties of f are reflected in the rate of decay of $\widehat{f}$ at infinity, and vice versa. Parts (a), (c), (e), and (f) of this theorem are valid also on $\mathbb{T}^n$, as is (b) provided that T leaves the lattice $\mathbb{Z}^n$ invariant (Exercise 12).

8.23 Corollary. *$\mathcal{F}$ maps the Schwartz class $\mathcal{S}$ continuously into itself.*

Proof. If $f \in \mathcal{S}$, then $x^\alpha\partial^\beta f \in L^1 \cap C_0$ for all α, β, so by Theorem 8.22d,e, $\widehat{f}$ is C^∞ and

$$(x^\alpha\partial^\beta f)\widehat{\ } = (-1)^{|\alpha|}(2\pi i)^{|\beta|-|\alpha|}\partial^\alpha(\xi^\beta\widehat{f}).$$

Thus $\partial^\alpha(\xi^\beta\widehat{f})$ is bounded for all α, β, whence $\widehat{f} \in \mathcal{S}$ by Proposition 8.3. Moreover, since $\int(1+|x|)^{-n-1}\,dx < \infty$,

$$\|(x^\alpha\partial^\beta f)\widehat{\ }\|_u \le \|x^\alpha\partial^\beta f\|_1 \le C\|(1+|x|)^{n+1}x^\alpha\partial^\beta f\|_u.$$

It then follows that $\|\widehat{f}\|_{(N,\beta)} \le C_{N,\beta}\sum_{|\gamma|\le|\beta|}\|f\|_{(N+n+1,\gamma)}$ by the proof of Proposition 8.3, so the Fourier transform is continuous on $\mathcal{S}$. ∎

At this point we need to compute an important specific Fourier transform.

8.24 Proposition. *If $f(x) = e^{-\pi a|x|^2}$ where $a > 0$, then $\widehat{f}(\xi) = a^{-n/2}e^{-\pi|\xi|^2/a}$.*

Proof. First consider the case $n = 1$. Since the derivative of $e^{-\pi ax^2}$ is $-2\pi ae^{-\pi ax^2}$, by Theorem 8.22d,e we have

$$(\widehat{f})'(\xi) = (-2\pi ixe^{-\pi ax^2})\widehat{\ }(\xi) = \frac{i}{a}(f')\widehat{\ }(\xi) = \frac{i}{a}(2\pi i\xi)\widehat{f}(\xi) = -\frac{2\pi}{a}\xi\widehat{f}(\xi).$$

It follows that $(d/d\xi)(e^{\pi\xi^2/a}\widehat{f}(\xi)) = 0$, so that $e^{\pi\xi^2/a}\widehat{f}(\xi)$ is constant. To evaluate the constant, set $\xi = 0$ and use Proposition 2.53:

$$\widehat{f}(0) = \int e^{-\pi ax^2}\,dx = a^{-1/2}.$$

The n-dimensional case follows by Fubini's theorem, since $|x|^2 = \sum_1^n x_j^2$:

$$\begin{aligned}\widehat{f}(\xi) &= \prod_1^n \int \exp(-\pi ax_j^2 - 2\pi i\xi_j x_j)\,dx_j \\ &= \prod_1^n \left[a^{-1/2}\exp(-\pi\xi_j^2/a)\right] = a^{-n/2}\exp(-\pi|\xi|^2/a).\end{aligned}$$

∎

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

$$f^\vee(x) = \widehat{f}(-x) = \int f(\xi)e^{2\pi i\xi\cdot x}\,d\xi,$$

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

$$(\widehat{f})^\vee(x) = \iint f(y)\,e^{-2\pi i\xi\cdot y}e^{2\pi i\xi\cdot x}\,dy\,d\xi$$

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

8.25 Lemma. *If $f, g \in L^1$ then $\int \widehat{f}g = \int f\widehat{g}$.*

Proof. Both integrals are equal to $\iint f(x)g(\xi)e^{-2\pi i\xi\cdot x}\,dx\,d\xi$. ∎

8.26 The Fourier Inversion Theorem. *If $f \in L^1$ and $\widehat{f} \in L^1$, then f agrees almost everywhere with a continuous function f_0, and $(\widehat{f})^\vee = (f^\vee)\widehat{\ } = f_0$.*

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

$$\phi(\xi) = \exp(2\pi i\xi \cdot x - \pi t^2|\xi|^2).$$

By Theorem 8.22a and Proposition 8.24,

$$\widehat{\phi}(y) = t^{-n}\exp(-\pi|x-y|^2/t^2) = g_t(x-y),$$

where $g(x) = e^{-\pi|x|^2}$ and the subscript t has the meaning in (8.13). By Lemma 8.25, then,

$$\int e^{-\pi t^2|\xi|^2}e^{2\pi i\xi\cdot x}\widehat{f}(\xi)\,d\xi = \int \widehat{f}\phi = \int f\widehat{\phi} = f * g_t(x).$$

Since $\int e^{-\pi|x|^2}\,dx = 1$, by Theorem 8.14 we have $f * g_t \to f$ in the L^1 norm as $t \to 0$. On the other hand, since $\widehat{f} \in L^1$ the dominated convergence theorem yields

$$\lim_{t\to 0}\int e^{-\pi t^2|\xi|^2}e^{2\pi i\xi\cdot x}\widehat{f}(\xi)\,d\xi = \int e^{2\pi i\xi\cdot x}\widehat{f}(\xi)\,d\xi = (\widehat{f})^\vee(x).$$

It follows that $f = (\widehat{f})^\vee$ a.e., and similarly $(f^\vee)\widehat{\ } = f$ a.e. Since $(\widehat{f})^\vee$ and $(f^\vee)\widehat{\ }$ are continuous, being Fourier transforms of L^1 functions, the proof is complete. ∎

8.27 Corollary. *If $f \in L^1$ and $\widehat{f} = 0$, then $f = 0$ a.e.*

8.28 Corollary. *$\mathcal{F}$ is an isomorphism of $\mathcal{S}$ onto itself.*

Proof. By Corollary 8.23, $\mathcal{F}$ maps $\mathcal{S}$ continuously into itself, and hence so does $f \mapsto f^\vee$, since $f^\vee(x) = \widehat{f}(-x)$. By the Fourier inversion theorem, these maps are inverse to each other. ∎

At last we are in a position to derive the analogue of Theorem 8.20 on $\mathbb{R}^n$.

8.29 The Plancherel Theorem. *If $f \in L^1 \cap L^2$, then $\widehat{f} \in L^2$; and $\mathcal{F}|(L^1 \cap L^2)$ extends uniquely to a unitary isomorphism on L^2.*

Proof. Let $\mathcal{X} = \{f \in L^1 : \widehat{f} \in L^1\}$. Since $\widehat{f} \in L^1$ implies $f \in L^\infty$, we have $\mathcal{X} \subset L^2$ by Proposition 6.10, and $\mathcal{X}$ is dense in L^2 because $\mathcal{S} \subset \mathcal{X}$ and $\mathcal{S}$ is dense in L^2 by Proposition 8.17. Given $f, g \in \mathcal{X}$, let $h = \overline{\widehat{g}}$. By the inversion theorem,

$$\widehat{h}(\xi) = \int e^{-2\pi i\xi\cdot x}\overline{\widehat{g}(x)}\,dx = \int \overline{e^{2\pi i\xi\cdot x}\widehat{g}(x)}\,dx = \overline{g(\xi)}.$$

Hence, by Lemma 8.25,

$$\int f\overline{g} = \int f\widehat{h} = \int \widehat{f}h = \widehat{f}\overline{\widehat{g}}.$$

Thus $\mathcal{F}|\mathcal{X}$ preserves the L^2 inner product; in particular, by taking $g = f$, we obtain $\|\widehat{f}\|_2 = \|f\|_2$. Since $\mathcal{F}(\mathcal{X}) = \mathcal{X}$ by the inversion theorem, $\mathcal{F}|\mathcal{X}$ extends by continuity to a unitary isomorphism on L^2.

It remains only to show that this extension agrees with $\mathcal{F}$ on all of $L^1 \cap L^2$. But if $f \in L^1 \cap L^2$ and $g(x) = e^{-\pi|x|^2}$ as in the proof of the inversion theorem, we have $f * g_t \in L^1$ by Young's inequality and $(f * g_t)\widehat{} \in L^1$ because $(f * g_t)\widehat{}(\xi) = e^{-\pi t^2|\xi|^2}\widehat{f}(\xi)$ and $\widehat{f}$ is bounded. Hence $f * g_t \in \mathcal{X}$; moreover, by Theorem 8.14, $f * g_t \to f$ in both the L^1 and L^2 norms. Therefore $(f * g_t)\widehat{} \to \widehat{f}$ both uniformly and in the L^2 norm, and we are done. ∎

We have thus extended the domain of the Fourier transform from L^1 to $L^1 + L^2$. Just as on $\mathbb{T}^n$, the Riesz-Thorin theorem yields the following result for the intermediate L^p spaces:

8.30 The Hausdorff-Young Inequality. *Suppose that $1 \le p \le 2$ and q is the conjugate exponent to p. If $f \in L^p(\mathbb{R}^n)$, then $\widehat{f} \in L^q(\mathbb{R}^n)$ and $\|\widehat{f}\|_q \le \|f\|_p$.*

If $f \in L^1$ and $\widehat{f} \in L^1$, the inversion formula

$$f(x) = \int \widehat{f}(\xi)e^{2\pi i\xi\cdot x}\,d\xi$$

exhibits f as a superposition of the basic functions $e^{2\pi i\xi\cdot x}$; it is often called the **Fourier integral** representation of f. This formula remains valid in spirit for all $f \in L^2$, although the integral (as well as the integral defining $\widehat{f}$) may not converge pointwise. The interpretation of the inversion formula will be studied further in the next section.

We conclude this section with a beautiful theorem that involves an interplay of Fourier series and Fourier integrals. To motivate it, consider the following problem: Given a function $f \in L^1(\mathbb{R}^n)$, how can one manufacture a periodic function (that is, a function on $\mathbb{T}^n$) from it? Two possible answers suggest themselves. One way is to "average" f over all periods, producing the series $\sum_{k\in\mathbb{Z}^n} f(x-k)$. This series, if it converges, will surely define a periodic function. The other way is to restrict f to the lattice $\mathbb{Z}^n$ and use it to form a Fourier series $\sum_{\kappa\in\mathbb{Z}^n}\widehat{f}(\kappa)e^{2\pi i\kappa\cdot x}$. The content of the following theorem is that these methods both work and both give the same answer.

8.31 Theorem. *If $f \in L^1(\mathbb{R}^n)$, the series $\sum_{k\in\mathbb{Z}^n}\tau_k f$ converges pointwise a.e. and in $L^1(\mathbb{T}^n)$ to a function Pf such that $\|Pf\|_1 \le \|f\|_1$. Moreover, for $\kappa \in \mathbb{Z}^n$, $(Pf)\widehat{}(\kappa)$ (Fourier transform on $\mathbb{T}^n$) equals $\widehat{f}(\kappa)$ (Fourier transform on $\mathbb{R}^n$).*

Proof. Let $Q = [-\frac{1}{2}, \frac{1}{2})^n$. Then $\mathbb{R}^n$ is the disjoint union of the cubes $Q + k = \{x + k : x \in Q\}$, $k \in \mathbb{Z}^n$, so

$$\int_Q \sum_{k\in\mathbb{Z}^n}|f(x-k)|\,dx = \sum_{k\in\mathbb{Z}^n}\int_{Q+k}|f(x)|\,dx = \int_{\mathbb{R}^n}|f(x)|\,dx.$$

Now apply Theorem 2.25. First, it shows that the series $\sum \tau_k f$ converges a.e. and in $L^1(\mathbb{T}^n)$ to a function $Pf \in L^1(\mathbb{T}^n)$ such that $\|Pf\|_1 \leq \|f\|_1$, since $\mathbb{T}^n$ is measure-theoretically identical to Q. Second, it yields

$$\begin{aligned}(Pf)\widehat{\ }(\kappa) &= \int_Q \sum_{k\in\mathbb{Z}^n} f(x-k)e^{-2\pi i\kappa\cdot x}\,dx = \sum_{k\in\mathbb{Z}^n}\int_{Q+k} f(x)e^{-2\pi i\kappa\cdot(x+k)}\,dx \\ &= \sum_{k\in\mathbb{Z}^N}\int_{Q+k} f(x)e^{-2\pi i\kappa\cdot x}\,dx = \int_{\mathbb{R}^n} f(x)e^{-2\pi i\kappa\cdot x}\,dx = \widehat{f}(\kappa).\end{aligned}$$

∎

If we impose conditions on f to guarantee that the series in question converge absolutely, we obtain a more refined result.

8.32 The Poisson Summation Formula. *Suppose $f \in C(\mathbb{R}^n)$ satisfies $|f(x)| \leq C(1+|x|)^{-n-\epsilon}$ and $|\widehat{f}(\xi)| \leq C(1+|\xi|)^{-n-\epsilon}$ for some $C, \epsilon > 0$. Then*

$$\sum_{k\in\mathbb{Z}^n} f(x+k) = \sum_{\kappa\in\mathbb{Z}^n} \widehat{f}(\kappa)e^{2\pi i\kappa\cdot x},$$

where both series converge absolutely and uniformly on $\mathbb{T}^n$. In particular, taking $x = 0$,

$$\sum_{k\in\mathbb{Z}^n} f(k) = \sum_{\kappa\in\mathbb{Z}^n} \widehat{f}(\kappa).$$

Proof. The absolute and uniform convergence of the series follows from the fact that $\sum_{k\in\mathbb{Z}^n}(1+|k|)^{-n-\epsilon} < \infty$, which can be seen by comparing the latter series to the convergent integral $\int(1+|x|)^{-n-\epsilon}\,dx$. Thus the function $Pf = \sum_k \tau_k f$ is in $C(\mathbb{T}^n)$ and hence in $L^2(\mathbb{T}^n)$, so Theorem 8.35 implies that the series $\sum \widehat{f}(\kappa)e^{2\pi i\kappa\cdot x}$ converges in $L^2(\mathbb{T}^n)$ to Pf. Since it also converges uniformly, its sum equals Pf pointwise. (The replacement of k by $-k$ in the formula for Pf is immaterial since the sum is over all $k \in \mathbb{Z}^n$.) ∎

Exercises

12. Work out the analogue of Theorem 8.22 for the Fourier transform on $\mathbb{T}^n$.

13. Let $f(x) = \frac{1}{2} - x$ on the interval $[0,1)$, and extend f to be periodic on $\mathbb{R}$.

a. $\widehat{f}(0) = 0$, and $\widehat{f}(\kappa) = (2\pi i\kappa)^{-1}$ if $\kappa \neq 0$.

b. $\sum_1^\infty k^{-2} = \pi^2/6$. (Use the Parseval identity.)

14. (**Wirtinger's Inequality**) If $f \in C^1([a,b])$ and $f(a) = f(b) = 0$, then

$$\int_a^b |f(x)|^2\,dx \leq \left(\frac{b-a}{\pi}\right)^2 \int_a^b |f'(x)|^2\,dx.$$

(By a change of variable it suffices to assume $a = 0$, $b = \frac{1}{2}$. Extend f to $[-\frac{1}{2}, \frac{1}{2}]$ by setting $f(-x) = -f(x)$, and then extend f to be periodic on $\mathbb{R}$. Check that f, thus extended, is in $C^1(\mathbb{T})$ and apply the Parseval identity.)

15. Let $\operatorname{sinc} x = (\sin \pi x)/\pi x$ ($\operatorname{sinc} 0 = 1$).
a. If $a > 0$, $\widehat{\chi}_{[-a,a]}(x) = \chi^{\vee}_{[a,a]}(x) = 2a \operatorname{sinc} 2ax$.
b. Let $\mathcal{H}_a = \{f \in L^2 : \widehat{f}(\xi) = 0$ (a.e.) for $|\xi| > a\}$. Then $\mathcal{H}$ is a Hilbert space and $\{\sqrt{2a} \operatorname{sinc}(2ax - k) : k \in \mathbb{Z}\}$ is an orthonormal basis for $\mathcal{H}$.
c. (**The Sampling Theorem**) If $f \in \mathcal{H}_a$, then $f \in C_0$ (after modification on a null set), and $f(x) = \sum_{-\infty}^{\infty} f(k/2a) \operatorname{sinc}(2ax - k)$, where the series converges both uniformly and in L^2. (In the terminology of signal analysis, a signal of bandwidth $2a$ is completely determined by sampling its values at a sequence of points $\{k/2a\}$ whose spacing is the reciprocal of the bandwidth.)

16. Let $f_k = \chi_{[-1,1]} * \chi_{[-k,k]}$.
a. Compute $f_k(x)$ explicitly and show that $\|f\|_u = 2$.
b. $f_k^{\vee}(x) = (\pi x)^{-2} \sin 2\pi kx \sin 2\pi x$, and $\|f_k^{\vee}\|_1 \to \infty$ as $k \to \infty$. (Use Exercise 15a, and substitute $y = 2\pi kx$ in the integral defining $\|f_k^{\vee}\|_1$.)
c. $\mathcal{F}(L^1)$ is a proper subset of C_0. (Consider $g_k = f_k^{\vee}$ and use the open mapping theorem.)

17. Given $a > 0$, let $f(x) = e^{-2\pi x} x^{a-1}$ for $x > 0$ and $f(x) = 0$ for $x \le 0$.
a. $f \in L^1$, and $f \in L^2$ if $a > \frac{1}{2}$.
b. $\widehat{f}(\xi) = \Gamma(a)[(2\pi)(1 + i\xi)]^{-a}$. (Here we are using the branch of z^a in the right half plane that is positive when z is positive. Cauchy's theorem may be used to justify the complex substitution $y = (1 + i\xi)x$ in the integral defining $\widehat{f}$.)
c. If $a, b > \frac{1}{2}$ then

$$\int_{-\infty}^{\infty} (1 - ix)^{-a}(1 + ix)^{-b}\, dx = \frac{2^{2-a-b}\pi\Gamma(a + b - 1)}{\Gamma(a)\Gamma(b)}.$$

18. Suppose $f \in L^2(\mathbb{R})$.
a. The L^2 derivative f' (in the sense of Exercises 8 and 9) exists iff $\xi\widehat{f} \in L^2$, in which case $\widehat{f'}(\xi) = 2\pi i \xi \widehat{f}(\xi)$.
b. If the L^2 derivative f' exists, then

$$\left[\int |f(x)|^2\, dx\right] \le 4 \int |xf(x)|^2\, dx \int |f'(x)|^2\, dx.$$

(If the integrals on the right are finite, one can integrate by parts to obtain $\int |f|^2 = -2 \operatorname{Re} \int x\overline{f} f'$.)
c. (**Heisenberg's Inequality**) For any $b, \beta \in \mathbb{R}$,

$$\int (x - b)^2 |f(x)|^2\, dx \int (\xi - \beta)^2 |\widehat{f}(\xi)|^2\, d\xi \ge \frac{\|f\|_2^4}{16\pi^2}.$$

(The inequality is trivial if either integral on the right is infinite; if not, reduce to the case $b = \beta = 0$ by considering $g(x) = e^{-2\pi i\beta x} f(x+b)$.) This inequality, a form of the quantum uncertainty principle, says that f and $\widehat{f}$ cannot both be sharply localized about single points b and β.

19. (A variation on the theme of Exercise 18) If $f \in L^2(\mathbb{R}^n)$ and the set $S = \{x : f(x) \neq 0\}$ has finite measure, then for any measurable $E \subset \mathbb{R}^n$, $\int_E |\widehat{f}|^2 \leq \|f\|_2^2 m(S) m(E)$.

20. If $f \in L^1(\mathbb{R}^{n+m})$, define $Pf(x) = \int f(x,y)\,dy$. (Here $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.) Then $Pf \in L^1(\mathbb{R}^n)$, $\|Pf\|_1 \leq \|f\|_1$, and $(Pf)\widehat{\ }(\xi) = \widehat{f}(\xi, 0)$.

21. State and prove a result that encompasses both Theorem 8.31 and Exercise 20, in the setting of Fourier transforms on closed subgroups and quotient groups of $\mathbb{R}^n$.

22. Since $\mathcal{F}$ commutes with rotations, the Fourier transform of a radial function is radial; that is, if $F \in L^1(\mathbb{R}^n)$ and $F(x) = f(|x|)$, then $\widehat{F}(\xi) = g(|\xi|)$, where f and g are related as follows.

a. Let $J(\xi) = \int_S e^{ix\xi} d\sigma(x)$ where σ is surface measure on the unit sphere S in $\mathbb{R}^n$ (Theorem 2.49). Then J is radial — say, $J(\xi) = j(|\xi|)$ — and $g(\rho) = \int_0^\infty j(2\pi r\rho) f(r) r^{n-1}\,dr$.

b. J satisfies $\sum_1^n \partial_k^2 J + J = 0$.

c. j satisfies $\rho j''(\rho) + (n-1)j'(\rho) + \rho j(\rho) = 0$. (This equation is a variant of Bessel's equation. The function j is completely determined by the fact that it is a solution of this equation, is smooth at $\rho = 0$, and satisfies $j(0) = \sigma(S) = 2\pi^{n/2}/\Gamma(n/2)$. In fact, $j(\rho) = (2\pi)^{n/2}\rho^{(2-n)/2} J_{(n-2)/2}(\rho)$ where J_α is the Bessel function of the first kind of order α.)

d. If $n = 3$, $j(\rho) = 4\pi\rho^{-1}\sin\rho$. (Set $f(\rho) = \rho j(\rho)$ and use (c) to show that $f'' + f = 0$. Alternatively, use spherical coordinates to compute the integral defining $J(0,0,\rho)$ directly.)

23. In this exercise we develop the theory of Hermite functions.

a. Define operators T, T^* on $\mathcal{S}(\mathbb{R})$ by $Tf(x) = 2^{-1/2}[xf(x) - f'(x)]$ and $T^*f(x) = 2^{-1/2}[xf(x) + f'(x)]$. Then $\int (Tf)\overline{g} = \int f(\overline{T^*g})$ and $T^*T^k - T^kT^* = kT^{k-1}$.

b. Let $h_0(x) = \pi^{-1/4}e^{-x^2/2}$, and for $k \geq 1$ let $h_k = (k!)^{-1/2}T^k h_0$. (h_k is the kth **normalized Hermite function**.) We have $Th_k = \sqrt{k+1}\,h_{k+1}$ and $T^*h_k = \sqrt{k}\,h_{k-1}$, and hence $TT^*h_k = kh_k$.

c. Let $S = 2TT^* + I$. Then $Sf(x) = x^2 f(x) - f''(x)$ and $Sh_k = (2k+1)h_k$. (S is called the **Hermite operator**.)

d. $\{h_k\}_0^\infty$ is an orthonormal set in $L^2(\mathbb{R})$. (Check directly that $\|h_0\|_2 = 1$, then observe that for $k > 0$, $\int h_k \overline{h}_m = k^{-1}\int (TT^*h_k)\overline{h}_m$ and use (a) and (b).)

e. We have

$$T^k f(x) = (-1)^k 2^{-k/2} e^{x^2/2} \left(\frac{d}{dx}\right)^k \left[e^{-x^2/2} f(x)\right]$$

(use induction on k), and in particular,

$$h_k(x) = \frac{(-1)^k}{[\pi^{1/2}2^k k!]^{1/2}} e^{x^2/2} \left(\frac{d}{dx}\right)^k e^{-x^2}.$$

f. Let $H_k(x) = e^{x^2/2}h_k(x)$. Then H_k is a polynomial of degree k, called the kth **normalized Hermite polynomial**. The linear span of $H_0, \ldots, H_m$ is the set of all polynomials of degree $\leq m$. (The kth Hermite polynomial as usually defined is $[\pi^{1/2}2^k k!]^{1/2}H_k$.)

g. $\{h_k\}_0^\infty$ is an orthonormal basis for $L^2(\mathbb{R})$. (Suppose $f \perp h_k$ for all k, and let $g(x) = f(x)e^{-x^2/2}$. Show that $\widehat{g} = 0$ by expanding $e^{-2\pi i\xi\cdot x}$ in its Maclaurin series and using (f).)

h. Define $A : L^2 \to L^2$ by $Af(x) = (2\pi)^{1/4}f(x\sqrt{2\pi})$, and define $\widetilde{f} = A^{-1}\mathcal{F}Af$ for $f \in L^2$. Then A is unitary and $\widetilde{f}(\xi) = (2\pi)^{-1/2}\int f(x)e^{-i\xi x}\,dx$. Moreover, $\widetilde{Tf} = -iT(\widetilde{f})$ for $f \in \mathcal{S}$, and $\widetilde{h}_0 = h_0$; hence $\widetilde{h}_k = (-i)^k h_k$. Therefore, if $\phi_k = Ah_k$, $\{\phi_k\}_0^\infty$ is an orthonormal basis for L^2 consisting of eigenfunctions for $\mathcal{F}$; namely, $\widehat{\phi}_k = (-i)^k\phi_k$.

8.4 SUMMATION OF FOURIER INTEGRALS AND SERIES

The Fourier inversion theorem shows how to express a function f on $\mathbb{R}^n$ in terms of $\widehat{f}$ provided that f and $\widehat{f}$ are in L^1. The same result holds for periodic functions. Namely, if $f \in L^1(\mathbb{T}^n)$ and $\widehat{f} \in l^1(\mathbb{Z}^n)$, then the Fourier series $\sum_\kappa \widehat{f}(\kappa)e^{2\pi i\kappa\cdot x}$ converges absolutely and uniformly to a function g. Since $l^1 \subset l^2$, it follows from Theorem 8.20 that $f \in L^2$ and that the series converges to f in the L^2 norm. Hence $f = g$ a.e., and $f = g$ everywhere if f is assumed continuous at the outset.

Two questions therefore arise. What conditions on f will guarantee that $\widehat{f}$ is integrable? And how can f be recovered from $\widehat{f}$ if $\widehat{f}$ is not integrable?

As for the first question, since $\widehat{f}$ is bounded for $f \in L^1$, the issue is the decay of $\widehat{f}$ at infinity, and this is related to the smoothness properties of f. For example, by Theorem 8.22e, if $f \in C^{n+1}(\mathbb{R}^n)$ and $\partial^\alpha f \in L^1 \cap C_0$ for $|\alpha| \leq n+1$, then $|\widehat{f}(\xi)| \leq C(1+|\xi|)^{-n-1}$ and hence $\widehat{f} \in L^1(\mathbb{R}^n)$ by Corollary 2.52. The same result holds for periodic functions, for the same reason: If $f \in C^{n+1}(\mathbb{T}^n)$, then $|\widehat{f}(\kappa)| \leq C(1+|\kappa|)^{-n-1}$ and hence $\widehat{f} \in l^1(\mathbb{Z}^n)$.

To obtain sharper results when $n > 1$ requires a generalized notion of partial derivatives, so we shall postpone this task until §9.3. (See Theorem 9.17.) However, for $n = 1$ we can easily obtain a better theorem that covers the useful case of functions that are continuous and piecewise C^1. We state it for periodic functions and leave the nonperiodic case to the reader (Exercise 24).

8.33 Theorem. *Suppose that f is periodic and absolutely continuous on $\mathbb{R}$, and that $f' \in L^p(\mathbb{T})$ for some $p > 1$. Then $\widehat{f} \in l^1(\mathbb{Z})$.*

Proof. Since $p > 1$, we have $C_p = \sum_1^\infty \kappa^{-p} < \infty$; and since $L^p(\mathbb{T}) \subset L^2(\mathbb{T})$ for $p > 2$, we may assume that $p \le 2$. Integration by parts (Theorem 3.36) shows that $(f')\widehat{\ }(\kappa) = 2\pi i\kappa \widehat{f}(\kappa)$. Hence, by the inequalities of Hölder and Hausdorff-Young, if q is the conjugate exponent to p,

$$\begin{aligned}\sum_{\kappa\neq 0} |\widehat{f}(\kappa)| &\le \Big[\sum_{\kappa\neq 0}(2\pi|\kappa|)^{-p}\Big]^{1/p}\Big[\sum_{\kappa\neq 0}(2\pi|\kappa\widehat{f}(\kappa)|)^q\Big]^{1/q} \\ &= \frac{(2C_p)^{1/p}}{2\pi}\|(f')\widehat{\ }\|_q \le \frac{(2C_p)^{1/p}}{2\pi}\|f'\|_p.\end{aligned}$$

Adding $|\widehat{f}(0)|$ to both sides, we see that $\|\widehat{f}\|_1 < \infty$. ∎

We now turn to the problem of recovering f from $\widehat{f}$ under minimal hypotheses on f, and we consider first the case of $\mathbb{R}^n$. The proof of the Fourier inversion theorem contains the essential idea: Replace the divergent integral $\int \widehat{f}(\xi)e^{2\pi i\xi\cdot x}\,d\xi$ by $\int \widehat{f}(\xi)\Phi(t\xi)e^{2\pi i\xi\cdot x}\,d\xi$ where Φ is a continuous function that vanishes rapidly enough at infinity to make the integral converge. If we choose Φ to satisfy $\Phi(0) = 1$, then $\Phi(t\xi) \to 1$ as $t \to 0$, and with any luck the corresponding integral will converge to f in some sense. One Φ that works is the function $\Phi(\xi) = e^{-\pi|\xi|^2}$ used in the proof of the inversion theorem, but we shall see below that there are others of independent interest. We therefore formulate a fairly general theorem, for which we need the following lemma that complements Theorem 8.22c.

8.34 Lemma. *If $f, g \in L^2(\mathbb{R}^n)$, then $(\widehat{f}\widehat{g})^\vee = f * g$.*

Proof. $\widehat{f}\widehat{g} \in L^1$ by Plancherel's theorem and Hölder's inequality, so $(\widehat{f}\widehat{g})^\vee$ makes sense. Given $x \in \mathbb{R}^n$, let $h(y) = \overline{g(x-y)}$. It is easily verified that $\widehat{h}(\xi) = \overline{\widehat{g}(\xi)}e^{-2\pi i\xi\cdot x}$, so since $\mathcal{F}$ is unitary on L^2,

$$f * g(x) = \int f\overline{h} = \int \widehat{f}\overline{\widehat{h}} = \int \widehat{f}(\xi)\widehat{g}(\xi)e^{2\pi i\xi\cdot x}\,d\xi = (\widehat{f}\widehat{g})^\vee(x).$$

∎

8.35 Theorem. *Suppose that $\Phi \in L^1 \cap C_0$, $\Phi(0) = 1$, and $\phi = \Phi^\vee \in L^1$. Given $f \in L^1 + L^2$, for $t > 0$ set*

$$f^t(x) = \int \widehat{f}(\xi)\Phi(t\xi)e^{2\pi i\xi\cdot x}\,d\xi.$$

- a. *If $f \in L^p$ $(1 \le p < \infty)$, then $f^t \in L^p$ and $\|f^t - f\|_p \to 0$ as $t \to 0$.*
- b. *If f is bounded and uniformly continuous, then so is f^t, and $f^t \to f$ uniformly as $t \to 0$.*
- c. *Suppose also that $|\phi(x)| \le C(1+|x|)^{-n-\epsilon}$ for some $C, \epsilon > 0$. Then $f^t(x) \to f(x)$ for every x in the Lebesgue set of f.*

Proof. We have $f = f_1 + f_2$ where $f_1 \in L^1$ and $f_2 \in L^2$. Since $\widehat{f_1} \in L^\infty$, $\widehat{f_2} \in L_2$, and $\Phi \in (L^1 \cap C_0) \subset (L^1 \cap L^2)$, the integral defining f^t converges absolutely for every x. Moreover, if $\phi_t(x) = t^{-n}\phi(t^{-1}x)$, we have $\Phi(t\xi) = (\phi_t)\widehat{\ }(\xi)$ by the inversion theorem and Theorem 8.22b, and $\int \phi(x)\,dx = \Phi(0) = 1$. Since $\phi, \Phi \in L^1$ we have $f_1 * \phi \in L^1$ and $\widehat{f_1}\Phi \in L^1$, so by Theorem 8.22c and the inversion formula,

$$\int \widehat{f_1}(\xi)\Phi(t\xi)e^{2\pi i\xi\cdot x}\,d\xi = f_1 * \phi_t(x).$$

Also, $\phi \in L^2$ by the Plancherel theorem, so by Lemma 8.34,

$$\int \widehat{f_2}(\xi)\Phi(t\xi)e^{2\pi i\xi\cdot x}\,d\xi = f_2 * \phi_t(\xi).$$

In short, $f^t = f * \phi_t$, so the assertions follow from Theorems 8.14 and 8.15. ∎

By combining this theorem with the Poisson summation formula, we obtain a corresponding result for periodic functions.

8.36 Theorem. *Suppose that $\Phi \in C(\mathbb{R}^n)$ satisfies $|\Phi(\xi)| \le C(1 + |\xi|)^{-n-\epsilon}$, $|\Phi^\vee(x)| \le C(1 + |x|)^{-n-\epsilon}$, and $\Phi(0) = 1$. Given $f \in L^1(\mathbb{T}^n)$, for $t > 0$ set*

$$f^t(x) = \sum_{\kappa\in\mathbb{Z}^n} \widehat{f}(\kappa)\Phi(t\kappa)e^{2\pi i\kappa\cdot x}$$

(which converges absolutely since $\sum_\kappa |\Phi(t\kappa)| < \infty$).

a. If $f \in L^p(\mathbb{T}^n)$ $(1 \le p < \infty)$, then $\|f^t - f\|_p \to 0$ as $t \to 0$, and if $f \in C(\mathbb{T}^n)$, then $f^t \to f$ uniformly as $t \to 0$.

b. $f^t(x) \to f(x)$ for every x in the Lebesgue set of f.

Proof. Let $\phi = \Phi^\vee$ and $\phi_t(x) = t^{-n}\phi(t^{-1}x)$. Then $(\phi_t)\widehat{\ }(\xi) = \Phi(t\xi)$, and ϕ_t satisfies the hypotheses of the Poisson summation formula, so

$$\sum_{k\in\mathbb{Z}^n} \phi_t(x - k) = \sum_{k\in\mathbb{Z}^n} \Phi(t\kappa)e^{2\pi i\kappa\cdot x}.$$

Let us denote the common value of these sums by $\psi_t(x)$. Then

$$(f * \psi_t)\widehat{\ }(\kappa) = \widehat{f}(\kappa)\widehat{\psi_t}(\kappa) = \widehat{f}(\kappa)\Phi(t\kappa) = (f^t)\widehat{\ }(\kappa),$$

so $f^t = f * \psi_t$. Hence, by Young's inequality and Theorem 8.31 we have

$$\|f^t\|_p \le \|f\|_p\|\psi_t\|_1 \le \|f\|_p\|\phi_t\|_1 = \|f\|_p\|\phi\|_1,$$

so the operators $f \mapsto f^t$ are uniformly bounded on L^p, $1 \le p \le \infty$.

Now, since Φ is continuous and $\Phi(0) = 1$, we clearly have $f^t \to f$ uniformly (and hence in $L^p(\mathbb{T}^n)$) if f is a trigonometric polynomial — that is, if $\widehat{f}(\kappa) = 0$ for all but finitely many κ. But the trigonometric polynomials are dense in $C(\mathbb{T}^n)$ in the uniform norm by the Stone-Weierstrass theorem, and hence also dense in $L^p(\mathbb{T}^n)$ in the L^p norm for $p < \infty$. Assertion (a) therefore follows from Proposition 5.17.

To prove (b), suppose that x is in the Lebesgue set of f; by translating f we may assume that $x = 0$, which simplifies the notation. With $Q = [-\frac{1}{2}, \frac{1}{2})^n$, we have

$$\begin{aligned} f^t(0) = f * \psi_t(0) &= \int_Q f(x)\psi_t(-x)\,dx \\ &= \int_Q f(x)\phi_t(-x)\,dx + \sum_{k\neq 0}\int_Q f(x)\phi_t(-x+k)\,dx. \end{aligned}$$

Since

$$|\phi_t(x)| \leq Ct^{-n}(1+t^{-1}|x|)^{-n-\epsilon} \leq Ct^{\epsilon}|x|^{-n-\epsilon},$$

for $x \in Q$ and $k \neq 0$ we have $|\phi_t(-x+k)| \leq C2^{n+\epsilon}t^{\epsilon}|k|^{-n-\epsilon}$, and hence

$$\sum_{k\neq 0}\int_Q |f(x)\phi_t(-x+k)|\,dx \leq \Big[C2^{n+\epsilon}\|f\|_1 \sum_{k\neq 0}|k|^{-n-\epsilon}\Big]t^{\epsilon},$$

which vanishes as $t \to 0$. On the other hand, if we define $g = f\chi_Q \in L^1(\mathbb{R}^n)$, then 0 is in the Lebesgue set of g (because 0 is in the interior of Q, and the condition that 0 be in the Lebesgue set of g depends only on the behavior of g near 0), so by Theorem 8.15,

$$\lim_{t\to 0}\int_Q f(x)\phi_t(-x)\,dx = \lim_{t\to 0} g * \phi_t(0) = g(0) = f(0).$$

∎

Let us examine some specific examples of functions Φ that can be used in Theorems 8.35 and 8.36. The first is the one already used in the proof of the inversion theorem,

$$\Phi(\xi) = e^{-\pi|\xi|^2}, \qquad \phi(x) = \Phi^{\vee}(x) = e^{-\pi|x|^2}.$$

This ϕ is called the **Gauss kernel** or **Weierstrass kernel**. It is important for a number of reasons, including its connection with the heat equation that we shall explain in §8.7. When $n = 1$, its periodized version

$$\psi_t(x) = \frac{1}{t}\sum_{k\in\mathbb{Z}} e^{-\pi|x-k|^2/t^2} = \sum_{\kappa\in\mathbb{Z}} e^{-\pi t^2\kappa^2}e^{2\pi i\kappa\cdot x},$$

in terms of which the f^t in Theorem 8.36 is given by $f^t = f * \psi_t$, is essentially one of the Jacobi theta functions, which are connected with elliptic functions and have applications in number theory.

The second example is $\Phi(\xi) = e^{-2\pi|\xi|}$, whose inverse Fourier transform ϕ is called the **Poisson kernel** on $\mathbb{R}^n$. When $n = 1$, we have

$$\begin{aligned} \phi(x) &= \int_{-\infty}^0 e^{2\pi(1+ix)\xi}\,d\xi + \int_0^{\infty} e^{2\pi(-1+ix)\xi}\,d\xi \\ &= \frac{1}{2\pi}\left[\frac{1}{1+ix} + \frac{1}{1-ix}\right] = \frac{1}{\pi(1+x^2)}. \end{aligned} \tag{8.37}$$

The formula for ϕ in higher dimensions is worked out in Exercise 26; it turns out that $\phi(x)$ is a constant multiple of $(1+|x|^2)^{-(n+1)/2}$. Like the Gauss kernel, the Poisson kernel has an interpretation in terms of partial differential equations that we shall explain in §8.7.

If we take $n = 1$ and $\Phi(\xi) = e^{-2\pi|\xi|}$ in Theorem 8.36, make the substitution $r = e^{-2\pi t}$, and write $A_r f$ in place of f^t, we obtain

$$
\begin{aligned}
A_r f(x) &= \sum_{\kappa\in\mathbb{Z}} r^{|\kappa|}\widehat{f}(\kappa)e^{-2\pi i\kappa x} \\
&= \widehat{f}(0) + \sum_{k=1}^{\infty} r^k\big[\widehat{f}(k)e^{2\pi ikx} + \widehat{f}(-k)e^{-2\pi ikx}\big].
\end{aligned}
\tag{8.38}
$$

This formula is a special case of one of the classical methods for summing a (possibly) divergent series. Namely, if $\sum_0^\infty a_k$ is a series of complex numbers, for $0 < r < 1$ its **rth Abel mean** is the series $\sum_0^\infty r^k a_k$. If the latter series converges for $r < 1$ to the sum $S(r)$ and the limit $S = \lim_{r\nearrow 1} S(r)$ exists, the series $\sum_0^\infty a_k$ is said to be **Abel summable** to S. If $\sum_0^\infty a_k$ converges to the sum S, then it is also Abel summable to S (Exercise 27), but the Abel sum may exist even when the series diverges.

In (8.38), $A_r f(x)$ is the rth Abel mean of the Fourier series of f, in which the kth and $(-k)$th terms are grouped together to make a series indexed by the nonnegative integers. It has the following complex-variable interpretation: If we set $z = re^{2\pi ix}$, then

$$A_r f(x) = \sum_0^\infty \widehat{f}(k)z^k + \sum_1^\infty \widehat{f}(-k)\bar{z}^k.$$

The two series on the right define, respectively, a holomorphic and an antiholomorphic function on the unit disc $|z| < 1$. In particular, $A_r f(x)$ is a harmonic function on the unit disc, and the fact that $A_r f \to f$ as $r \to 1$ means that f is the boundary value of this function on the unit circle. See also Exercise 28.

Our final example is the function $\Phi(\xi) = \max(1-|\xi|, 0)$ with $n = 1$. Its inverse Fourier transform is

$$
\begin{aligned}
\phi(x) &= \int_{-1}^{0}(1+\xi)e^{2\pi i\xi\cdot x}\,d\xi + \int_0^1 (1-\xi)e^{2\pi i\xi\cdot x}\,d\xi \\
&= \frac{e^{2\pi ix} + e^{-2\pi ix} - 2}{(2\pi ix)^2} = \left(\frac{\sin \pi x}{\pi x}\right)^2.
\end{aligned}
$$

If we use this Φ in Theorem 8.36, take $t = (m+1)^{-1}$ $(m = 0, 1, 2, \ldots)$, and write $\sigma_m f(x)$ for $f^{1/(m+1)}(x)$, we obtain

$$
\begin{aligned}
\sigma_m f(x) &= \sum_{\kappa=-m}^{m} \frac{m+1-|\kappa|}{m+1}\widehat{f}(\kappa)e^{2\pi i\kappa x} \\
&= \widehat{f}(0) + \sum_{k=1}^{m}\frac{m+1-k}{m+1}\big[\widehat{f}(k)e^{2\pi ikx} + \widehat{f}(-k)e^{-2\pi ikx}\big].
\end{aligned}
\tag{8.39}
$$

This is an instance of another classical method for summing divergent series. Namely, if $\sum_0^\infty a_k$ is a series of complex numbers, its **mth Cesàro mean** is the average of its first $m+1$ partial sums, $(m+1)^{-1}\sum_0^m S_n$, where $S_n = \sum_0^n a_k$. If the sequence of Cesàro means converges as $m \to \infty$ to a limit S, the series is said to be **Cesàro summable** to S. It is easily verified that if $\sum_0^\infty a_k$ converges to S, then it is Cesàro summable to S (but perhaps not conversely), and that $\sigma_m f(x)$ is the mth Cesàro mean of the Fourier series of f with the kth and $(-k)$th terms grouped together. See Exercise 29, and also Exercise 33 in the next section.

Exercises

24. State and prove an analogue of Theorem 8.33 for functions on $\mathbb{R}$. (In addition to the hypotheses that f be locally absolutely continuous and that $f' \in L^p$ for some $p > 1$, you will need some further conditions f and/or f' at infinity to make the argument work. Make them as mild as possible.)

25. For $0 < \alpha \leq 1$, let $\Lambda_\alpha(\mathbb{T})$ be the space of Hölder continuous functions on $\mathbb{T}$ of exponent α as in Exercise 11 in §5.1. Suppose $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$.

a. If f satisfies the hypotheses of Theorem 8.33, then $f \in \Lambda_{1/q}(\mathbb{T})$, but f need not lie in $\Lambda_\alpha(\mathbb{T})$ for any $\alpha > 1/q$. (Hint: $f(b) - f(a) = \int_a^b f'(t)\,dt$.)

b. If $\alpha < 1$, $\Lambda_\alpha(\mathbb{T})$ contains functions that are not of bounded variation and hence are not absolutely continuous. (But cf. Exercise 37 in §3.5.)

26. The aim of this exercise is to show that the inverse Fourier transform of $e^{-2\pi|\xi|}$ on $\mathbb{R}^n$ is

$$\phi(x) = \frac{\Gamma(\frac{1}{2}(n+1))}{\pi^{(n+1)/2}}\left(1+|x|^2\right)^{-(n+1)/2}.$$

a. If $\beta \geq 0$, $e^{-\beta} = \pi^{-1}\int_{-\infty}^{\infty}(1+t^2)^{-1}e^{-i\beta t}\,dt$. (Use (8.37).)

b. If $\beta \geq 0$, $e^{-\beta} = \int_0^\infty (\pi s)^{-1/2}e^{-s}e^{-\beta^2/4s}\,ds$. (Use (a), Proposition 8.24, and the formula $(1+t^2)^{-1} = \int_0^\infty e^{-(1+t^2)s}\,ds$.)

c. Let $\beta = 2\pi|\xi|$ where $\xi \in \mathbb{R}^n$; then the formula in (b) expresses $e^{-2\pi|\xi|}$ as a superposition of dilated Gauss kernels. Use Proposition 8.24 again to derive the asserted formula for ϕ.

27. Suppose that the numerical series $\sum_0^\infty a_k$ is convergent.

a. Let $S_m^n = \sum_m^n a_k$. Then $\sum_m^n r^k a_k = \sum_m^{n-1} S_m^j(r^j - r^{j+1}) + S_m^n r^n$ for $0 \leq r \leq 1$ ("summation by parts").

b. $\left|\sum_m^n r^k a_k\right| \leq \sup_{j \geq m}|S_m^j|$.

c. The series $\sum_0^\infty r^k a_k$ is uniformly convergent for $0 \leq r \leq 1$, and hence its sum $S(r)$ is continuous there. In particular, $\sum_0^\infty a_k = \lim_{r \nearrow 1} S(r)$.

28. Suppose that $f \in L^1(\mathbb{T})$, and let $A_r f$ be given by (8.38).

a. $A_r f = f * P_r$ where $P_r(x) = \sum_{-\infty}^{\infty} r^{|\kappa|}e^{2\pi i\kappa x}$ is the **Poisson kernel** for $\mathbb{T}$.

b. $P_r(x) = (1-r^2)/(1+r^2-2r\cos 2\pi x)$.

29. Given $\{a_k\}_0^\infty \subset \mathbb{C}$, let $S_n = \sum_0^n a_k$ and $\sigma_m = (m+1)^{-1}\sum_0^m S_n$.

a. $\sigma_m = (m+1)^{-1}\sum_0^m (m+1-k)a_k$.
b. If $\lim_{n\to\infty} S_n = \sum_0^\infty a_k$ exists, then so does $\lim_{m\to\infty}\sigma_m$, and the two limits are equal.
c. The series $\sum_0^\infty (-1)^k$ diverges but is Abel and Cesàro summable to $\frac{1}{2}$.

30. If $f \in L^1(\mathbb{R}^n)$, f is continuous at 0, and $\widehat{f} \geq 0$, then $\widehat{f} \in L^1$. (Use Theorem 8.35c and Fatou's lemma.)

31. Suppose $a > 0$. Use (8.37) to show that

$$\sum_{-\infty}^{\infty} \frac{1}{k^2+a^2} = \frac{\pi}{a}\frac{1+e^{-2\pi a}}{1-e^{-2\pi a}}.$$

Then subtract a^{-2} from both sides and let $a \to 0$ to show that $\sum_1^\infty k^{-2} = \pi^2/6$.

32. A C^∞ function f on $\mathbb{R}$ is **real-analytic** if for every $x \in \mathbb{R}$, f is the sum of its Taylor series based at x in some neighborhood of x. If f is periodic and we regard f as a function on $S = \{z \in \mathbb{C} : |z| = 1\}$, this condition is equivalent to the condition that f be the restriction to S of a holomorphic function on some neighborhood of S. Show that $f \in C^\infty(\mathbb{T})$ is real-analytic iff $|\widehat{f}(\kappa)| \leq Ce^{-\epsilon|\kappa|}$ for some $C, \epsilon > 0$. (See the discussion of the Abel means $A_r f$ in the text, and note that $\overline{z} = z^{-1}$ when $|z| = 1$.)

8.5 POINTWISE CONVERGENCE OF FOURIER SERIES

The techniques and results of the previous two sections, involving such things as L^p norms and summability methods, are relatively modern; they were preceded historically by the study of pointwise convergence of one-dimensional Fourier series. Although the latter is one of the oldest parts of Fourier analysis, it is also one of the most difficult — unfortunately for the mathematicians who developed it, but fortunately for us who are the beneficiaries of the ideas and techniques they invented in doing so. A thorough study of this issue is beyond the scope of this book, but we would be remiss not to present a few of the classic results.

To set the stage, suppose $f \in L^1(\mathbb{T})$. We denote by $S_m f$ the mth symmetric partial sum of the Fourier series of f:

$$S_m f(x) = \sum_{-m}^{m} \widehat{f}(k) e^{2\pi i k x}.$$

From the definition of $\widehat{f}(k)$, we have

$$S_m f(x) = \sum_{-m}^{m} \int_0^1 f(y) e^{2\pi i k(x-y)}\, dy = f * D_m(x),$$

where D_m is the mth **Dirichlet kernel**:

$$D_m(x) = \sum_{-m}^{m} e^{2\pi ikx}.$$

The terms in this sum form a geometric progression, so

$$D_m(x) = e^{-2\pi imx} \sum_{0}^{2m} e^{2\pi ikx} = e^{-2\pi imx} \frac{e^{2\pi(2m+1)x} - 1}{e^{2\pi ix} - 1}.$$

Multiplying top and bottom by $e^{-\pi ix}$ yields the standard closed formula for D_m:

$$(8.40) \qquad D_m(x) = \frac{e^{(2m+1)\pi ix} - e^{-(2m+1)\pi ix}}{e^{\pi ix} - e^{-\pi ix}} = \frac{\sin(2m+1)\pi x}{\sin \pi x}.$$

The difficulty with the partial sums $S_m f$, as opposed to (for example) the Abel or Cesàro means, can be summed up in a nutshell as follows. $S_m f$ can be regarded as a special case of the construction in Theorem 8.36; in fact, with the notation used there, $S_m f = f^{1/m}$ if we take $\Phi = \chi_{[-1,1]}$. *But $\chi_{[-1,1]}$ does not satisfy the hypotheses of Theorem 8.36,* because its inverse Fourier transform $(\pi x)^{-1} \sin 2\pi x$ (Exercise 15a) is not in $L^1(\mathbb{R})$. On the level of periodic functions, this is reflected in the fact that although $D_m \in L^1(\mathbb{T})$ for all m, $\|D_m\|_1 \to \infty$ as $m \to \infty$ (Exercise 34).

Among the consequences of this is that the Fourier series of a continuous function f need not converge pointwise, much less uniformly, to f; see Exercise 35. (This does not contradict the fact that trigonometric polynomials are dense in $C(\mathbb{T})$! It just means that if one wants to approximate a function $f \in C(\mathbb{T})$ uniformly by trigonometric polynomials, one should not count on the partial sums $S_m f$ to do the job; the Cesàro means defined by (8.39) work much better in general.) To obtain positive results for pointwise convergence, one must look in other directions.

The first really general theorem about pointwise convergence of Fourier series was obtained in 1829 by Dirichlet, who showed that $S_m f(x) \to \frac{1}{2}[f(x+) + f(x-)]$ for every x provided that f is piecewise continuous and piecewise monotone. Later refinements of the argument showed that what is really needed is for f to be of bounded variation. We now prove this theorem, for which we need two lemmas. The first one is a slight generalization of one of the more arcane theorems of elementary calculus, the "second mean value theorem for integrals."

8.41 Lemma. *Let ϕ and ψ be real-valued functions on $[a, b]$. Suppose that ϕ is monotone and right continuous on $[a, b]$ and ψ is continuous on $[a, b]$. Then there exists $\eta \in [a, b]$ such that*

$$\int_a^b \phi(x)\psi(x)\,dx = \phi(a) \int_a^\eta \psi(x)\,dx + \phi(b) \int_\eta^b \psi(x)\,dx.$$

Proof. Adding a constant c to ϕ changes both sides of the equation by the amount $c \int_a^b \psi(x)\,dx$, so we may assume that $\phi(a) = 0$. We may also assume that ϕ

is increasing; otherwise replace ϕ by $-\phi$. Let $\Psi(x) = \int_x^b \psi(t)\,dt$ (so that $\Psi' = -\psi$) and apply Theorem 3.36:

$$\int_a^b \phi(x)\psi(x)\,dx = -\phi(x)\Psi(x)\big|_a^b + \int_{(a,b]} \Psi(x)\,d\phi(x).$$

The endpoint evaluations vanish since $\phi(a) = \Psi(b) = 0$. Since ϕ is increasing and $\int_{(a,b]} d\phi = \phi(b) - \phi(a) = \phi(b)$, if m and M are the minimum and maximum values of Ψ on $[a,b]$ we have $m\phi(b) \le \int_{(a,b]} \Psi\,d\phi \le M\phi(b)$. By the intermediate value theorem, then, there exists $\eta \in [a,b]$ such that $\int_{(a,b]} \Psi\,d\phi = \Psi(\eta)\phi(b)$, which is the desired result. ∎

8.42 Lemma. *There is a constant $C < \infty$ such that for every $m \ge 0$ and every $[a,b] \subset [-\frac{1}{2}, \frac{1}{2}]$,*

$$\left|\int_a^b D_m(x)\,dx\right| \le C.$$

Moreover, $\int_{-1/2}^0 D_m(x)\,dx = \int_0^{1/2} D_m(x)\,dx = \frac{1}{2}$ for all m.

Proof. By (8.40),

$$\int_a^b D_m(x)\,dx = \int_a^b \frac{\sin(2m+1)\pi x}{\pi x}\,dx + \int_a^b \sin(2m+1)\pi x\left[\frac{1}{\sin \pi x} - \frac{1}{\pi x}\right]dx.$$

Since $(\sin \pi x)^{-1} - (\pi x)^{-1}$ is bounded on $[-\frac{1}{2}, \frac{1}{2}]$ and $|\sin(2m+1)\pi x| \le 1$, the second integral on the right is bounded in absolute value by a constant. With the substitution $y = (2m+1)\pi x$, the first one becomes

$$\int_{(2m+1)\pi a}^{(2m+1)\pi b} \frac{\sin y}{\pi y}\,dy = \frac{\mathrm{Si}[(2m+1)\pi b] - \mathrm{Si}[(2m+1)\pi a]}{\pi}$$

where $\mathrm{Si}(x) = \int_0^x y^{-1}\sin y\,dy$. But $\mathrm{Si}(x)$ is continuous and approaches the finite limits $\pm\frac{1}{2}\pi$ as $x \to \pm\infty$ (see Exercise 59b in §2.6), so $\mathrm{Si}(x)$ is bounded. This proves the first assertion. As for the second one,

$$\int_{-1/2}^{1/2} D_m(x)\,dx = \sum_{-m}^{m} \int_{-1/2}^{1/2} e^{2\pi i k x}\,dx = 1$$

(only the term with $k = 0$ is nonzero), so since D_m is even,

$$\int_{-1/2}^{0} D_m(x)\,dx = \int_0^{\frac{1}{2}} D_m(x)\,dx = \tfrac{1}{2}.$$

∎

8.43 Theorem. *If $f \in BV(\mathbb{T})$ — that is, if f is periodic on $\mathbb{R}$ and of bounded variation on $[-\frac{1}{2}, \frac{1}{2}]$ — then*

$$\lim_{m\to\infty} S_m f(x) = \tfrac{1}{2}\big[f(x+) + f(x-)\big] \textit{ for every } x.$$

In particular, $\lim_{m\to\infty} S_m f(x) = f(x)$ *at every x at which f is continuous.*

Proof. We begin by making some reductions. In examining the convergence of $S_m f(x)$, we may assume that $x = 0$ (by replacing f with the translated function $\tau_{-x} f$), that f is real-valued (by considering the real and imaginary parts separately), and that f is right continuous (since replacing $f(t)$ by $f(t+)$ affects neither $S_m f$ nor $\frac{1}{2}[f(0+) + f(0-)]$). In this case, by Theorem 3.27b, on the interval $[-\frac{1}{2}, \frac{1}{2})$ we can write f as the difference of two right continuous increasing functions g and h. If these functions are extended to $\mathbb{R}$ by periodicity, they are again of bounded variation, and it is enough to show that $S_m g(0) \to \frac{1}{2}[g(0+) + g(0-)]$ and likewise for h.

In short, it suffices to consider the case where $x = 0$ and f is increasing and right continuous on $[-\frac{1}{2}, \frac{1}{2})$. Since D_m is even, we have $S_m f(0) = f * D_m(0) = \int_{-1/2}^{1/2} f(x) D_m(x)\,dx$, so by Lemma 8.42,

$$\begin{aligned} &S_m f(0) - \tfrac{1}{2}\big[f(0+) + f(0-)\big] \\ &\qquad = \int_0^{1/2} \big[f(x) - f(0+)\big] D_m(x)\,dx + \int_{-1/2}^{0} \big[f(x) - f(0-)\big] D_m(x)\,dx. \end{aligned}$$

We shall show that the first integral on the right tends to zero as $m \to \infty$; a similar argument shows that the second integral also tends to zero, thereby completing the proof.

Given $\epsilon > 0$, choose $\delta > 0$ small enough so that $f(\delta) - f(0+) < \epsilon/C$ where C is as in Lemma 8.42. Then by Lemma 8.41, for some $\eta \in [0, \delta]$,

$$\left|\int_0^{\delta} \big[f(x) - f(0+)\big] D_m(x)\,dx\right| = \big[f(\delta) - f(0+)\big]\left|\int_\eta^{\delta} D_m(x)\,dx\right|,$$

which is less than ϵ. On the other hand, by (8.40),

$$\int_\delta^{1/2} \big[f(x) - f(0+)\big] D_m(x)\,dx = \widehat{g}_+(-m) - \widehat{g}_-(m),$$

where $g_\pm$ is the periodic function given on the interval $[-\frac{1}{2}, \frac{1}{2})$ by

$$g_\pm(x) = \frac{\big[f(x) - f(0+)\big] e^{\pm\pi i x}}{2i \sin \pi x} \chi_{[\delta,1/2)}(x).$$

But $g_\pm \in L^1(\mathbb{T})$, so $\widehat{g}_\pm(\mp m) \to 0$ as $m \to \infty$ by the Riemann-Lebesgue lemma (the periodic analogue of Theorem 8.22f). Therefore,

$$\limsup_{m\to\infty} \left|\int_0^{1/2} \big[f(x) - f(0+)\big] D_m(x)\,dx\right| < \epsilon$$

for every $\epsilon > 0$, and we are done. ∎

One of the less attractive features of Fourier series is that bad behavior of a function at one point affects the behavior of its Fourier series at all points. For example, if f has even one jump discontinuity, then $\widehat{f}$ cannot be in $l^1(\mathbb{Z})$ and so the series $\sum \widehat{f}(k)e^{2\pi ikx}$ cannot converge absolutely at any point. However, to a limited extent the convergence of the series at a point x depends only on the behavior of f near x, as explained in the following localization theorem.

8.44 Theorem. *If f and g are in $L^1(\mathbb{T})$ and $f = g$ on an open interval I, then $S_m f - S_m g \to 0$ uniformly on compact subsets of I.*

Proof. It is enough to assume that $g = 0$ (consider $f - g$), and by translating f we may assume that I is centered at 0, say $I = (-c, c)$ where $c \le \frac{1}{2}$. Fix $\delta < c$; we shall show that if $f = 0$ on I then $S_m f \to 0$ uniformly on $[-\delta, \delta]$.

The first step is to show that $S_m f \to 0$ pointwise on $[-\delta, \delta]$, and the argument is similar to the preceding proof. Namely, by (8.40) we have

$$S_m f(x) = \int_{-1/2}^{1/2} f(x-y) D_m(y)\, dy = \widehat{g}_{x,+}(-m) - \widehat{g}_{x,-}(m),$$

where

$$g_{x,\pm}(y) = \frac{f(x-y)e^{\pm \pi i y}}{2i \sin \pi y}.$$

Since $f(x-y) = 0$ on a neighborhood of the zeros of $\sin \pi y$, the functions $g_{x,\pm}$ are in $L^1(\mathbb{T})$, so $\widehat{g}_{x,\pm}(\mp m) \to 0$ by the Riemann-Lebesgue lemma.

The next step is to show that if $x_1, x_2 \in [-\delta, \delta]$, then $S_m f(x_1) - S_m f(x_2)$ vanishes as $x_1 - x_2 \to 0$, uniformly in m. By (8.40) again,

$$S_m f(x_1) - S_m f(x_2) = \int_{-1/2}^{1/2} \frac{\sin(2m+1)\pi y}{\sin \pi y} \left[f(x_1 - y) - f(x_2 - y) \right] dy.$$

But $f(x_1 - y) - f(x_2 - y) = 0$ for $|y| < c - \delta$, and for $c - \delta \le |y| \le \frac{1}{2}$ we have

$$\left| \frac{\sin(2m+1)\pi y}{\sin \pi y} \right| \le \frac{1}{\sin \pi (c - \delta)} = A,$$

where A is independent of m. Hence

$$|S_m f(x_1) - S_m f(x_2)| \le A \int_{-1/2}^{1/2} |f(x_1 - y) - f(x_2 - y)|\, dy = A \|\tau_{x_1} f - \tau_{x_2} f\|_1,$$

which vanishes as $x_1 - x_2 \to 0$ by (the periodic analogue of) Proposition 8.5.

Now, given $\epsilon > 0$, we can choose η small enough so that if $x_1, x_2 \in [-\delta, \delta]$ and $|x_1 - x_2| < \eta$, then $|S_m f(x_1) - S_m f(x_2)| < \epsilon/2$. Choose $x_1, \ldots, x_k \in [-\delta, \delta]$ so that the intervals $|x - x_j| < \eta$ cover $[-\delta, \delta]$. Since $S_m f(x_j) \to 0$ for each j, we can choose M large enough so that $|S_m f(x_j)| < \epsilon/2$ for $m > M$ and $1 \le j \le k$. If $|x| \le \delta$, then, we have $|x - x_j| < \eta$ for some j, so

$$|S_m f(x)| \le |S_m f(x) - S_m f(x_j)| + |S_m f(x_j)| < \epsilon$$

for $m > M$, and we are done. ∎

8.45 Corollary. *Suppose that $f \in L^1(\mathbb{T})$ and I is an open interval of length ≤ 1.*

a. *If f agrees on I with a function g such that $\widehat{g} \in l^1(\mathbb{Z})$, then $S_m f \to f$ uniformly on compact subsets of I.*

b. *If f is absolutely continuous on I and $f' \in L^p(I)$ for some $p > 1$, then $S_m f \to f$ uniformly on compact subsets of I.*

Proof. If $f = g$ on I, then $S_m f - f = S_m f - g = (S_m f - S_m g) + (S_m g - g)$ on I, and if $\widehat{g} \in l^1(\mathbb{Z})$, then $S_m g \to g$ uniformly on $\mathbb{R}$; (a) follows. As for (b), given $[a_0, b_0] \subset I$, pick $a < a_0$ and $b > b_0$ so that $[a, b] \subset I$, and let g be the continuous periodic function that equals f on $[a, b]$ and is linear on $[b, a+1]$ (which is unique since $g(b) = f(b)$ and $g(a+1) = g(a) = f(a)$). Under the hypotheses of (b), g is absolutely continuous on $\mathbb{R}$ and $g' \in L^p(\mathbb{T})$, so $\widehat{g} \in l^1(\mathbb{Z})$ by Theorem 8.33. Thus $S_m f \to f$ uniformly on $[a_0, b_0]$ by (a). ∎

Finally, we discuss the behavior of $S_m f$ near a jump discontinuity of f. Let us first consider a simple example: Let

$$\phi(x) = \tfrac{1}{2} - x - [x] \qquad ([x] = \text{ greatest integer } \leq x). \tag{8.46}$$

Then ϕ is periodic and is C^∞ except for jump discontinuities at the integers, where $\phi(j+) - \phi(j-) = 1$. It is easy to check that $\widehat{\phi}(0) = 0$ and $\widehat{\phi}(k) = (2\pi i k)^{-1}$ for $k \neq 0$ (Exercise 13a), so that

$$S_m \phi(x) = \sum_{0<|k|\leq m} \frac{e^{2\pi i k x}}{2\pi i k} = \sum_1^m \frac{\sin 2\pi k x}{\pi k}.$$

From Corollary 8.45 it follows that $S_m \phi \to \phi$ uniformly on any compact set not containing an integer, and it is obvious that $S_m \phi(x) = 0$ when x is an integer. But near the integers a peculiar thing happens: $S_m \phi$ contains a sequence of spikes that overshoot and undershoot ϕ, as shown in Figure 8.1, and as $m \to \infty$ the spikes tend to zero in width but *not* in height. In fact, when m is large the value of $S_m \phi$ at its first maximum to the right of 0 is about 0.5895, about 18% greater than $\phi(0+) = \frac{1}{2}$. This is known as the **Gibbs phenomenon**; the precise statement and proof are given in Exercise 37.

Now suppose that f is any periodic function on $\mathbb{R}$ having a jump discontinuity at $x = a$ (that is, $f(a+)$ and $f(a-)$ exist and are unequal). Then the function

$$g(x) = f(x) - \big[f(a+) - f(a-)\big]\phi(x - a)$$

is continuous at every point where f is, and also at $x = a$ provided that we (re)define $g(a)$ to be $\frac{1}{2}[f(a+) + f(a-)]$, as the jumps in f and ϕ cancel out. If g satisfies one of the hypotheses of Corollary 8.45 on an interval I containing a, the Fourier series of g will converge uniformly near a, and hence the Fourier series of f will exhibit the same Gibbs phenomenon as that of ϕ.

Finally, suppose that f is periodic and continuous except at finitely many points $a_1, \ldots, a_k \in \mathbb{T}$, where f has jump discontinuities. We can then subtract off all the

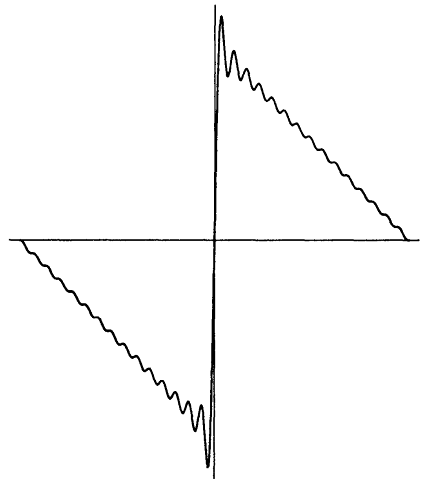

Fig. 8.1 The Gibbs phenomenon: the graph $y = \sum_1^{30} (\pi k)^{-1} \sin 2\pi kx$, $-\frac{1}{2} \leq x \leq \frac{1}{2}$.

jumps to form a continuous function g:

$$g(x) = f(x) - \sum [f(a_j+) - f(a_j-)]\phi(x - a_j)$$

If f satisfies some mild smoothness conditions — for example, if f is absolutely continuous on any interval not containing any a_j and $f' \in L^p$ for some $p > 1$ — then $\widehat{g}$ will be in $l^1(\mathbb{Z})$. Conclusion: $S_m f \to f$ uniformly on any interval not containing any a_j, $S_m(a_j) \to \frac{1}{2}[f(a_j+) + f(a_j-)]$, and $S_m f$ exhibits the Gibbs phenomenon near every a_j.

Exercises

33. Let $\sigma_m f$ be the Cesàro means of the Fourier series of f given by (8.39).
a. $\sigma_m f = f * F_m$ where $F_m = (m+1)^{-1} \sum_0^m D_k$ and D_k is the kth Dirichlet kernel. (See Exercise 29a.) F_m is called the mth **Fejér kernel.**
b. $F_m(x) = \sin^2(m+1)\pi x/(m+1)\sin^2 \pi x$. (Use (8.40) and the fact that $\sin(2k+1)\pi x = \operatorname{Im} e^{(2k+1)\pi i x}$.)

34. If D_m is the mth Dirichlet kernel, $\|D_m\|_1 \to \infty$ as $m \to \infty$. (Make the substitution $y = (2m+1)\pi x$ and use Exercise 59a in §2.6.)

35. The purpose of this exercise is to show that the Fourier series of "most" continuous functions on $\mathbb{T}$ do not converge pointwise.
a. Define $\phi_m(f) = S_m f(0)$. Then $\phi \in C(\mathbb{T})^*$ and $\|\phi\| = \|D_m\|_1$.
b. The set of all $f \in C(\mathbb{T})$ such that the sequence $\{S_m f(0)\}$ converges is meager in $C(\mathbb{T})$. (Use Exercise 34 and the uniform boundedness principle.)
c. There exist $f \in C(\mathbb{T})$ (in fact, a residual set of such f's) such that $\{S_m f(x)\}$ diverges for every x in a dense subset of $\mathbb{T}$. (The result of (b) holds if the point 0 is replaced by any other point in $\mathbb{T}$. Apply Exercise 40 in §5.3.)

36. The Fourier transform is not surjective from $L^1(\mathbb{T})$ to $C_0(\mathbb{Z})$. (Use Exercise 34, and cf. Exercise 16c.)

37. Let ϕ be given by (8.46) and let $\Delta_m = S_m\phi - \phi$.

a. $(d/dx)\Delta_m(x) = D_m(x)$ for $x \notin \mathbb{Z}$.

b. The first maximum of Δ_m to the right of 0 occurs at $x = (2m+1)^{-1}$, and

$$\lim_{m\to\infty} \Delta_m\left(\frac{1}{2m+1}\right) = \frac{1}{\pi}\int_0^\pi \frac{\sin t}{t}\,dt - \frac{1}{2} \cong 0.0895.$$

(Use (8.40) and the fact that $\Delta_m(x) = \int_0^x \Delta_m'(t)\,dt - \frac{1}{2}$.)

c. More generally, the jth critical point of Δ_m to the right of 0 occurs at $x = j/(2m+1)$ ($j = 1, \ldots, 2m$), and

$$\lim_{m\to\infty} \Delta_m\left(\frac{j}{2m+1}\right) = \frac{1}{\pi}\int_0^{j\pi} \frac{\sin t}{t}\,dt - \frac{1}{2}.$$

These numbers are positive for j odd and negative for j even. (See Exercise 59b in §2.6.)

8.6 FOURIER ANALYSIS OF MEASURES

We recall that $M(\mathbb{R}^n)$ is the space of complex Borel measures on $\mathbb{R}^n$ (which are automatically Radon measures by Theorem 7.8), and we embed $L^1(\mathbb{R}^n)$ into $M(\mathbb{R}^n)$ by identifying $f \in L^1$ with the measure $d\mu = f\,dm$. We shall need to define products of complex measures on Cartesian product spaces, which can easily be done in terms of products of positive measures by using Radon-Nikodym derivatives. Namely, if $\mu, \nu \in M(\mathbb{R}^n)$, we define $\mu \times \nu \in M(\mathbb{R}^n \times \mathbb{R}^n)$ by

$$d(\mu \times \nu)(x,y) = \frac{d\mu}{d|\mu|}(x)\frac{d\nu}{d|\nu|}(y)\,d(|\mu| \times |\nu|)(x,y).$$

If $\mu, \nu \in M(\mathbb{R}^n)$, we define their convolution $\mu * \nu \in M(\mathbb{R}^n)$ by $\mu * \nu(E) = \mu \times \nu(\alpha^{-1}(E))$ where $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is addition, $\alpha(x,y) = x + y$. In other words,

$$\mu \times \nu(E) = \iint \chi_E(x+y)\,d\mu(x)\,d\nu(y). \tag{8.47}$$

8.48 Proposition.

a. Convolution of measures is commutative and associative.

b. For any bounded Borel measurable function h,

$$\int h\,d(\mu * \nu) = \iint h(x+y)\,d\mu(x)\,d\nu(y).$$

c. $\|\mu * \nu\| \leq \|\mu\| \, \|\nu\|$.

d. *If* $d\mu = f\,dm$ *and* $d\nu = g\,dm$, *then* $d(\mu * \nu) = (f * g)\,dm$*; that is, on* L^1 *the new and old definitions of convolution coincide.*

Proof. Commutativity is obvious from Fubini's theorem, as is associativity, for $\lambda * \mu * \nu$ is unambiguously defined by the formula

$$\lambda * \mu * \nu(E) = \iiint \chi_E(x + y + z)\, d\lambda(x)\, d\mu(y)\, d\nu(z).$$

Assertion (b) follows from (8.47) by the usual linearity and approximation arguments. In particular, taking $h = d|\mu * \nu|/d(\mu * \nu)$, since $|h| = 1$ we obtain

$$\|\mu * \nu\| = \int h\, d(\mu * \nu) \leq \iint |h|\, d|\mu|\, d|\nu| = \|\mu\| \, \|\nu\|,$$

which proves (c). Finally, if $d\mu = f\,dm$ and $d\nu = g\,dm$, for any bounded measurable h we have

$$\begin{aligned}\int h\, d(\mu * \nu) &= \iint h(x + y) f(x) g(y)\, dx\, dy \\ &= \iint h(x) f(x - y) g(y)\, dx\, dy = \int h(x)(f * g)(x)\, dx,\end{aligned}$$

whence $d(\mu * \nu) = (f * g)\,dm$. ∎

We can also define convolutions of measures with functions in $L^p(\mathbb{R}^n, m)$, which we implicitly assume to be Borel measurable. (By Proposition 2.12, this is no restriction.)

8.49 Proposition. *If* $f \in L^p(\mathbb{R}^n)$ $(1 \leq p \leq \infty)$ *and* $\mu \in M(\mathbb{R}^n)$, *then the integral* $f * \mu(x) = \int f(x - y)\, d\mu(y)$ *exists for a.e.* x, $f * \mu \in L^p$, *and* $\|f * \mu\|_p \leq \|f\|_p \|\mu\|$. *(Here "*$L^p$*" and "a.e." refer to Lebesgue measure.)*

Proof. If f and μ are nonnegative, then $f * \mu(x)$ exists (possibly being equal to ∞) for every x, and by Minkowski's inequality for integrals,

$$\|f * \mu\|_p \leq \int \|f(\cdot - y)\|_p\, d\mu(y) = \|f\|_p \|\mu\|.$$

In particular, $f * \mu(x) < \infty$ for a.e. x. In the general case this argument applies to $|f|$ and $|\mu|$, and the result follows easily. ∎

In the case $p = 1$, the definition of $f * \mu$ in Proposition 8.49 coincides with the definition given earlier in which f is identified with $f\,dm$, for

$$\int_E f * \mu(x)\, dx = \iint \chi_E(x) f(x - y)\, d\mu(y)\, dx = \iint \chi_E(x + y) f(x)\, dx\, d\mu(y)$$

for any Borel set E. Thus $L^1(\mathbb{R}^n)$ is not merely a subalgebra of $M(\mathbb{R}^n)$ with respect to convolution but an ideal.

We extend the Fourier transform from $L^1(\mathbb{R}^n)$ to $M(\mathbb{R}^n)$ in the obvious way: If $\mu \in M(\mathbb{R}^n)$, $\widehat{\mu}$ is the function defined by

$$\widehat{\mu}(\xi) = \int e^{-2\pi i \xi \cdot x}\, d\mu(x).$$

(The Fourier transform on measures is sometimes called the **Fourier-Stieltjes transform.**) Since $e^{-2\pi i \xi \cdot x}$ is uniformly continuous in x, it is clear that $\widehat{\mu}$ is a bounded continuous function and that $\|\widehat{\mu}\|_u \leq \|\mu\|$. Moreover, by taking $h(x) = e^{-2\pi i \xi \cdot x}$ in Proposition 8.48b, one sees immediately that $(\mu * \nu)\widehat{\ } = \widehat{\mu}\widehat{\nu}$.

We conclude by giving a useful criterion for vague convergence of measures in terms of Fourier transforms.

8.50 Proposition. *Suppose that $\mu_1, \mu_2, \ldots$, and μ are in $M(\mathbb{R}^n)$. If $\|\mu_k\| \leq C < \infty$ for all k and $\widehat{\mu}_k \to \widehat{\mu}$ pointwise, then $\mu_k \to \mu$ vaguely.*

Proof. If $f \in \mathcal{S}$,then $f^\vee \in \mathcal{S}$ (Corollary 8.23), so by the Fourier inversion theorem,

$$\int f\, d\mu_k = \iint f^\vee(y) e^{-2\pi i y \cdot x}\, dy\, d\mu_k(x) = \int f^\vee(y)\widehat{\mu}_k(y)\, dy.$$

Since $f^\vee \in L^1$ and $\|\widehat{\mu}_k\|_u \leq C$, the dominated convergence theorem implies that $\int f\, d\mu_k \to \int f\, d\mu$. But $\mathcal{S}$ is dense in $C_0(\mathbb{R}^n)$ (Proposition 8.17), so by Proposition 5.17, $\int f\, d\mu_k \to \int f\, d\mu$ for all $f \in C_0(\mathbb{R}^n)$, that is, $\mu_k \to \mu$ vaguely. ∎

This result has a partial converse: If $\mu_k \to \mu$ vaguely and $\|\mu_k\| \to \|\mu\|$, then $\widehat{\mu}_k \to \widehat{\mu}$ pointwise. This follows from Exercise 26 in §7.3.

Exercises

38. Work out the analogues of the results in this section for measures on the torus $\mathbb{T}^n$.

39. If μ is a positive Borel measure on $\mathbb{T}$ with $\mu(\mathbb{T}) = 1$, then $|\widehat{\mu}(k)| < 1$ for all $k \neq 0$ unless μ is a linear combination, with positive coefficients, of the point masses at $0, \frac{1}{m}, \ldots, \frac{m-1}{m}$ for some $m \in \mathbb{N}$, in which case $\widehat{\mu}(jm) = 1$ for all $j \in \mathbb{Z}$.

40. $L^1(\mathbb{R}^n)$ is vaguely dense in $M(\mathbb{R}^n)$. (If $\mu \in M(\mathbb{R}^n)$, consider $\phi_t * \mu$ where $\{\phi_t\}_{t>0}$ is an approximate identity.)

41. Let Δ be the set of finite linear combinations of the point masses δ_x, $x \in \mathbb{R}^n$. Then Δ is vaguely dense in $M(\mathbb{R}^n)$. (If f is in the dense subset $C_c(\mathbb{R}^n)$ of $L^1(\mathbb{R}^n)$ and $g \in C_0(\mathbb{R}^n)$, approximate $\int fg$ by Riemann sums. Then use Exercise 40.)

42. A function ϕ on $\mathbb{R}^n$ that satisfies $\sum_{j,k=1}^m z_j \bar{z}_k \phi(x_j - x_k) \geq 0$ for all $z_1, \ldots, z_m \in \mathbb{C}$ and all $x_1, \ldots, x_m \in \mathbb{R}^n$, for any $m \in \mathbb{N}$, is called **positive definite**. If $\mu \in M(\mathbb{R}^n)$ is positive, then $\widehat{\mu}$ is positive definite.

8.7 APPLICATIONS TO PARTIAL DIFFERENTIAL EQUATIONS

In this section we present a few of the many applications of Fourier analysis to the theory of partial differential equations; others will be found in Chapter 9. We shall use the term **differential operator** to mean a linear partial differential operator with smooth coefficients, that is, an operator L of the form

$$Lf(x) = \sum_{|\alpha|\leq m} a_\alpha(x)\partial^\alpha f(x), \qquad a_\alpha \in C^\infty.$$

If the a_α's are constants, we call L a **constant-coefficient** operator. In this case, if for all sufficiently well-behaved functions f (for example, $f \in \mathcal{S}$) we have

$$(Lf)\widehat{\ }(\xi) = \sum_{|\alpha|\leq m} a_\alpha(2\pi i\xi)^\alpha \widehat{f}(\xi).$$

It is therefore convenient to write L in a slightly different form: We set $b_\alpha = (2\pi i)^{|\alpha|}a_\alpha$ and introduce the operators

$$D^\alpha = (2\pi i)^{-|\alpha|}\partial^\alpha,$$

so that

$$L = \sum_{|\alpha|\leq m} b_\alpha D^\alpha, \qquad (Lf)\widehat{\ } = \sum_{|\alpha|\leq m} b_\alpha \xi^\alpha \widehat{f}.$$

Thus, if P is any polynomial in n complex variables, say $P(\xi) = \sum_{|\alpha|\leq m} b_\alpha\xi^\alpha$, we can form the constant-coefficient operator $P(D) = \sum_{|\alpha|\leq m} b_\alpha D^\alpha$, and we then have $[P(D)f]\widehat{\ } = P\widehat{f}$. The polynomial P is called the **symbol** of the operator $P(D)$.

Clearly, one potential application of the Fourier transform is in finding solutions of the differential equation $P(D)u = f$. Indeed, application of the Fourier transform to both sides yields $\widehat{u} = P^{-1}\widehat{f}$, whence $u = (P^{-1}\widehat{f})^\vee$. Moreover, if P^{-1} is the Fourier transform of a function ϕ, we can express u directly in terms of f as $u = f * \phi$. For these calculations to make sense, however, the functions f and $P^{-1}\widehat{f}$ (or P^{-1}) must be ones to which the Fourier transform can be applied, which is a serious limitation within the theory we have developed so far. The full power of this method becomes available only when the the domain of the Fourier transform is substantially extended. We shall do this in §9.2; for the time being, we invite the reader to work out a fairly simple example in Exercise 43. (It must also be pointed out that even when this method works, $u = (P^{-1}\widehat{f})^\vee$ is far from being the only solution of $P(D)u = f$; there are others that grow too fast at infinity to be within the scope even of the extended Fourier transform.)

Let us turn to some more concrete problems. The most important of all partial differential operators is the **Laplacian**

$$\Delta = \sum_1^n \frac{\partial^2}{\partial x_j^2} = -4\pi^2\sum_1^n D_j^2 = P(D) \text{ where } P(\xi) = -4\pi^2|\xi|^2.$$

The reason for this is that Δ is essentially the only (scalar) differential operator that is invariant under translations and rotations. (If one considers operators on vector-valued functions, there are others, such as the familiar grad, curl, and div of 3-dimensional vector analysis.) More precisely, we have:

8.51 Theorem. *A differential operator L satisfies $L(f \circ T) = (Lf) \circ T$ for all translations and rotations T iff there is a polynomial P in one variable such that $L = P(\Delta)$.*

Proof. Clearly L is translation-invariant iff L has constant coefficients, in which case $L = Q(D)$ for some polynomial Q in n variables. Moreover, since $(Lf)\widehat{\ } = Q\widehat{f}$ and the Fourier transform commutes with rotations, L commutes with rotations iff Q is rotation-invariant. Let $Q = \sum_0^m Q_j$ where Q_j is homogeneous of degree j; then it is easy to see that Q is rotation-invariant iff each Q_j is rotation-invariant. (Use induction on j and the fact that $Q_j(\xi) = \lim_{r\to 0} r^{-j} \sum_j^m Q_i(r\xi)$.) But this means that $Q_j(\xi)$ depends only on $|\xi|$, so $Q_j(\xi) = c_j|\xi|^j$ by homogeneity. Moreover, $|\xi|^j$ is a polynomial precisely when j is even, so $c_j = 0$ for j odd. Setting $b_k = (-4\pi^2)^{-k}c_{2k}$, then, we have $Q(\xi) = \sum b_k(-4\pi^2|\xi|^2)^k$, that is, $L = \sum b_k\Delta^k$. ∎

One of the basic boundary value problems for the Laplacian is the **Dirichlet problem**: Given an open set $\Omega \subset \mathbb{R}^n$ and a function f on its boundary $\partial\Omega$, find a function u on $\overline{\Omega}$ such that $\Delta u = 0$ on Ω and $u|\partial\Omega = f$. (This statement of the problem is deliberately a bit imprecise.) We shall solve the Dirichlet problem when Ω is a half-space.

For this purpose it will be convenient to replace n by $n+1$ and to denote the coordinates on $\mathbb{R}^{n+1}$ by $x_1, \ldots, x_n, t$. We continue to use the symbol Δ to denote the Laplacian on $\mathbb{R}^n$, and we set

$$\partial_t = \frac{\partial}{\partial t},$$

so the Laplacian on $\mathbb{R}^{n+1}$ is $\Delta + \partial_t^2$. We take the half-space Ω to be $\mathbb{R}^n \times (0, \infty)$. Thus, given a function f on $\mathbb{R}^n$, satisfying conditions to be made more precise below, we wish to find a function u on $\mathbb{R}^n \times [0, \infty)$ such that $(\Delta + \partial_t^2)u = 0$ and $u(x, 0) = f(x)$.

The idea is to apply the Fourier transform on $\mathbb{R}^n$, thus converting the partial differential equation $(\Delta + \partial_t^2)u = 0$ into the simple ordinary differential equation $(-4\pi^2|\xi|^2 + \partial_t^2)\widehat{u} = 0$. The general solution of this equation is

$$\widehat{u}(\xi, t) = c_1(\xi)e^{-2\pi t|\xi|} + c_2(\xi)e^{2\pi t|\xi|}, \tag{8.52}$$

and we require that $\widehat{u}(\xi, 0) = \widehat{f}(\xi)$. We therefore obtain a solution to our problem by taking $c_1(\xi) = \widehat{f}(\xi)$, $c_2(\xi) = 0$ (more about the reasons for this choice below); this gives $\widehat{u}(\xi, t) = \widehat{f}(\xi)e^{-2\pi t|\xi|}$, or $u(x, t) = (f * P_t)(x)$ where $P_t = (e^{-2\pi t|\xi|})^\vee$ is the Poisson kernel introduced in §8.4. As we calculated in Exercise 26,

$$P_t(x) = \frac{\Gamma(\frac{1}{2}(n+1))}{\pi^{(n+1)/2}} \frac{t}{(t^2 + |x|^2)^{-(n+1)/2}}.$$

So far this is all formal, since we have not specified conditions on f to ensure that these manipulations are justified. We now give a precise result.

8.53 Theorem. *Suppose $f \in L^p(\mathbb{R}^n)$ $(1 \leq p \leq \infty)$. Then the function $u(x,t) = (f * P_t)(x)$ satisfies $(\Delta + \partial_t^2)u = 0$ on $\mathbb{R}^n \times (0,\infty)$, and $\lim_{t\to 0} u(x,t) = f(x)$ for a.e. x and for every x at which f is continuous. Moreover, $\lim_{t\to 0} \|u(\cdot,t) - f\|_p = 0$ provided $p < \infty$.*

Proof. P_t and all of its derivatives are in $L^q(\mathbb{R}^n)$ for $1 \leq q \leq \infty$, since a rough calculation shows that $|\partial_x^\alpha P_t(x)| \leq C_\alpha |x|^{-n-1-|\alpha|}$ and $|\partial_t^j P_t(x)| \leq C_j |x|^{-n-1}$ for large x. Also, $(\Delta + \partial_t^2)P_t(x) = 0$, as can be verified by direct calculation or (more easily) by taking the Fourier transform. Hence $f * P_t$ is well defined and

$$(\Delta + \partial_t^2)(f * P_t) = f * (\Delta + \partial_t^2)P_t = 0.$$

Since $P_t(x) = t^{-n}P_1(t^{-1}x)$ and $\int P_1(x)\,dx = \widehat{P}_1(0) = 1$, the remaining assertions follow from Theorems 8.14 and 8.15. ∎

The function $u(x,t) = (f * P_t)(x)$ is not the only one satisfying the conclusions of Theorem 8.53; for example, $v(x,t) = u(x,t) + ct$ also works, for any $c \in \mathbb{C}$. For $f \in L^1$, we could also obtain a large family of solutions by taking c_2 in (8.52) to be an arbitrary function in C_c^∞ and $c_1 = \widehat{f} - c_2$. (But there is no nice convolution formula for the resulting function u, because $e^{2\pi t|\xi|}$ is not the Fourier transform of a function or even a distribution.) The solution $u(x,t) = (f * P_t)(x)$ is distinguished, however, by its regularity at infinity; for example, it can be shown that if $f \in BC(\mathbb{R}^n)$, then u is the unique solution in $BC(\mathbb{R}^n \times [0,\infty))$.

The same idea can be used to solve the **heat equation**

$$(\partial_t - \Delta)u = 0$$

on $\mathbb{R}^n \times (0,\infty)$ subject to the initial condition $u(x,0) = f(x)$. (Physical interpretation: $u(x,t)$ represents the temperature at position x and time t in a homogeneous isotropic medium, given that the temperature at time 0 is $f(x)$.) Indeed, Fourier transformation leads to the ordinary differential equation $(\partial_t + 4\pi^2|\xi|^2)\widehat{u} = 0$ with initial condition $\widehat{u}(\xi,0) = \widehat{f}(\xi)$. The unique solution of the latter problem is $\widehat{u}(\xi,t) = \widehat{f}(\xi)e^{-4\pi^2 t|\xi|^2}$. In view of Proposition 8.24, this yields

$$u(x,t) = f * G_t(x), \qquad G_t(x) = (4\pi t)^{-n/2}e^{-|x|^2/4t}.$$

Here we have $G_t(x) = t^{-n/2}G_1(t^{-1/2}x)$, so after the change of variable $s = \sqrt{t}$, Theorems 8.14 and 8.15 apply again, and we obtain an exact analogue of Theorem 8.53 for the initial value problem $(\partial_t - \Delta)u = 0$, $u(x,0) = f(x)$. Actually, in the present case the hypotheses on f can be relaxed considerably because $G_t \in \mathcal{S}$; see Exercise 44.

Another fundamental equation of mathematical physics is the **wave equation**

$$(\partial_t^2 - \Delta)u = 0.$$

(Physical interpretation: $u(x,t)$ is the amplitude at position x and time t of a wave traveling in a homogeneous isotropic medium, with units chosen so that the speed of propagation is 1.) Here it is appropriate to specify both $u(x,0)$ and $\partial_t u(x,0)$:

$$(8.54) \qquad (\partial_t^2 - \Delta)u = 0, \qquad u(x,0) = f(x), \qquad \partial_t u(x,0) = g(x).$$

After applying the Fourier transform, we obtain

$$(\partial_t^2 + 4\pi^2|\xi|^2)\widehat{u}(\xi,t) = 0, \qquad \widehat{u}(\xi,0) = \widehat{f}(\xi), \qquad \partial_t\widehat{u}(\xi,0) = \widehat{g}(\xi),$$

the solution to which is

$$(8.55) \qquad \widehat{u}(\xi,t) = (\cos 2\pi t|\xi|)\widehat{f}(\xi) + \frac{\sin 2\pi t|\xi|}{2\pi|\xi|}\widehat{g}(\xi).$$

Since

$$\cos 2\pi t|\xi| = \frac{\partial}{\partial t}\left[\frac{\sin 2\pi t|\xi|}{2\pi|\xi|}\right],$$

it follows that

$$u(x,t) = f * \partial_t W_t(x) + g * W_t(x), \text{ where } W_t = \left[\frac{\sin 2\pi t|\xi|}{2\pi|\xi|}\right]^\vee.$$

But here there is a problem: $(2\pi|\xi|)^{-1}\sin 2\pi t|\xi|$ is the Fourier transform of a function only when $n \leq 2$ and the Fourier transform of a measure only when $n \leq 3$; for these cases the resulting solution of the wave equation is worked out in Exercises 45–47. To carry out this analysis in higher dimensions requires the theory of distributions, which we shall examine in Chapter 9. (We shall not, however, derive the explicit formula for W_t, which becomes increasingly complicated as n increases.)

Exercises

43. Let $\phi(x) = e^{-|x|/2}$ on $\mathbb{R}$. Use the Fourier transform to derive the solution $u = f * \phi$ of the differential equation $u - u'' = f$, and then check directly that it works. What hypotheses are needed on f?

44. Let $G_t(x) = (4\pi t)^{-n/2}e^{-|x|^2/4t}$, and suppose that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ satisfies $|f(x)| \leq C_\epsilon e^{\epsilon|x|^2}$ for every $\epsilon > 0$. Then $u(x,t) = f * G_t(x)$ is well defined for all $x \in \mathbb{R}^n$ and $t > 0$; $(\partial_t - \Delta)u = 0$ on $\mathbb{R}^n \times (0,\infty)$; and $\lim_{t\to 0} u(x,t) = f(x)$ for a.e. x and for every x at which f is continuous. (To show $u(x,t) \to f(x)$ a.e. on a bounded open set V, write $f = \phi f + (1-\phi)f$ where $\phi \in C_c$ and $\phi = 1$ on V, and show that $[(1-\phi)f] * G_t \to 0$ on V.)

45. Let $n = 1$. Use (8.55) and Exercise 15a to derive d'Alembert's solution to the initial value problem (8.54):

$$u(x,t) = \frac{1}{2}[f(x+t) + f(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} g(s)\,ds.$$

Under what conditions on f and g does this formula actually give a solution?

46. Let $n = 3$, and let σ_t denote surface measure on the sphere $|x| = t$. Then

$$\frac{\sin 2\pi t|\xi|}{2\pi|\xi|} = (4\pi t)^{-1}\widehat{\sigma}_t(\xi).$$

(See Exercise 22d.) What is the resulting solution of the initial value problem (8.54), expressed in terms of convolutions? What conditions on f and g ensure its validity?

47. Let $n = 2$. If $\xi \in \mathbb{R}^2$, let $\widetilde{\xi} = (\xi, 0) \in \mathbb{R}^3$. Rewrite the result of Exercise 46,

$$\frac{\sin 2\pi t|\widetilde{\xi}|}{2\pi|\widetilde{\xi}|} = \frac{1}{4\pi t}\int_{|x|=t} e^{-2\pi i\widetilde{\xi}\cdot x}\, d\sigma_t(x),$$

in terms of an integral over the disc $D_t = \{y : |y| \le t\}$ in $\mathbb{R}^2$ by projecting the upper and lower hemispheres of the sphere $|x| = t$ in $\mathbb{R}^3$ onto the equatorial plane. Conclude that $(2\pi|\xi|)^{-1}\sin 2\pi t|\xi|$ is the Fourier transform of

$$W_t(x) = (2\pi)^{-1}(t^2 - |x|^2)^{-1/2}\chi_{D_t}(x),$$

and write out the resulting solution of the initial value problem (8.54).

48. Solve the following initial value problems in terms of Fourier series, where f, g, and $u(\cdot, t)$ are periodic functions on $\mathbb{R}$:

a. $(\partial_t^2 + \partial_x^2)u = 0$, $u(x, 0) = f(x)$. (Cf. the discussion of Abel means in §8.4.)

b. $(\partial_t - \partial_x^2)u = 0$, $u(x, 0) = f(x)$.

c. $(\partial_t^2 - \partial_x^2)u = 0$, $u(x, 0) = f(x)$, $\partial_t u(x, 0) = g(x)$.

49. In this exercise we discuss heat flow on an interval.

a. Solve $(\partial_t - \partial_x^2)u = 0$ on $(a, b) \times (0, \infty)$ with boundary conditions $u(x, 0) = f(x)$ for $x \in (a, b)$, $u(a, t) = u(b, t) = 0$ for $t > 0$, in terms of Fourier series. (This describes heat flow on (a, b) when the endpoints are held at a constant temperature. It suffices to assume $a = 0$, $b = \frac{1}{2}$; extend f to $\mathbb{R}$ by requiring f to be odd and periodic, and use Exercise 48b.)

b. Solve the same problem with the condition $u(a, t) = u(b, t) = 0$ replaced by $\partial_x u(a, t) = \partial_x u(b, t) = 0$. (This describes heat flow on (a, b) when the endpoints are insulated. This time, extend f to be even and periodic.)

50. Solve $(\partial_t^2 - \partial_x^2)u = 0$ on $(a, b) \times (0, \infty)$ with boundary conditions $u(x, 0) = f(x)$ and $\partial_t u(x, 0) = g(x)$ for $x \in (a, b)$, $u(a, t) = u(b, t) = 0$ for $t > 0$, in terms of Fourier series by the method of Exercise 49a. (This problem describes the motion of a vibrating string that is fixed at the endpoints. It can also be solved by extending f to be odd and periodic and using Exercise 45. That form of the solution tells you what you see when you look at a vibrating string; this one tells you what you hear when you listen to it.)

8.8 NOTES AND REFERENCES

The scope of Fourier analysis is much wider than we have been able to indicate in this chapter. Dym and McKean [36] gives a more comprehensive treatment with many interesting applications. Also recommended are Körner's delightful book [87], which discusses various aspects of classical Fourier analysis and their role in science, and the excellent collection of expository articles edited by Ash [7], which gives a broader view of the mathematical ramifications of the subject. On the more advanced level, the reader should consult Zygmund [167] for the classical theory and Stein [140], [141] and Stein and Weiss [142] for some of the more recent developments.

§8.1: The formulas given in most calculus books for the remainder term $R_k(x) = f(x) - P_k(x)$ in Taylor's formula (where P_k is the Taylor polynomial of degree k) require f to possess derivatives of order $k+1$, but this is not really necessary. The version of Taylor's theorem stated in the text is derived in Folland [45].

§8.3: Trigonometric series and integrals have a very long history, but modern Fourier analysis only became possible after the invention of the Lebesgue integral. When that tool became available, the L^2 theory was quickly established: the Riesz-Fischer theorem [44], [114] for Fourier series (essentially Theorem 8.20), and the Plancherel theorem [110] for Fourier integrals. Since then the subject has developed in many directions.

There is no universal agreement on where to put the factors of 2π in the definition of the Fourier transform. Other common conventions are

$$\mathcal{F}_1 f(\xi) = \int e^{-i\xi\cdot x} f(x)\,dx, \qquad \mathcal{F}_2 f(\xi) = (2\pi)^{-n/2}\int e^{-i\xi\cdot x}\,dx,$$

whose inverse transforms are

$$\mathcal{F}_1^{-1} g(x) = (2\pi)^{-n}\int e^{i\xi\cdot x} g(\xi)\,d\xi, \qquad \mathcal{F}_2^{-1} g(\xi) = (2\pi)^{-n/2}\int e^{i\xi\cdot x} g(\xi)\,d\xi.$$

$\mathcal{F}_1$ has the disadvantage of not being unitary ($\|\mathcal{F}_1 f\|_2 = (2\pi)^{n/2}\|f\|_2$), whereas $\mathcal{F}_2$ is unitary but does not convert convolution into multiplication ($\mathcal{F}_2(f * g) = (2\pi)^{n/2}(\mathcal{F}_2 f)(\mathcal{F}_2 g)$). To make both L^2 norms and convolutions come out right, one can either put the 2π's in the exponent, as we have done, or omit them from the exponent but replace Lebesgue measure dx by $(2\pi)^{-n/2}\,dx$ in defining both the Fourier transform (as in $\mathcal{F}_2$) and convolutions.

The Hausdorff-Young inequality $\|\widehat{f}\|_q \le \|f\|_p$ ($1 \le p \le 2$, $p^{-1} + q^{-1} = 1$) is sharp on $\mathbb{T}^n$, since equality holds when f is a constant function; but on $\mathbb{R}^n$ the optimal result, a deep theorem of Beckner [14], is that $\|\widehat{f}\|_q \le p^{n/2p} q^{-n/2q}\|f\|_p$.

One of the fundamental qualitative features of the Fourier transform is the fact that, roughly speaking, a nonzero function and its Fourier transform cannot both be sharply localized, that is, they cannot both be negligibly small outside of small sets. This general principle has a number of different precise formulations, two of which are derived in Exercises 18 and 19; see Folland and Sitaram [50] for a comprehensive discussion.

A nice complex-variable proof of the fact that the Fourier transform is injective on L^1 can be found in Newman [106].

§§8.4–5: The theory of convergence of one-dimensional Fourier series really began (as mentioned in the text) with Dirichlet's theorem in 1829. The first construction of a continuous function whose Fourier series does not converge pointwise was obtained by du Bois Reymond in 1876, and the fact that the Fourier series of a continuous function f is uniformly Cesàro summable to f was proved by Fejér in 1904. In 1926 Kolmogorov produced an $f \in L^1(\mathbb{T})$ such that $\{S_m f(x)\}$ diverges at *every* x; on the other hand, in 1927 M. Riesz proved that for $1 < p < \infty$, $\|S_m f - f\|_p \to 0$ for every $f \in L^p(\mathbb{T})$. The culmination of this subject is the theorem of Carleson (1966, for $p = 2$) and Hunt (1967, for general p) that if $f \in L^p(\mathbb{T})$ where $p > 1$, then $S_m f \to f$ almost everywhere.

For more information, see Zygmund [167], the articles by Zygmund and Hunt in Ash [7], and Fefferman [42]. Also see Hewitt and Hewitt [74] for an interesting historical discussion of the Gibbs phenomenon.

Convergence of Fourier series in n variables is an even trickier subject. In the first place, one must decide what one means by a partial sum of a series indexed by $\mathbb{Z}^n$. It is a straightforward consequence of the Riesz and Carleson-Hunt theorems that if $f \in L^p(\mathbb{T}^n)$ with $p > 1$, the "cubical partial sums"

$$S_m^c f(x) = \sum_{\|\kappa\| \le m} \widehat{f}(\kappa) e^{2\pi i \kappa \cdot x} \qquad (\|\kappa\| = \max(|\kappa_1|, \ldots, |\kappa_n|))$$

converge to f a.e. and (if $p < \infty$) in the L^p norm. On the other hand, C. Fefferman proved the rather shocking result that for the "spherical partial sums"

$$S_r f(x) = \sum_{|\kappa| < r} \widehat{f}(\kappa) e^{2\pi i \kappa \cdot x} \qquad \Big(|\kappa|^2 = \sum_1^n \kappa_j^2\Big),$$

the convergence $\lim_{r \to \infty} \|S_r f - f\|_p = 0$ holds for all $f \in L^p$ *only* when $p = 2$, if $n > 1$. Of course, one can consider modifications of the spherical partial sums in the hope of obtaining positive results; the most intensively studied of these are the **Bochner-Riesz means**

$$\sigma_r^\alpha f(x) = \sum_{|\kappa| < r} \big(1 - |r^{-1}\kappa|^2\big)^\alpha \widehat{f}(\kappa) e^{2\pi i \kappa \cdot x}$$

obtained by taking $\Phi(\xi) = \big[\max(1 - |\xi|^2), 0\big]^\alpha$ in Theorem 8.36. (When $n = 1$, $\sigma_{m+1}^1 f$ is essentially equivalent to the Cesàro mean $\sigma_m f$.) These Φ's satisfy the hypotheses of Theorem 8.36 when $\alpha > \frac{1}{2}(n-1)$, and some positive results are also known for smaller values of α.

Davis and Chang [30] is a good source for all of this material; see also Stein and Weiss [142] and Ash [7].

§8.7: The solution of the initial value problem for the wave equation in arbitrary dimensions can be found in Folland [48]; see also Folland [49]. Further applications of Fourier analysis to differential equations can be found in Folland [46], [48], Körner [87], and Taylor [147].