

Chapter VI

Fourier Series, Fourier Transform, and Characteristic Functions

Fourier series and Fourier transform provide one of the most important tools for analysis and partial differential equations, with widespread applications to physics in particular and science in general. This is (up to a scalar multiple) a norm-preserving (i.e., isometry), linear transformation on the Hilbert space of square-integrable complex-valued functions. It turns the integral operation of convolution of functions into the elementary algebraic operation of the product of the transformed functions, and that of differentiation of a function into multiplication by its Fourier frequency.

Although beyond our scope, this powerful and elegant theory extends beyond functions on finite-dimensional Euclidean space to infinite-dimensional spaces and locally compact abelian groups.¹ From this point of view, Fourier series is the Fourier transform on the circle group.

This chapter develops the basic properties of Fourier series and the Fourier transform with applications to the central limit theorem and to transience and recurrence of random walks.

Consider a real- or complex-valued periodic function on the real line. By changing the scale if necessary, one may take the period to be 2π . Is it possible to represent f as a superposition of the periodic functions (“waves”) $\cos nx$, $\sin nx$ of frequency n ($n = 0, 1, 2, \dots$)? In view of **Weierstrass approximation theorem**, every continuous periodic function f of period 2π is the limit (in the sense of uniform convergence of functions) of a sequence of **trigonometric polynomials**, i.e., functions of the form

$$\sum_{n=-T}^T c_n e^{inx} = c_0 + \sum_{n=1}^T (a_n \cos nx + b_n \sin nx);$$

the Bernstein polynomials in e^{ix} illustrate one such approximation.

¹Extensions of the theory can be found in the following standard references, among others: Rudin (1967), Grenander (1963), Parthasarathy (1967).

As will be seen, Theorem 6.1 below gives an especially useful version of the approximation from the perspective of *Fourier series*. In a Fourier series, the coefficients of the polynomials are especially defined according to an $L^2[-\pi, \pi]$ -orthogonality of the complex exponentials $e^{inx} = \cos(nx) + i \sin(nx)$ as explained below. For this special choice of coefficients the theory of Fourier series yields, among other things, that with the weaker notion of L^2 -convergence the approximation holds for a wider class of functions, namely for all square-integrable functions f on $[-\pi, \pi]$; here square integrability means that f is measurable and that $\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$; denoted $f \in L^2[-\pi, \pi]$. It should be noted that in general, we consider integrals of complex-valued functions in this section, and the $L^p = L^p(dx)$ spaces are those of complex-valued functions (see Exercise 36 of Chapter I).

The successive coefficients c_n for this approximation are the so-called **Fourier coefficients**:

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \quad (n = 0, \pm 1, \pm 2, \dots). \quad (6.1)$$

The main point of Theorem 6.1 in this context is to provide a tool for *uniformly approximating* continuous functions by trigonometric polynomials whose coefficients more closely approximate Fourier coefficients than alternatives such as Bernstein polynomials.

As remarked above, the functions e^{inx} ($n = 0, \pm 1, \pm 2, \dots$) form an **orthonormal set**:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx} dx = \begin{cases} 0, & \text{for } n \neq m, \\ 1 & \text{for } n = m, \end{cases} \quad (6.2)$$

so that the **Fourier series of** f is written formally, without regard to convergence for the time being, as

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}. \quad (6.3)$$

As such, this is a representation of f as a superposition of orthogonal components. To make matters precise we first prove the following useful class of Fejér polynomials; see Exercise 1 for an alternative approach.

Theorem 6.1 Let f be a continuous periodic function of period 2π . Then, given $\delta > 0$, there exists a trigonometric polynomial, specifically a **Fejér average** $\sum_{n=-N}^N d_n e^{inx}$, where

$$d_n = \left(1 - \frac{|n|}{N+1}\right) \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad n = 0, \pm 1, \pm 2, \dots,$$

such that

$$\sup_{x \in \mathbb{R}^1} \left| f(x) - \sum_{n=-N}^N d_n e^{inx} \right| < \delta.$$

Proof For each positive integer N , introduce the **Fejér kernel**

$$k_N(x) := \frac{1}{2\pi} \sum_{n=-N}^N \left(1 - \frac{|n|}{N+1}\right) e^{inx}. \quad (6.4)$$

This may also be expressed as

$$\begin{aligned} 2\pi(N+1)k_N(x) &= \sum_{0 \leq j, k \leq N} e^{i(j-k)x} = \left| \sum_{j=0}^N e^{ijx} \right|^2 \\ &= \frac{2\{1 - (\cos(N+1)x)\}}{2(1 - \cos x)} = \left(\frac{\sin\{\frac{1}{2}(N+1)x\}}{\sin \frac{1}{2}x} \right)^2. \end{aligned} \quad (6.5)$$

At $x = 2n\pi$ ($n = 0, \pm 1, \pm 2, \dots$), the right side is taken to be $(N+1)^2$. The first equality in (6.5) follows from the fact that there are $N+1 - |n|$ pairs (j, k) in the sum such that $j - k = n$. It follows from (6.5) that k_N is a positive continuous periodic function with period 2π . Also, k_N is a pdf on $[-\pi, \pi]$, since nonnegativity follows from (6.5) and normalization from (6.4) on integration. For every $\varepsilon > 0$ it follows from (6.5) that $k_N(x)$ goes to zero uniformly on $[-\pi, -\varepsilon] \cup [\varepsilon, \pi]$, so that

$$\int_{[-\pi, -\varepsilon] \cup [\varepsilon, \pi]} k_N(x) dx \rightarrow 0 \quad \text{as } N \rightarrow \infty. \quad (6.6)$$

In other words, $k_N(x) dx$ converges weakly to $\delta_0(dx)$, the point mass at 0, as $N \rightarrow \infty$.

Consider now the approximation f_N of f defined by

$$f_N(x) := \int_{-\pi}^{\pi} f(y) k_N(x-y) dy = \sum_{n=-N}^N \left(1 - \frac{|n|}{N+1}\right) c_n e^{inx}, \quad (6.7)$$

where c_n is the n th Fourier coefficient of f . By changing variables and using the periodicity of f and k_N , one may express f_N as

$$f_N(x) = \int_{-\pi}^{\pi} f(x-y) k_N(y) dy.$$

Therefore, writing $M = \sup\{|f(x)| : x \in \mathbb{R}^k\}$, and $\delta_\varepsilon = \sup\{|f(y) - f(y')| : |y - y'| < \varepsilon\}$, one has

$$|f(x) - f_N(x)| \leq \int_{-\pi}^{\pi} |f(x-y) - f(x)| k_N(y) dy \leq 2M \int_{[-\pi, -\varepsilon] \cup [\varepsilon, \pi]} k_N(y) dy + \delta_\varepsilon. \quad (6.8)$$

It now follows from (6.6) that $f - f_N$ converges to zero uniformly as $N \rightarrow \infty$. Now write $d_n = (1 - |n|/(N + 1))c_n$. ■

Remark 6.1 The representation of the approximating trigonometric polynomial for f as a convolution $f * k_N$, where k_N is a nonnegative kernel such that $k_N \Rightarrow \delta_0$ is a noteworthy consequence of the proof of Theorem 6.1. The advantages of such polynomials over an approximation by Bernstein polynomials will become evident in the context of unique determination of an integrable periodic function, or even a finite measure on the circle, from its Fourier coefficients; see Proposition 6.3 and Theorem 6.4 below.

The first task is to establish the convergence of the Fourier series (6.3) to f in L^2 . Here the norm $\|\cdot\|$ is $\|\cdot\|_2$ as defined by (6.10) below. If $f(x) = \sum_{n=-N}^N a_n e^{inx}$ is a trigonometric polynomial then the proof is immediate. The general case follows by a uniform approximation of 2π -periodic continuous function by such trigonometric polynomials, and finally the density of such continuous functions in $L^2[-\pi, \pi]$.

Theorem 6.2

- For every f in $L^2[-\pi, \pi]$, the Fourier series of f converges to f in L^2 -norm, and the identity $\|f\| = (\sum_{-\infty}^{\infty} |c_n|^2)^{1/2}$ holds for its Fourier coefficients c_n . Here $\|\cdot\|$ is defined in (6.10).
- If (i) f is differentiable, (ii) $f(-\pi) = f(\pi)$, and (iii) f' is square-integrable, then the Fourier series of f also converges uniformly to f on $[-\pi, \pi]$.

Proof (a) Note that for every square-integrable f and all positive integers N ,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(f(x) - \sum_{-N}^N c_n e^{inx} \right) e^{-imx} dx = c_m - c_m = 0 \quad (m = 0, \pm 1, \dots, \pm N). \quad (6.9)$$

Therefore, if one defines the **norm** (or “length”) of a function g in $L^2[-\pi, \pi]$ by

$$\|g\| = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(x)|^2 dx \right)^{1/2} \equiv \|g\|_2, \quad (6.10)$$

then, writing \bar{z} for the complex conjugate of z ,

$$\begin{aligned} 0 &\leq \|f - \sum_{-N}^N c_n e^{inx}\|^2 \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(f(x) - \sum_{-N}^N c_n e^{inx} \right) \left(\bar{f}(x) - \sum_{-N}^N \bar{c}_n e^{-inx} \right) dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x) - \sum_{-N}^N c_n e^{inx}) \bar{f}(x) dx \\
&= \|f\|^2 - \sum_{-N}^N c_n \bar{c}_n = \|f\|^2 - \sum_{-N}^N |c_n|^2.
\end{aligned} \tag{6.11}$$

This shows that $\|f - \sum_{-N}^N c_n e^{inx}\|^2$ decreases as N increases and that

$$\lim_{N \rightarrow \infty} \|f - \sum_{-N}^N c_n e^{inx}\|^2 = \|f\|^2 - \sum_{-\infty}^{\infty} |c_n|^2. \tag{6.12}$$

To prove that the right side of (6.12) vanishes, first assume that f is continuous and $f(-\pi) = f(\pi)$. Given $\varepsilon > 0$, there exists, by Theorem 6.1, a trigonometric polynomial $\sum_{-N_0}^{N_0} d_n e^{inx}$ such that

$$\max_x \left| f(x) - \sum_{-N_0}^{N_0} d_n e^{inx} \right| < \varepsilon.$$

This implies

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \sum_{-N_0}^{N_0} d_n e^{inx} \right|^2 dx < \varepsilon^2. \tag{6.13}$$

But by (6.9), $f(x) - \sum_{-N_0}^{N_0} c_n \exp\{inx\}$ is orthogonal to e^{imx} ($m = 0, \pm 1, \dots, \pm N_0$), so that

$$\begin{aligned}
&\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \sum_{-N_0}^{N_0} d_n e^{inx} \right|^2 dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \sum_{-N_0}^{N_0} c_n e^{inx} + \sum_{-N_0}^{N_0} (c_n - d_n) e^{inx} \right|^2 dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \sum_{-N_0}^{N_0} c_n e^{inx} \right|^2 dx \\
&\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{-N_0}^{N_0} (c_n - d_n) e^{inx} \right|^2 dx.
\end{aligned} \tag{6.14}$$

Hence, by (6.13), (6.14), and (6.11),

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \sum_{-N_0}^{N_0} c_n e^{inx} \right|^2 dx < \varepsilon^2, \quad \lim_{N \rightarrow \infty} \left\| f - \sum_{-N}^N c_n e^{in \cdot} \right\|^2 \leq \varepsilon^2. \quad (6.15)$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$\lim_{N \rightarrow \infty} \left\| f(x) - \sum_{-N}^N c_n e^{inx} \right\| = 0, \quad (6.16)$$

and by (6.12),

$$\|f\|^2 = \sum_{-\infty}^{\infty} |c_n|^2. \quad (6.17)$$

This completes the proof of convergence for continuous periodic f . Now it may be shown that given a square-integrable f and $\varepsilon > 0$, there exists a continuous periodic g such that $\|f - g\| < \varepsilon/2$ (Exercise 1). Also, letting $\sum a_n e^{inx}$, $\sum c_n e^{inx}$ be the Fourier series of g , f , respectively, there exists N_1 such that

$$\left\| g - \sum_{-N_1}^{N_1} a_n e^{in \cdot} \right\| < \frac{\varepsilon}{2}.$$

Hence (see (6.14))

$$\begin{aligned} \left\| f - \sum_{-N_1}^{N_1} c_n e^{in \cdot} \right\| &\leq \left\| f - \sum_{-N_1}^{N_1} a_n e^{in \cdot} \right\| \leq \|f - g\| + \left\| g - \sum_{-N_1}^{N_1} a_n e^{in \cdot} \right\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned} \quad (6.18)$$

Since $\varepsilon > 0$ is arbitrary and $\|f(\cdot) - \sum_{-N}^N c_n e^{in \cdot}\|^2$ decreases to $\|f\|^2 - \sum_{-\infty}^{\infty} |c_n|^2$ as $N \uparrow \infty$ (see (6.12)), one has

$$\lim_{N \rightarrow \infty} \left\| f - \sum_{-N}^N c_n e^{in \cdot} \right\| = 0; \quad \|f\|^2 = \sum_{-\infty}^{\infty} |c_n|^2. \quad (6.19)$$

To prove part (b), let f be as specified. Let $\sum c_n e^{inx}$ be the Fourier series of f , and $\sum c_n^{(1)} e^{inx}$ that of f' . Then

$$\begin{aligned} c_n^{(1)} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f'(x) e^{-inx} dx = \frac{1}{2\pi} f(x) e^{-inx} \Big|_{-\pi}^{\pi} + \frac{in}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \\ &= 0 - inc_n = -inc_n. \end{aligned} \quad (6.20)$$

Since f' is square-integrable,

$$\sum_{-\infty}^{\infty} |nc_n|^2 = \sum_{-\infty}^{\infty} |c_n^{(1)}|^2 < \infty. \quad (6.21)$$

Therefore, by the Cauchy–Schwarz inequality,

$$\sum_{-\infty}^{\infty} |c_n| = |c_0| + \sum_{n \neq 0} \frac{1}{|n|} |nc_n| \leq |c_0| + \left(\sum_{n \neq 0} \frac{1}{n^2} \right)^{1/2} \left(\sum_{n \neq 0} |nc_n|^2 \right)^{1/2} < \infty. \quad (6.22)$$

But this means that $\sum c_n e^{inx}$ is uniformly absolutely convergent, since

$$\max_x \left| \sum_{|n| > N} c_n e^{inx} \right| \leq \sum_{|n| > N} |c_n| \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Since the continuous functions $\sum_{-N}^N c_n e^{inx}$ converge uniformly (as $N \rightarrow \infty$) to $\sum_{-\infty}^{\infty} c_n e^{inx}$, the latter must be a continuous function, say h . Uniform convergence to h also implies convergence in norm to h . Since $\sum_{-\infty}^{\infty} c_n e^{inx}$ also converges in norm to f , $f(x) = h(x)$ for all x . If the two continuous functions f and h are not identically equal, then

$$\int_{-\pi}^{\pi} |f(x) - h(x)|^2 dx > 0.$$

■

Definition 6.1 For a finite measure (or a finite-signed measure) μ on the circle $[-\pi, \pi)$ (identifying $-\pi$ and π), the n th **Fourier coefficient** of μ is defined by

$$c_n = \frac{1}{2\pi} \int_{[-\pi, \pi)} e^{-inx} \mu(dx) \quad (n = 0, \pm 1, \dots). \quad (6.23)$$

If μ has a density f , then (6.23) is the same as the n th Fourier coefficient of f given by (6.1).

Proposition 6.3 A finite measure μ on the circle is determined by its Fourier coefficients.

Proof Approximate the measure $\mu(dx)$ by $g_N(x) dx$, where

$$g_N(x) := \int_{[-\pi, \pi)} k_N(x - y) \mu(dy) = \sum_{-N}^N \left(1 - \frac{|n|}{N+1} \right) c_n e^{inx}, \quad (6.24)$$

with c_n defined by (6.23). For every continuous periodic function h (i.e., for every continuous function on the circle),

$$\int_{[-\pi, \pi)} h(x) g_N(x) dx = \int_{[-\pi, \pi)} \left(\int_{[-\pi, \pi)} h(x) k_N(x-y) dx \right) \mu(dy). \quad (6.25)$$

As $N \rightarrow \infty$, the probability measure $k_N(x-y) dx = k_N(y-x) dx$ on the circle converges weakly to $\delta_y(dx)$. Hence, the inner integral on the right side of (6.25) converges to $h(y)$. Since the inner integral is bounded by $\sup\{|h(y)| : y \in \mathbb{R}\}$, Lebesgue's dominated convergence theorem implies that

$$\lim_{N \rightarrow \infty} \int_{[-\pi, \pi)} h(x) g_N(x) dx = \int_{[-\pi, \pi)} h(y) \mu(dy). \quad (6.26)$$

This means that μ is determined by $\{g_N : N \geq 1\}$. The latter in turn are determined by $\{c_n\}_{n \in \mathbb{Z}}$. \blacksquare

We are now ready to answer an important question: When is a given sequence $\{c_n : n = 0, \pm 1, \dots\}$ the sequence of Fourier coefficients of a finite measure on the circle? A sequence of complex numbers $\{c_n : n = 0, \pm 1, \pm 2, \dots\}$ is said to be **positive-definite** if for any finite sequence of complex numbers $\{z_j : 1 \leq j \leq N\}$, one has

$$\sum_{1 \leq j, k \leq N} c_{j-k} z_j \bar{z}_k \geq 0. \quad (6.27)$$

Theorem 6.4 (Herglotz Theorem) $\{c_n : n = 0, \pm 1, \dots\}$ is the sequence of Fourier coefficients of a probability measure on the circle if and only if it is positive-definite, and $c_0 = \frac{1}{2\pi}$.

Proof Necessity If μ is a probability measure on the circle, and $\{z_j : 1 \leq j \leq N\}$ a given finite sequence of complex numbers, then

$$\begin{aligned} \sum_{1 \leq j, k \leq N} c_{j-k} z_j \bar{z}_k &= \frac{1}{2\pi} \sum_{1 \leq j, k \leq N} z_j \bar{z}_k \int_{[-\pi, \pi)} e^{i(j-k)x} \mu(dx) \\ &= \frac{1}{2\pi} \int_{[-\pi, \pi)} \left(\sum_1^N z_j e^{ijx} \right) \left(\sum_1^N \bar{z}_k e^{-ikx} \right) \mu(dx) \\ &= \frac{1}{2\pi} \int_{[-\pi, \pi)} \left| \sum_1^N z_j e^{ijx} \right|^2 \mu(dx) \geq 0. \end{aligned} \quad (6.28)$$

Also,

$$c_0 = \frac{1}{2\pi} \int_{[-\pi, \pi)} \mu(dx) = \frac{1}{2\pi}.$$

Sufficiency. Take $z_j = e^{i(j-1)x}$, $j = 1, 2, \dots, N+1$, in (6.27) to get

$$g_N(x) := \frac{1}{N+1} \sum_{0 \leq j, k \leq N} c_{j-k} e^{i(j-k)x} \geq 0. \quad (6.29)$$

Again, since there are $N+1 - |n|$ pairs (j, k) such that $j - k = n$ ($-N \leq n \leq N$) it follows that (6.29) becomes

$$0 \leq g_N(x) = \sum_{-N}^N \left(1 - \frac{|n|}{N+1}\right) e^{inx} c_n. \quad (6.30)$$

In particular, using (6.2),

$$\int_{[-\pi, \pi]} g_N(x) dx = 2\pi c_0 = 1. \quad (6.31)$$

Hence g_N is a pdf on $[-\pi, \pi]$. By Proposition 7.6, there exists a subsequence $\{g_{N'}\}$ such that $g_{N'}(x) dx$ converges weakly to a probability measure $\mu(dx)$ on $[-\pi, \pi]$ as $N' \rightarrow \infty$. Also, again using (6.2) yields

$$\int_{[-\pi, \pi]} e^{-inx} g_N(x) dx = 2\pi \left(1 - \frac{|n|}{N+1}\right) c_n \quad (n = 0, \pm 1, \dots, \pm N). \quad (6.32)$$

For each fixed n , restrict to the subsequence $N = N'$ in (6.32) and let $N' \rightarrow \infty$. Then, since for each n , $\cos(nx)$, $\sin(nx)$ are bounded continuous functions,

$$2\pi c_n = \lim_{N' \rightarrow \infty} 2\pi \left(1 - \frac{|n|}{N'+1}\right) c_n = \int_{[-\pi, \pi]} e^{-inx} \mu(dx) \quad (n = 0, \pm 1, \dots). \quad (6.33)$$

In other words, c_n is the n th Fourier coefficient of μ . ■

Corollary 6.5 A sequence $\{c_n : n = 0, \pm 1, \dots\}$ of complex numbers is the sequence of Fourier coefficients of a finite measure on the circle $[-\pi, \pi]$ if and only if $\{c_n : n = 0, \pm 1, \dots\}$ is positive-definite.

Proof Since the measure $\mu = 0$ has Fourier coefficients $c_n = 0$ for all n , and the latter is trivially a positive-definite sequence, it is enough to prove the correspondence between nonzero positive-definite sequences and nonzero finite measures. It follows from Theorem 6.4, by normalization, that this correspondence is 1-1 between positive-definite sequences $\{c_n : n = 0, \pm 1, \dots\}$ with $c_0 = c > 0$ and measures on the circle having total mass 2π . ■

It is instructive to consider the **Fourier transform** \hat{f} of an integrable function f on \mathbb{R} , defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{i\xi y} f(y) dy, \quad \xi \in \mathbb{R}. \quad (6.34)$$

as a limiting version of a Fourier series. In particular, if f is differentiable and vanishes outside a finite interval, and if f' is square-integrable, then one may use the Fourier series of f (scaled to be defined on $(-\pi, \pi]$) to obtain (see Exercise 6) the **Fourier inversion formula**,

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(y) e^{-izy} dy. \quad (6.35)$$

Moreover, any f that vanishes outside a finite interval and is square-integrable is automatically integrable, and for such an f one has the **Plancherel identity** (see Exercise 6)

$$\|\hat{f}\|_2^2 := \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 d\xi = 2\pi \int_{-\infty}^{\infty} |f(y)|^2 dy = 2\pi \|f\|_2^2. \quad (6.36)$$

The extension of this theory relating to Fourier series and Fourier transforms in higher dimensions is straightforward along the following lines. The Fourier series of a square-integrable function f on $[-\pi, \pi) \times [-\pi, \pi) \times \cdots \times [-\pi, \pi) = [-\pi, \pi)^k$ is defined by $\sum_v c_v \exp\{i v \cdot x\}$, where the summation is over all *integral vectors* (or *multi-indices*) $v = (v^{(1)}, v^{(2)}, \dots, v^{(k)})$, each $v^{(i)}$ being an integer. Also, $v \cdot x = \sum_{i=1}^k v^{(i)} x^{(i)}$ is the usual Euclidean inner product on \mathbb{R}^k between two vectors $v = (v^{(1)}, \dots, v^{(k)})$ and $x = (x^{(1)}, x^{(2)}, \dots, x^{(k)})$. The *Fourier coefficients* are given by

$$c_v = \frac{1}{(2\pi)^k} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(x) e^{-iv \cdot x} dx. \quad (6.37)$$

The extensions of Theorems (and Proposition) 6.1–6.4 are fairly obvious. Similarly, the *Fourier transform* of an integrable function (with respect to Lebesgue measure on \mathbb{R}^k) f is defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{i\xi \cdot y} f(y) dy \quad (\xi \in \mathbb{R}^k), \quad (6.38)$$

the Fourier inversion formula becomes

$$f(z) = \frac{1}{(2\pi)^k} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \hat{f}(\xi) e^{-iz \cdot \xi} d\xi, \quad (6.39)$$

which holds when $f(x)$ and $\hat{f}(\xi)$ are integrable. The Plancherel identity (6.36) becomes

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 d\xi = (2\pi)^k \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |f(y)|^2 dy, \quad (6.40)$$

which holds whenever f is integrable and square-integrable, i.e., Theorem 6.7 below.

Definition 6.2 The **Fourier transform**² of an integrable (real- or complex-valued) function f on \mathbb{R}^k is the function \hat{f} on \mathbb{R}^k defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^k} e^{i\xi \cdot y} f(y) dy, \quad \xi \in \mathbb{R}^k. \quad (6.41)$$

As a special case, take $k = 1$, $f = \mathbf{1}_{(c,d]}$. Then,

$$\hat{f}(\xi) = \frac{e^{i\xi d} - e^{i\xi c}}{i\xi}, \quad (6.42)$$

so that $\hat{f}(\xi) \rightarrow 0$ as $|\xi| \rightarrow \infty$. Such “decay” in the Fourier transform is to be generally expected for integrable functions as follows.

Proposition 6.6 (Riemann–Lebesgue Lemma) The Fourier transform $\hat{f}(\xi)$ of an integrable function f on \mathbb{R}^k tends to zero in the limit as $|\xi| \rightarrow \infty$.

Proof The convergence to zero as $\xi \rightarrow \pm\infty$ illustrated by (6.42) is clearly valid for arbitrary *step functions*, i.e., finite linear combinations of indicator functions of finite rectangles. Now let f be an arbitrary integrable function. Given $\varepsilon > 0$ there exists a step function f_ε such that (see Remark following Proposition 2.5)

$$\|f_\varepsilon - f\|_1 := \int_{\mathbb{R}^k} |f_\varepsilon(y) - f(y)| dy < \varepsilon. \quad (6.43)$$

Now it follows from (6.41) that $|\hat{f}_\varepsilon(\xi) - \hat{f}(\xi)| \leq \|f_\varepsilon - f\|_1$ for all ξ . Since $\hat{f}_\varepsilon(\xi) \rightarrow 0$ as $|\xi| \rightarrow \infty$, one has $\limsup_{|\xi| \rightarrow \infty} |\hat{f}(\xi)| \leq \varepsilon$. Since $\varepsilon > 0$ is arbitrary,

$$\hat{f}(\xi) \rightarrow 0 \quad \text{as } |\xi| \rightarrow \infty.$$

■

Let us now check that (6.35), (6.36), in fact, hold under the following more general conditions

Theorem 6.7

- If f and \hat{f} are both integrable, then the Fourier inversion formula (6.35) holds.
- If f is integrable as well as square-integrable, then the Plancherel identity (6.36) holds.

²There are several different ways in which Fourier transforms can be parameterized and/or normalized by extra constant factors and/or a different sign in the exponent. The definition given here follows the standard conventions of probability theory.

Proof (a) Let f, \hat{f} be integrable. Assume for simplicity that f is continuous. Note that this assumption is innocuous since the inversion formula yields a continuous (version of) f (see Exercise 7(i) for the steps of the proof without this a priori continuity assumption for f). Let φ_{ε^2} denote the pdf of the Gaussian distribution with mean zero and variance $\varepsilon^2 > 0$. Then writing Z to denote a standard normal random variable,

$$f * \varphi_{\varepsilon^2}(x) = \int_{\mathbb{R}} f(x-y)\varphi_{\varepsilon^2}(y)dy = \mathbb{E}f(x-\varepsilon Z) \rightarrow f(x), \quad (6.44)$$

as $\varepsilon \rightarrow 0$. On the other hand (see Exercise 3),

$$\begin{aligned} f * \varphi_{\varepsilon^2}(x) &= \int_{\mathbb{R}} f(x-y)\varphi_{\varepsilon^2}(y)dy = \int_{\mathbb{R}} f(x-y) \left\{ \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi y} e^{-\varepsilon^2 \xi^2/2} d\xi \right\} dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\varepsilon^2 \xi^2/2} \left\{ \int_{\mathbb{R}} e^{i\xi(x-y)} f(x-y) dy \right\} e^{-i\xi x} d\xi \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} e^{-\varepsilon^2 \xi^2/2} \hat{f}(\xi) d\xi \rightarrow \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} \hat{f}(\xi) d\xi \end{aligned} \quad (6.45)$$

as $\varepsilon \rightarrow 0$. The inversion formula (6.35) follows from (6.44), (6.45). For part (b) see Exercise 7(ii). \blacksquare

Remark 6.2 Since $L^1(\mathbb{R}, dx) \cap L^2(\mathbb{R}, dx)$ is dense in $L^2(\mathbb{R}, dx)$ in the L^2 -metric, the Plancherel identity (6.36) may be extended to all of $L^2(\mathbb{R}, dx)$, extending in this process the definition of the Fourier transform \hat{f} of $f \in L^2(\mathbb{R}, dx)$. However, we do not make use of this extension in this text.

Suppose $k = 1$ to start. If f is continuously differentiable and f, f' are both integrable, then integration by parts yields (Exercise 2(b))

$$\hat{f}'(\xi) = -i\xi \hat{f}(\xi). \quad (6.46)$$

The boundary terms in deriving (6.46) vanish, if f' is integrable (as well as f) then $f(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. More generally, if f is r -times continuously differentiable and $f^{(j)}, 0 \leq j \leq r$, are all integrable, then one may repeat the relation (6.46) to get by induction (Exercise 2(b))

$$\widehat{f^{(r)}}(\xi) = (-i\xi)^r \hat{f}(\xi). \quad (6.47)$$

In particular, (6.47) implies that if f, f', f'' are integrable then \hat{f} is integrable. Similar formulae are readily obtained for dimensions $k > 1$ using integration by parts. From this and the Riemann–Lebesgue lemma one may therefore observe a clear sense in which the *smoothness* of the function f is related to the *rate of decay* of the Fourier

transform at ∞ . The statements of smoothness in higher dimensions use the multi-index notation for derivatives: For a k -tuple of positive integers $\alpha = (\alpha_1, \dots, \alpha_k)$ $|\alpha| = \sum_{j=1}^k \alpha_j$, $\partial^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_k}}{\partial x_k^{\alpha_k}}$.

Theorem 6.8

- a. Suppose f in $L^1(\mathbb{R}^k)$. For $|\alpha| \leq m$, $\hat{f} \in C^m$, and $\partial^\alpha \hat{f} = (ix)^\alpha \hat{f}$.
- b. If (i) $x^\alpha f \in C^m$, (ii) $\partial^\alpha f \in L^1$ for $\alpha \leq m$, and (iii) $\partial^\alpha f \in C_0$ for $|\alpha| \leq m - 1$, then $\partial^{\hat{\alpha}} f(\xi) = (i\xi)^\alpha \hat{f}(\xi)$.

Proof To establish part (i) requires differentiation under the integral and induction on $|\alpha|$. The differentiation is justified by the dominated convergence theorem. Integration by parts yields part (ii) in the case $|\alpha| = 1$, as indicated above. The result then follows by induction on $|\alpha|$. ■

Definition 6.3 The **Fourier transform** $\hat{\mu}$ of a finite measure μ on \mathbb{R}^k , with Borel σ -field \mathcal{B}^k , is defined by

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^k} e^{i\xi \cdot x} d\mu(x). \quad (6.48)$$

If μ is a finite-signed measure, i.e., $\mu = \mu_1 - \mu_2$ where μ_1, μ_2 are finite measures, then also one defines $\hat{\mu}$ by (6.48) directly, or by setting $\hat{\mu} = \hat{\mu}_1 - \hat{\mu}_2$. In particular, if $\mu(dx) = f(x) dx$, where f is real-valued and integrable, then $\hat{\mu} = \hat{f}$. If μ is a probability measure, then $\hat{\mu}$ is also called the **characteristic function** of μ , or of any random vector $X = (X_1, \dots, X_k)$ on (Ω, \mathcal{F}, P) whose distribution is $\mu = P \circ X^{-1}$. In this case, by the change of variable formula, one has the equivalent definition

$$\hat{\mu}(\xi) = \mathbb{E}e^{i\xi \cdot X}, \quad \xi \in \mathbb{R}^k. \quad (6.49)$$

In the case that $\hat{Q} \in L^1(\mathbb{R}^k)$ the Fourier inversion formula yields a density function for $Q(dx)$, i.e., integrability of \hat{Q} implies absolute continuity of Q with respect to Lebesgue measure.

We next consider the **convolution** of two integrable functions f, g :

$$f * g(x) = \int_{\mathbb{R}^k} f(x - y)g(y) dy \quad (x \in \mathbb{R}^k). \quad (6.50)$$

Since by the Tonelli part of the Fubini–Tonelli theorem,

$$\begin{aligned} \int_{\mathbb{R}^k} |f * g(x)| dx &= \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} |f(x - y)||g(y)| dy dx \\ &= \int_{\mathbb{R}^k} |f(x)| dx \int_{-\infty}^{\infty} |g(y)| dy, \end{aligned} \quad (6.51)$$

$f * g$ is integrable. Its Fourier transform is

$$\begin{aligned}
(f * g)^\wedge(\xi) &= \int_{\mathbb{R}^k} e^{i\xi \cdot x} \left(\int_{\mathbb{R}^k} f(x-y)g(y) dy \right) dx \\
&= \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} e^{i\xi \cdot (x-y)} e^{i\xi \cdot y} f(x-y)g(y) dy dx \\
&= \int_{\mathbb{R}^k} \int_{\mathbb{R}^k} e^{i\xi \cdot z} e^{i\xi \cdot y} f(z)g(y) dy dz = \hat{f}(\xi)\hat{g}(\xi), \quad (6.52)
\end{aligned}$$

a result of importance in both probability and analysis. By iteration, one defines the n -fold convolution $f_1 * \cdots * f_n$ of n integrable functions f_1, \dots, f_n and it follows from (6.52) that $(f_1 * \cdots * f_n)^\wedge = \hat{f}_1 \hat{f}_2 \cdots \hat{f}_n$. Note also that if f, g are real-valued integrable functions and one defines the measures μ, ν by $\mu(dx) = f(x) dx$, $\nu(dx) = g(x) dx$, and $\mu * \nu$ by $(f * g)(x) dx$, then

$$\begin{aligned}
(\mu * \nu)(B) &= \int_B (f * g)(x) dx = \int_{\mathbb{R}^k} \left(\int_B f(x-y) dx \right) g(y) dy \\
&= \int_{\mathbb{R}^k} \mu(B-y)g(y) dy = \int_{\mathbb{R}^k} \mu(B-y)d\nu(y), \quad (6.53)
\end{aligned}$$

for every interval (or, more generally, for every Borel set) B . Here $B-y$ is the translate of B by $-y$, obtained by subtracting from each point in B the number y . Also $(\mu * \nu)^\wedge = (f * g)^\wedge = \hat{f}\hat{g} = \hat{\mu}\hat{\nu}$. In general (i.e., whether or not finite-signed measures μ and/or ν have densities), the last expression in (6.53) defines the convolution $\mu * \nu$ of finite-signed measures μ and ν . The Fourier transform of this finite-signed measure is still given by $(\mu * \nu)^\wedge = \hat{\mu}\hat{\nu}$. Recall that if X_1, X_2 are independent k -dimensional random vectors on some probability space (Ω, \mathcal{A}, P) and have distributions Q_1, Q_2 , respectively, then the distribution of $X_1 + X_2$ is $Q_1 * Q_2$. The characteristic function (i.e., Fourier transform) may also be computed from

$$(Q_1 * Q_2)^\wedge(\xi) = \mathbb{E}e^{i\xi \cdot (X_1 + X_2)} = \mathbb{E}e^{i\xi \cdot X_1} \mathbb{E}e^{i\xi \cdot X_2} = \hat{Q}_1(\xi)\hat{Q}_2(\xi). \quad (6.54)$$

This argument extends to finite-signed measures, and is an alternative way of thinking about (or deriving) the result $(\mu * \nu)^\wedge = \hat{\mu}\hat{\nu}$.

Theorem 6.9 (Uniqueness) Let Q_1, Q_2 be probabilities on the Borel σ -field of \mathbb{R}^k . Then $\hat{Q}_1(\xi) = \hat{Q}_2(\xi)$ for all $\xi \in \mathbb{R}^k$ if and only if $Q_1 = Q_2$.

Proof For each $\xi \in \mathbb{R}^k$, one has by definition of the characteristic function that $e^{-i\xi \cdot x} \hat{Q}_1(\xi) = \int_{\mathbb{R}^k} e^{i\xi \cdot (y-x)} Q_1(dy)$. Thus, integrating with respect to Q_2 , one obtains the duality relation

$$\int_{\mathbb{R}^k} e^{-i\xi \cdot x} \hat{Q}_1(\xi) Q_2(d\xi) = \int_{\mathbb{R}^k} \hat{Q}_2(y-x) Q_1(dy). \quad (6.55)$$

Let $\varphi_{1/\sigma^2}(x) = \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{\sigma^2 x^2}{2}}$, $x \in \mathbb{R}$, denote the Gaussian pdf with variance $1/\sigma^2$ centered at 0, and take $Q_2(dx) \equiv \Phi_{1/\sigma^2}(dx) := \prod_{j=1}^k \varphi_{1/\sigma^2}(x_j) dx_1 \cdots dx_k$ in (6.55).

Then $\hat{Q}_2(\xi) = \hat{\Phi}_{1/\sigma^2}(\xi) = e^{-\sum_{j=1}^k \frac{\xi_j^2}{2\sigma^2}} = (\sqrt{2\pi\sigma^2})^k \prod_{j=1}^k \varphi_{\sigma^2}(\xi_j)$ so that the right-hand side may be expressed as $(\sqrt{2\pi\sigma^2})^k$ times the pdf of $\Phi_{\sigma^2} * Q_1$. In particular, one has

$$\frac{1}{2\pi} \int_{\mathbb{R}^k} e^{-i\xi \cdot x} \hat{Q}_1(\xi) e^{-\sum_{j=1}^k \frac{\sigma^2 \xi_j^2}{2}} d\xi_j = \int_{\mathbb{R}^k} \prod_{j=1}^k \varphi_{\sigma^2}(y_j - x_j) Q_1(dy).$$

The right-hand side may be viewed as the pdf of the distribution of the sum of independent random vectors $X_{\sigma^2} + Y$ with respective distributions Φ_{σ^2} and Q_1 . Also, by the Chebyshev inequality, $X_{\sigma^2} \rightarrow 0$ in probability as $\sigma^2 \rightarrow 0$. Thus the distribution of $X_{\sigma^2} + Y$ converges weakly to Q_1 . Equivalently, the pdf of $X_{\sigma^2} + Y$ is given by the expression on the left side, involving Q_1 only through \hat{Q}_1 . In this way \hat{Q}_1 uniquely determines Q_1 . ■

Remark 6.3 Equation (6.55) may be viewed as a form of **Parseval's relation**.

The following version of **Parseval relation** is easily established by an application of the Fubini–Tonelli theorem and definition of characteristic function.

Proposition 6.10 (Parseval Relation) Let Q_1 and Q_2 be probabilities on \mathbb{R}^k with characteristic functions \hat{Q}_1 and \hat{Q}_2 , respectively. Then

$$\int_{\mathbb{R}^k} \hat{Q}_1(\xi) Q_2(d\xi) = \int_{\mathbb{R}^k} \hat{Q}_2(\xi) Q_1(d\xi).$$

At this point we have established that the map $Q \in \mathcal{P}(\mathbb{R}^k) \rightarrow \hat{Q} \in \hat{\mathcal{P}}(\mathbb{R}^k)$ is one to one, and transforms convolution as pointwise multiplication. Some additional basic properties of this map are presented in the exercises. We next consider important special cases of an *inversion formula* for absolutely continuous finite (signed) measures $\mu(dx) = f(x)dx$ on \mathbb{R}^k . This is followed by a result on the *continuity* of the map $Q \rightarrow \hat{Q}$ for respectively the weak topology on $\mathcal{P}(\mathbb{R}^k)$ and the topology of pointwise convergence on $\hat{\mathcal{P}}(\mathbb{R}^k)$. Finally the identification of the *range* of the Fourier transform of finite positive measures is provided. Such results are of notable theoretical and practical value.

Next we will see that the correspondence $Q \mapsto \hat{Q}$, on the set of probability measures with the weak topology onto the set of characteristic functions with the topology of pointwise convergence is *continuous*, thus providing a basic tool for obtaining weak convergence of probabilities on the finite-dimensional space \mathbb{R}^k .

Theorem 6.11 (Cramér–Lévy Continuity Theorem) Let $P_n (n \geq 1)$ be probability measures on $(\mathbb{R}^k, \mathcal{B}^k)$.

- a. If P_n converges weakly to P , then $\hat{P}_n(\xi)$ converges to $\hat{P}(\xi)$ for every $\xi \in \mathbb{R}^k$.

- b. If for some continuous function φ one has $\hat{P}_n(\xi) \rightarrow \varphi(\xi)$ for every ξ , then φ is the characteristic function of a probability P , and P_n converges weakly to P .

Proof (a) Since $\hat{P}_n(\xi)$, $\hat{P}(\xi)$ are the integrals of the bounded continuous function $\exp\{i\xi \cdot x\}$ with respect to P_n and P , it follows from the definition of weak convergence that $\hat{P}_n(\xi) \rightarrow \hat{P}(\xi)$. (b) We will show that $\{P_n : n \geq 1\}$ is tight. First, let $k = 1$. For $\delta > 0$ one has

$$\begin{aligned} \frac{1}{2\delta} \int_{-\delta}^{\delta} (1 - \hat{P}_n(\xi)) d\xi &= \frac{1}{2\delta} \int_{\mathbb{R}} \left\{ \int_{-\delta}^{\delta} (1 - e^{i\xi x}) d\xi \right\} P_n(dx) \\ &= \frac{1}{2\delta} \int_{\mathbb{R}} (2\delta - \xi \left[\frac{\sin(\xi x)}{\xi x} \right]_{-\delta}^{\delta}) P_n(dx) \\ &= \frac{1}{2\delta} \int_{\mathbb{R}} \left(2\delta - 2\delta \frac{\sin(\delta x)}{\delta x} \right) P_n(dx) \\ &= \int_{\mathbb{R}} \left(1 - \frac{\sin(\delta x)}{\delta x} \right) P_n(dx) \\ &\geq \frac{1}{2} P_n(\{x : |\delta x| \geq 2\}) = \frac{1}{2} P_n \left(\left\{ x : |x| \geq \frac{2}{\delta} \right\} \right). \end{aligned}$$

Hence, by assumption,

$$P_n \left(\left\{ x : |x| \geq \frac{2}{\delta} \right\} \right) \leq \frac{2}{2\delta} \int_{-\delta}^{\delta} (1 - \hat{P}_n(\xi)) d\xi \rightarrow \frac{2}{2\delta} \int_{-\delta}^{\delta} (1 - \varphi(\xi)) d\xi,$$

as $n \rightarrow \infty$. Since φ is continuous and $\varphi(0) = 1$, given any $\varepsilon > 0$ one may choose $\delta > 0$ such that $(1 - \varphi(\xi)) \leq \varepsilon/4$ for $|\xi| \leq \delta$. Then the limit in (6.56) is no more than $\varepsilon/2$, proving tightness. For $k > 1$, consider the distribution $P_{j,n}$ under P_n of the one-dimensional projections $x = (x_1, \dots, x_k) \mapsto x_j$ for each $j = 1, \dots, k$. Then $\hat{P}_{j,n}(\xi_j) = \hat{P}_n(0, \dots, 0, \xi_j, 0, \dots, 0) \rightarrow \varphi_j(\xi_j) := \varphi(0, \dots, 0, \xi_j, 0, \dots, 0)$ for all $\xi_j \in \mathbb{R}^1$. The previous argument shows that $\{P_{j,n} : n \geq 1\}$ is a tight family for each $j = 1, \dots, k$. Hence there is a $\delta > 0$ such that $P_n(\{x \in \mathbb{R}^k : |x_j| \leq 2/\delta, j = 1, \dots, k\}) \geq 1 - \sum_{j=1}^k P_{j,n}(\{x_j : |x_j| \geq 2/\delta\}) \geq 1 - k\varepsilon/2$ for all sufficiently large n , establishing the desired tightness. By Prohorov's Theorem (Theorem 7.11), there exists a subsequence of $\{P_n\}_{n=1}^{\infty}$, say $\{P_{n_m}\}_{m=1}^{\infty}$, that converges weakly to some probability P . By part (a), $\hat{P}_{n_m}(\xi) \rightarrow \hat{P}(\xi)$, so that $\hat{P}(\xi) = \varphi(\xi)$ for all $\xi \in \mathbb{R}^k$. Since the limit characteristic function $\varphi(\xi)$ is the same regardless of the subsequence $\{P_{n_m}\}_{m=1}^{\infty}$, it follows that P_n converges weakly to P as $n \rightarrow \infty$. ■

The **law of rare events**, or **Poisson approximation to the binomial distribution**, provides a simple illustration of the Cramér–Lévy continuity Theorem 6.11.

Proposition 6.12 (*Law of Rare Events*) For each $n \geq 1$, suppose that $X_{n,1}, \dots, X_{n,n}$ is a sequence of n i.i.d. 0 or 1-valued random variables with $p_n = P(X_{n,k} = 1)$, $q_n = P(X_{n,k} = 0)$, where $\lim_{n \rightarrow \infty} np_n = \lambda > 0$, $q_n = 1 - p_n$. Then $Y_n = \sum_{k=1}^n X_{n,k}$ converges in distribution to Y , where Y is distributed by the Poisson law

$$P(Y = m) = \frac{\lambda^m}{m!} e^{-\lambda},$$

$m = 0, 1, 2, \dots$

Proof Using the basic fact that $\lim_{n \rightarrow \infty} (1 + \frac{a_n}{n})^n = e^{\lim_n a_n}$ whenever $\{a_n\}_{n=1}^\infty$ is a sequence of complex numbers such that $\lim_n a_n$ exists, one has by independence, and in the limit as $n \rightarrow \infty$,

$$\mathbb{E} e^{i\xi Y_n} = (q_n + p_n e^{i\xi})^n = \left(1 + \frac{np_n(e^{i\xi} - 1)}{n}\right)^n \rightarrow \exp(\lambda(e^{i\xi} - 1)), \quad \xi \in \mathbb{R}.$$

One may simply check that this is the characteristic function of the asserted limiting Poisson distribution. \blacksquare

The development of tools for Fourier analysis of probabilities is concluded with an application of the Herglotz theorem (Theorem 6.4) to identify the **range** of the Fourier transform of finite positive measures.

Definition 6.4 A complex-valued function φ on \mathbb{R}^k is said to be **positive-definite** if for every positive integer n and finite sequences $\{\xi_1, \xi_2, \dots, \xi_n\} \subset \mathbb{R}^k$ and $\{z_1, z_2, \dots, z_n\} \subset \mathbb{C}$ (the set of complex numbers), one has

$$\sum_{1 \leq j, k \leq n} z_j \bar{z}_k \varphi(\xi_j - \xi_k) \geq 0. \quad (6.56)$$

Theorem 6.13 (*Bochner's theorem*) A function φ on \mathbb{R}^k is the Fourier transform of a finite measure on \mathbb{R}^k if and only if it is positive-definite and continuous.

Proof We give the proof in the case $k = 1$ and leave $k > 1$ to the reader. The proof of necessity is entirely analogous to (6.28). It is sufficient to consider the case $\varphi(0) = 1$. For each positive integer N , $c_{j,N} := \varphi(-j2^{-N})$, $j = 0, \pm 1, \pm 2, \dots$, is positive-definite in the sense of (6.27). Hence, by the Herglotz theorem, there exists a probability γ_N on $[-\pi, \pi)$ such that $c_{j,N} = (2\pi)^{-1} \int_{[-\pi, \pi)} e^{-ijx} \gamma_N(dx)$ for each j . By the change of variable $x \rightarrow 2^N x$, one has $\varphi(j2^{-N}) = (2\pi)^{-1} \int_{[-2^N\pi, 2^N\pi)} e^{ij2^{-N}x} \mu_N(dx)$ for some probability $\mu_N(dx)$ on $[-2^N\pi, 2^N\pi)$. The characteristic function $\hat{\mu}_N(\xi) := \int_{\mathbb{R}} e^{i\xi x} \mu_N(dx)$ agrees with φ at all dyadic rational points $j2^{-N}$, $j \in \mathbb{Z}$, dense in \mathbb{R} . To conclude the proof we note that one may use the continuity of $\varphi(\xi)$ to see that the family of functions $\hat{\mu}_N(\xi)$ is equicontinuous by the lemma below. With this it will follow by the Arzelà–Ascoli theorem (Appendix B) that there is a subsequence that converges pointwise to a continuous function g on \mathbb{R} . Since g and φ agree on a dense subset of \mathbb{R} , it follows that $g = \varphi$. \blacksquare

Lemma 1 (*An Equicontinuity Lemma*)

- Let φ_N , $N \geq 1$, be a sequence of characteristic functions of probabilities μ_N . If the sequence is equicontinuous at $\xi = 0$ then it is equicontinuous at all $\xi \in \mathbb{R}$.

- b. In the notation of the above proof of Bochner's theorem, let μ_N be the probability on $[-2^N\pi, 2^N\pi]$ with characteristic function $\varphi_N = \hat{\mu}_N$, where $\varphi_N(\xi) = \varphi(\xi)$ for $\xi = j2^{-N}$, $j \in \mathbb{Z}$. Then, (i) for $h \in [-1, 1]$, $0 \leq 1 - \operatorname{Re} \varphi_N(h2^{-N}) \leq 1 - \operatorname{Re} \varphi(2^{-N})$. (ii) φ_N is equicontinuous at 0, and hence at all points of \mathbb{R} (by (i)).

Proof For the first assertion (a) simply use the Cauchy–Schwarz inequality to check that $|\varphi_N(\xi) - \varphi_N(\xi + \eta)|^2 \leq 2|\varphi_N(0) - \operatorname{Re} \varphi_N(\eta)|$.

For (i) of the second assertion (b), write the formula and note that $1 - \cos(hx) \leq 1 - \cos(x)$ for $-\pi \leq x \leq \pi$, $0 \leq h \leq 1$. For (ii), given $\varepsilon > 0$ find $\delta > 0$, ($0 < \delta < 1$) such that $|1 - \varphi(\theta)| < \varepsilon$ for all $|\theta| < \delta$. Now express each such θ as $\theta = (h_N + k_N)2^{-N}$, where $k_N = [2^N\theta]$ is the integer part of $2^N\theta$, and $h_N = 2^N\theta - [2^N\theta] \in [-1, 1]$. Using the inequality $|a+b|^2 \leq 2|a|^2 + 2|b|^2$ together with the inequality in the proof of (a), one has that $|1 - \varphi_N(\theta)|^2 = |1 - \varphi_N((h_N + k_N)2^{-N})|^2 \leq 2|1 - \varphi(k_N 2^{-N})|^2 + 4|1 - \operatorname{Re} \varphi(2^{-N})| \leq 2\varepsilon^2 + 4\varepsilon$. ■

We will illustrate the use of characteristic functions in two probability applications. For the first, let us recall the general random walk on \mathbb{R}^k from Chapter II. A basic consideration in the probabilistic analysis of the long-run behavior of a stochastic evolution involves frequencies of visits to specific states.

Let us consider the random walk $S_n := Z_1 + \cdots + Z_n$, $n \geq 1$, starting at $S_0 = 0$. The state 0 is said to be **neighborhood recurrent** if for every $\varepsilon > 0$, $P(S_n \in B_\varepsilon \text{ i.o.}) = 1$, where $B_\varepsilon = \{x \in \mathbb{R}^k : |x| < \varepsilon\}$. It will be convenient for the calculations to use the **rectangular norm** $|x| := \max\{|x_j| : j = 1, \dots, k\}$, for $x = (x_1, \dots, x_k)$. All finite-dimensional norms being equivalent, there is no loss of generality in this choice.

Observe that if 0 is not neighborhood recurrent, then for some $\varepsilon > 0$, $P(S_n \in B_\varepsilon \text{ i.o.}) < 1$, and therefore by the Hewitt–Savage 0-1 law, $P(S_n \in B_\varepsilon \text{ i.o.}) = 0$. Much more may be obtained with regard to recurrence dichotomies, expected return times, nonrecurrence, etc., which is postponed to a fuller treatment of stochastic processes. However, the following lemma is required for the result given here. As a warm-up, note that by the Borel–Cantelli lemma I, if $\sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) < \infty$ for some $\varepsilon > 0$ then 0 cannot be neighborhood recurrent. In fact one has the following basic result.

Lemma 2 (Chung–Fuchs) 0 is neighborhood recurrent if and only if for all $\varepsilon > 0$, $\sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) = \infty$.

Proof As noted above, if for some $\varepsilon > 0$, $\sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) < \infty$, then with probability one, S_n will visit B_ε at most finitely often by the Borel–Cantelli lemma I. So it suffices to show that if $\sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) = \infty$ for every $\varepsilon > 0$ then S_n will

visit any given neighborhood of zero infinitely often with probability one. The proof is based on establishing the following two calculations:

$$(A) \quad \sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) = \infty \Rightarrow P(S_n \in B_{2\varepsilon} \text{ i.o.}) = 1,$$

$$(B) \quad \sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) \geq \frac{1}{(2m)^k} \sum_{n=1}^{\infty} P(S_n \in B_{m\varepsilon}), \quad m \geq 2.$$

In particular, $\sum_{n=0}^{\infty} P(S_n \in B_\varepsilon) = \infty$ for some $\varepsilon > 0$, then from (B), $\sum_{n=0}^{\infty} P(S_n \in B_{\varepsilon'}) = \infty$ for all $\varepsilon' < \varepsilon$. In view of (A) this would make 0 neighborhood recurrent. To prove (A), let $N_\varepsilon := \text{card}\{n \geq 0 : S_n \in B_\varepsilon\}$ count the number of visits to B_ε . Also let $T_\varepsilon := \sup\{n : S_n \in B_\varepsilon\}$ denote the (possibly infinite) time of the last visit to B_ε . To prove (A) we will show that if $\sum_{m=0}^{\infty} P(S_m \in B_\varepsilon) = \infty$, then $P(T_{2\varepsilon} = \infty) = 1$. Let r be an arbitrary positive integer. One has

$$\begin{aligned} & P(|S_m| < \varepsilon, |S_n| \geq \varepsilon, \forall n \geq m+r) \\ &= P(m \leq T_\varepsilon < m+r) \\ &= P(T_\varepsilon = m) + P(T_\varepsilon = m+1) + \cdots + P(T_\varepsilon = m+r-1). \end{aligned}$$

Hence,

$$\sum_{m=1}^{\infty} P(|S_m| < \varepsilon, |S_n| \geq \varepsilon, \forall n \geq m+r) = \sum_{m=1}^{\infty} P(T_\varepsilon = m) + \cdots + \sum_{m=1}^{\infty} P(T_\varepsilon = m+r-1) \leq r.$$

Thus,

$$\begin{aligned} r &\geq \sum_{m=0}^{\infty} P(S_m \in B_\varepsilon, |S_n| \geq \varepsilon \forall n \geq m+r) \\ &\geq \sum_{m=0}^{\infty} P(S_m \in B_\varepsilon, |S_n - S_m| \geq 2\varepsilon \forall n \geq m+r) \\ &= \sum_{m=0}^{\infty} P(S_m \in B_\varepsilon) P(|S_n| \geq 2\varepsilon \forall n \geq r). \end{aligned} \tag{6.57}$$

Assuming $\sum_{m=0}^{\infty} P(S_m \in B_\varepsilon) = \infty$, one must therefore have $P(T_{2\varepsilon} \leq r) \leq P(|S_n| \geq 2\varepsilon \forall n \geq r) = 0$. Thus $P(T_{2\varepsilon} < \infty) = 0$. For the proof of (B), let $m \geq 2$ and for $x = (x_1, \dots, x_k) \in \mathbb{R}^k$, define $\tau_x = \inf\{n \geq 0 : S_n \in R_\varepsilon(x)\}$, where $R_\varepsilon(x) := [0, \varepsilon)^k + x := \{y \in \mathbb{R}^k : 0 \leq y_i - x_i < \varepsilon, i = 1, \dots, k\}$ is the translate of $[0, \varepsilon)^k$ by x , i.e., “square with lower left corner at x of side lengths ε .” For arbitrary fixed $x \in \{-m\varepsilon, -(m-1)\varepsilon, \dots, (m-1)\varepsilon\}^k$,

$$\begin{aligned}
\sum_{n=0}^{\infty} P(S_n \in R_\varepsilon(x)) &= \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} P(S_n \in R_\varepsilon(x), \tau_x = m) \\
&\leq \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} P(|S_n - S_m| < \varepsilon, \tau_x = m) \\
&= \sum_{m=0}^{\infty} P(\tau_x = m) \sum_{j=0}^{\infty} P(S_j \in B_\varepsilon) \\
&\leq \sum_{j=0}^{\infty} P(S_j \in B_\varepsilon).
\end{aligned}$$

Thus, it now follow that

$$\begin{aligned}
\sum_{n=0}^{\infty} P(S_n \in B_{m\varepsilon}) &\leq \sum_{n=0}^{\infty} \sum_{x \in \{-m\varepsilon, -(m-1)\varepsilon, \dots, (m-1)\varepsilon\}^k} P(S_n \in R_\varepsilon(x)) \\
&= \sum_{x \in \{-m\varepsilon, -(m-1)\varepsilon, \dots, (m-1)\varepsilon\}^k} \sum_{n=0}^{\infty} P(S_n \in R_\varepsilon(x)) \\
&\leq (2m)^k \sum_{n=0}^{\infty} P(S_n \in B_\varepsilon).
\end{aligned}$$

■

Remark 6.4 On a countable state space such as \mathbb{Z}^d , the topology is discrete and $\{j\}$ is an open neighborhood of j for every state j . Hence neighborhood recurrence is equivalent to *point recurrence*. Using the so-called *strong Markov property* discussed in Chapter XI, one may show that if a state i of a Markov chain on a countable state space is point recurrent, then the probability of reaching a state j , starting from i , is *one*, provided that the n -step transition probability from i to j , $p_{ij}^{(n)}$, is nonzero for some n ; see Example 1 below, and Exercise 5 of Chapter XI.

Example 1 (Polya's Theorem) The simple symmetric random walk $\{S_n : n = 0, 1, 2, \dots\}$ on \mathbb{Z}^k starting at $S_0 = \mathbf{0}$ is defined by the random walk with the discrete displacement distribution $Q(\{\mathbf{e}_j\}) = Q(\{-\mathbf{e}_j\}) = \frac{1}{2k}$, $j = 1, 2, \dots, k$, where \mathbf{e}_j is the j th standard basis vector, i.e., j th column of the $k \times k$ identity matrix. For $k = 1$ the recurrence follows easily from Lemma 2 by the combinatorial identity $P(S_{2n} = 0) = \binom{2n}{n} 2^{-2n}$ and Stirling's formula. For $k = 2$, one may rotate the coordinate axis by $\pi/4$ to map the simple symmetric two-dimensional random walk onto a random walk on the rotated lattice having *independent* one-dimensional simple symmetric random walk coordinates. It then follows for the two-dimensional walk that $P(S_{2n} = 0) = (\binom{2n}{n} 2^{-2n})^2$, from which the point recurrence also follows in two dimensions. Combinatorial arguments for the transience in three or more dimensions

are also possible, but quite a bit more involved. An alternative approach by Fourier analysis is given below.

We turn now to conditions on the distribution of the displacements for neighborhood recurrence in terms of Fourier transforms. If, for example, $\mathbb{E}Z_1$ exists and is nonzero, then it follows from the strong law of large numbers that a.s. $|S_n| \rightarrow \infty$. The following is a complete characterization of neighborhood recurrence in terms of the distribution of the displacements. A simpler warm-up version for random walks on the integer lattice is given in Exercise 25. In the following theorem $\operatorname{Re}(z)$ refers to the *real part* of a complex number z .

Theorem 6.14 (*Chung–Fuchs Recurrence Criterion*) Let Z_1, Z_2, \dots be an i.i.d. sequence of random vectors in \mathbb{R}^k with common distribution Q . Let $\{S_n = Z_1 + \dots + Z_n : n \geq 1\}$, $S_0 = 0$, be a random walk on \mathbb{R}^k starting at 0. Then 0 is a neighborhood-recurrent state if and only if for every $\varepsilon > 0$,

$$\sup_{0 < r < 1} \int_{B_\varepsilon} \operatorname{Re} \left(\frac{1}{1 - r\hat{Q}(\xi)} \right) d\xi = \infty.$$

Proof First observe that the “triangular probability density function” $\hat{f}(\xi) = (1 - |\xi|)^+$, $\xi \in \mathbb{R}$, has the characteristic function $f(x) = 2 \frac{1 - \cos(x)}{x^2}$, $x \in \mathbb{R}$, and therefore, $\frac{1}{2\pi} f(x)$ has characteristic function $\hat{f}(\xi)$ (Exercise 23). One may also check that $f(x) \geq 1/2$ for $|x| \leq 1$ (Exercise 23). Also $\mathbf{f}(\mathbf{x}) := \prod_{j=1}^k f(x_j)$, $\mathbf{x} = (x_1, \dots, x_k)$, has characteristic function $\hat{\mathbf{f}}(\xi) = \prod_{j=1}^k \hat{f}(\xi_j)$, and $\hat{\mathbf{f}}$ has characteristic function $(2\pi)^k \mathbf{f}$. In view of Parseval’s relation (Proposition 6.10), one may write

$$\int_{\mathbb{R}^k} \mathbf{f} \left(\frac{\mathbf{x}}{\lambda} \right) Q^{*n}(d\mathbf{x}) = \lambda^k \int_{\mathbb{R}^k} \hat{\mathbf{f}}(\lambda\xi) \hat{Q}^n(\xi) d\xi,$$

for any $\lambda > 0$, $n \geq 1$. Using the Fubini–Tonelli theorem one therefore has for $0 < r < 1$ that

$$\int_{\mathbb{R}^k} \mathbf{f} \left(\frac{\mathbf{x}}{\lambda} \right) \sum_{n=0}^{\infty} r^n Q^{*n}(d\mathbf{x}) = \lambda^k \int_{\mathbb{R}^k} \frac{\hat{\mathbf{f}}(\lambda\xi)}{1 - r\hat{Q}(\xi)} d\xi.$$

Also, since the integral on the left is real, the right side must also be a real integral. For what follows note that when an indicated integral is real, one may replace the integrand by its respective real part. Suppose that for some $\varepsilon > 0$,

$$\sup_{0 < r < 1} \int_{B_{\frac{1}{\varepsilon}}} \operatorname{Re} \left(\frac{1}{1 - r\hat{Q}(\xi)} \right) d\xi < \infty.$$

Then, it follows that

$$\begin{aligned}
\sum_{n=1}^{\infty} P(S_n \in B_\varepsilon) &= \sum_{n=1}^{\infty} Q^{*n}(B_\varepsilon) \leq 2^k \int_{\mathbb{R}^k} \mathbf{f}\left(\frac{\mathbf{x}}{\varepsilon}\right) \sum_{n=0}^{\infty} Q^{*n}(d\mathbf{x}) \\
&\leq 2^k \varepsilon^k \sup_{0 < r < 1} \int_{\mathbb{R}^k} \frac{\hat{\mathbf{f}}(\varepsilon\xi)}{1 - r\hat{Q}(\xi)} d\xi \\
&\leq 2^k \varepsilon^k \sup_{0 < r < 1} \int_{B_{\frac{1}{\varepsilon}}} \operatorname{Re} \left(\frac{1}{1 - r\hat{Q}(\xi)} \right) d\xi < \infty.
\end{aligned}$$

Thus, in view of Borel–Cantelli I, 0 cannot be neighborhood recurrent.

For the converse, suppose that 0 is not neighborhood recurrent. Then, by Lemma 2, one must have for any $\varepsilon > 0$ that $\sum_{n=1}^{\infty} Q^{*n}(B_\varepsilon) < \infty$.

Let $\varepsilon > 0$. Then, again using the Parseval relation with $(2\pi)^k \hat{\mathbf{f}}$ as the Fourier transform of \mathbf{f} ,

$$\begin{aligned}
\sup_{0 < r < 1} \int_{B_\varepsilon} \operatorname{Re} \left(\frac{1}{1 - r\hat{Q}(\xi)} \right) d\xi &\leq 2^k \sup_{0 < r < 1} \int_{B_\varepsilon} \operatorname{Re} \left(\frac{\mathbf{f}\left(\frac{\mathbf{x}}{\varepsilon}\right)}{1 - r\hat{Q}(\mathbf{x})} \right) d\mathbf{x} \\
&\leq 2^k (2\pi)^k \varepsilon^k \sup_{0 < r < 1} \int_{\mathbb{R}^k} \hat{\mathbf{f}}(\varepsilon\mathbf{x}) \sum_{n=0}^{\infty} r^n Q^{*n}(d\mathbf{x}) \\
&\leq 2^k (2\pi)^k \varepsilon^k \int_{B_{\varepsilon^{-1}}} \hat{\mathbf{f}}(\varepsilon\mathbf{x}) \sum_{n=0}^{\infty} Q^{*n}(d\mathbf{x}) \\
&\leq (4\varepsilon\pi)^k \sum_{n=1}^{\infty} Q^{*n}(B_{\varepsilon^{-1}}) < \infty.
\end{aligned}$$

■

Corollary 6.15 If $\int_{B_\varepsilon} \operatorname{Re} \left(\frac{1}{1 - \hat{Q}(\xi)} \right) d\xi = \infty$ for $\varepsilon > 0$, then the random walk with displacement distribution Q is neighborhood recurrent.³ [Hint: Pass to the limit as $r \rightarrow 1$ in $0 \leq \operatorname{Re} \left(\frac{1}{1 - r\hat{Q}(\xi)} \right)$, using the Chung–Fuchs criterion]

Example 2 (Gaussian Random Walk) Suppose that Q is the k -dimensional standard normal distribution. Then $\hat{Q}(\xi) = e^{-\frac{k|\xi|^2}{2}}$, $\xi \in \mathbb{R}^k$.

We now turn to a hallmark application of Theorem 6.11 in probability to prove the celebrated Theorem 6.16 below. First, we need an estimate on the error in the Taylor polynomial approximation to the exponential function. The following lemma exploits the special structure of the exponential to obtain two bounds: a “good small x bound” and a “good large x bound”, each of which is valid for all x .

³That the converse is also true was independently established in Stone, C. J. (1969): On the potential operator for one-dimensional recurrent random walks, *Trans. AMS*, **136** 427–445, and Ornstein, D. (1969): Random walks, *Trans. AMS*, **138**, 1–60.

Lemma 3 (*Taylor Expansion of Characteristic Functions*) Suppose that X is a random variable defined on a probability space (Ω, \mathcal{F}, P) such that $\mathbb{E}|X|^m < \infty$. Then

$$\left| \mathbb{E}e^{i\xi X} - \sum_{k=0}^m \frac{(i\xi)^k}{k!} \mathbb{E}X^k \right| \leq \mathbb{E} \min \left\{ \frac{|\xi|^{m+1}|X|^{m+1}}{(m+1)!}, 2 \frac{|\xi|^m |X|^m}{m!} \right\}, \quad \xi \in \mathbb{R}.$$

Proof Let $f_m(x) = e^{ix} - \sum_{j=0}^m \frac{(ix)^j}{j!}$. Note that $f_m(x) = i \int_0^x f_{m-1}(y) dy$. Iteration yields a succession of $m-1$ iterated integrals with integrand of modulus $|f_0(y_{m-1})| = |e^{iy_{m-1}} - 1| \leq 2$. The iteration of the integrals is therefore at most $2 \frac{|x|^m}{m!}$. To obtain the other bound note the following integration by parts identity:

$$\int_0^x (x-y)^m e^{iy} dy = \frac{x^{m+1}}{m+1} + \frac{i}{m+1} \int_0^x (x-y)^{m+1} e^{iy} dy.$$

This defines a recursive formula that by induction leads to the expansion

$$e^{ix} = \sum_{j=0}^m \frac{(ix)^j}{j!} + \frac{i^{m+1}}{m!} \int_0^x (x-y)^m e^{iy} dy. \quad (6.58)$$

For $x \geq 0$, bound the modulus of the integrand by $|x-y|^m \leq y^m$ to get the bound on the modulus of the integral term by $\frac{|x|^{m+1}}{(m+1)!}$. Similarly for $x < 0$. Since both bounds hold for all x , the smaller of the two also holds for all x . Now replace x by $|\xi X|$ and take expected values to complete the proof. ■

Theorem 6.16 (*The Classical Central Limit Theorem*) Let $\mathbf{X}_n, n \geq 1$, be i.i.d. k -dimensional random vectors with (common) mean μ and a finite covariance matrix D . Then the distribution of $(\mathbf{X}_1 + \cdots + \mathbf{X}_n - n\mu)/\sqrt{n}$ converges weakly to Φ_D , the normal distribution on \mathbb{R}^k with mean zero and covariance matrix D .

Proof It is enough to prove the result for $\mu = \mathbf{0}$ and $D = I$, the $k \times k$ identity matrix I , since the general result then follows by an affine linear (and hence continuous) transformation. First, consider the case $k = 1, \{X_n : n \geq 1\}$ i.i.d. $\mathbb{E}X_n = 0, \mathbb{E}X_n^2 = 1$. Let φ denote the (common) characteristic function of X_n . Then the characteristic function, say φ_n , of $(X_1 + \cdots + X_n)/\sqrt{n}$ is given at a fixed ξ by

$$\varphi_n(\xi) = \varphi^n(\xi/\sqrt{n}) = \left(1 - \frac{\xi^2}{2n} + o\left(\frac{1}{n}\right) \right)^n, \quad (6.59)$$

where $no(\frac{1}{n}) = o(1) \rightarrow 0$ as $n \rightarrow \infty$. The limit of (6.59) is $e^{-\frac{\xi^2}{2}}$, the characteristic function of the standard normal distribution, which proves the theorem for the case $k = 1$, using Theorem 6.11(b).

For $k > 1$, let $\mathbf{X}_n, n \geq 1$, be i.i.d. with mean zero and covariance matrix I . Then for each fixed $\xi \in \mathbb{R}^k, \xi \neq \mathbf{0}, Y_n = \xi \cdot \mathbf{X}_n, n \geq 1$, defines an i.i.d. sequence of real-valued random variables with mean zero and variance $\sigma_\xi^2 = \xi \cdot \xi$. Hence by the preceding,

$Z_n := (Y_1 + \cdots + Y_n)/\sqrt{n}$ converges in distribution to the one-dimensional normal distribution with mean zero and variance $\xi \cdot \xi$, so that the characteristic function of Z_n converges to the function $\eta \mapsto \exp\{-(\xi \cdot \xi)\eta^2/2\}$, $\eta \in \mathbb{R}$. In particular, at $\eta = 1$, the characteristic function of Z_n is

$$\mathbb{E}e^{i\xi \cdot (X_1 + \cdots + X_n)/\sqrt{n}} \rightarrow e^{-\xi \cdot \xi/2}. \quad (6.60)$$

Since (6.60) holds for every $\xi \in \mathbb{R}^k$, the proof is complete by the Cramér–Lévy continuity theorem. \blacksquare

Let us now establish the Berry–Esseen bound on the rate of convergence first noted in Chapter IV.⁴

Theorem 6.17 (Berry–Esseen Convergence Rate) Let X_1, X_2, \dots be an i.i.d. sequence of random variables having finite third moments $\rho = \mathbb{E}|X_1|^3 < \infty$, with mean μ and variance σ^2 . Then, for $S_n = X_1 + \cdots + X_n$, $n \geq 1$, one has

$$\sup_{x \in \mathbb{R}} |P(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x) - \Phi(x)| \leq \frac{3\mathbb{E}|X_1|^3}{\sigma^3\sqrt{n}}.$$

The proof rests on the following lemma⁵ exploiting the fact that for any $T > 0$, the clearly integrable function $\omega_T(\xi) := 1 - \frac{|\xi|}{T}$, $|\xi| \leq T$, and zero on $|\xi| \geq T$, is by Bochner’s theorem the characteristic function of a probability distribution. In fact, one can exhibit this distribution as $v_T(x) := \frac{1}{\pi} \frac{1 - \cos(Tx)}{T^2 x^2}$, $x \in \mathbb{R}$.

Lemma 4 Let F be a distribution function on \mathbb{R} , and G any function on \mathbb{R} such that $\lim_{x \rightarrow -\infty} G(x) = 0$, $\lim_{x \rightarrow \infty} G(x) = 1$, and having bounded derivative $|G'(x)| \leq m < \infty$. Then, for $T > 0$,

$$\sup_{y \in \mathbb{R}} \left| \int_{\mathbb{R}} (F(y-x) - G(y-x)) \frac{1}{\pi} \frac{1 - \cos(Tx)}{Tx^2} dx \right| \geq \frac{1}{2} \sup_{x \in \mathbb{R}} |F(x) - G(x)| - \frac{12m}{\pi T}.$$

Proof Let $\Delta(x) = F(x) - G(x)$, $x \in \mathbb{R}$. Since G is continuous and F has left and right limits at any point $x \in \mathbb{R}$, so does $\Delta(x)$. Also $\Delta(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. So there is an x_0 such that either $|\Delta(x_0^+)|$ or $|\Delta(x_0^-)|$ takes the maximum value $\eta = \sup_{x \in \mathbb{R}} |\Delta(x)|$. Say $|\Delta(x_0)| = \eta$. We take $\Delta(x_0) = \eta$, by changing $F - G$ to $G - F$ in the desired inequality, if necessary. Since F is nondecreasing $|G'(x)|$ is bounded by m , $\Delta(x_0 + s) \geq \eta - ms$, $s > 0$. Taking $h = \eta/2m$, $y = x_0 + h$, $x = h - s$, for $|x| \leq h$ one has

$$\Delta(y - x) \geq \frac{\eta}{2} + mx.$$

⁴A comprehensive account of errors of normal approximation for the clt in general multidimensions may be found in Bhattacharya, R. and R. Ranga Rao (2010).

⁵The proof given here follows that given in Feller, W. (1971), vol 2. Feller refers to this particular estimate, attributed to A.C. Berry, as the *smoothing inequality*.

For $|x| > h$, $\Delta(y - x) \geq -\eta$. This, and the properties that v_T is symmetric about $x = 0$, and $\int_{|x|>h} v_T(x) dx \leq \frac{4}{\pi Th}$, provides the asserted bounds as follows:

$$\begin{aligned} & \sup_{y \in \mathbb{R}} \left| \int_{\mathbb{R}} (F(y-x) - G(y-x)) \frac{1}{\pi} \frac{1 - \cos(Tx)}{Tx^2} dx \right| \\ & \geq \int_{\mathbb{R}} \Delta(y-x) v_T(x) dx \\ & \geq \frac{\eta}{2} \left(1 - \frac{4}{\pi Th}\right) - \eta \frac{4}{\pi Th}. \end{aligned} \quad (6.61)$$

This is the asserted lower bound. ■

Proof of Berry–Esseen theorem Let $Q(dx)$ denote the distribution of X_1 . Apply Lemma 4 to $F(x) = F_n(x) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right)$, $x \in \mathbb{R}$, and $G(x) = \Phi(x)$, $x \in \mathbb{R}$, with, using Liapounov inequality,

$$T = \frac{4}{3} \frac{\sigma^3}{\rho} \sqrt{n} \leq \frac{4}{3} \sqrt{n}.$$

The integral on the left side of Lemma 4 is the distribution function of the signed measure $(F_n - \Phi) * v_T$ whose density is given by Fourier inversion as

$$\frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} (\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right) - e^{-\frac{\xi^2}{2}}) \hat{v}_T(\xi) d\xi = \frac{d}{dx} \int_{\mathbb{R}} \frac{e^{-i\xi x}}{-i\xi} (\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right) - e^{-\frac{\xi^2}{2}}) \hat{v}_T(\xi) d\xi.$$

Thus the integral on the right equals the integral on the left in the lemma. Since $|\Phi'(x)| = m < 2/5$, the smoothing lemma now yields

$$\pi |F_n(x) - \Phi(x)| \leq \int_{-T}^T \left| \varphi^n\left(\frac{\xi}{\sigma n}\right) - e^{-\frac{\xi^2}{2}} \right| \frac{d\xi}{|\xi|} + \frac{9.6}{T}. \quad (6.62)$$

Recall (6.58) from which it follows that $|e^{ix} - \sum_{j=0}^{n-1} \frac{(ix)^j}{j!}| \leq \frac{x^n}{n!}$, $x > 0$, $n = 1, 2, \dots$. Thus,

$$|\varphi(x) - 1 + \frac{1}{2}\sigma^2 x^2| = \left| \int_{\mathbb{R}} (e^{ixy} - 1 - ixy + \frac{1}{2}y^2 x^2) Q(dy) \right| \leq \frac{1}{6} \rho |x|^3. \quad (6.63)$$

Since $e^{-x} - 1 + x \leq \frac{x^2}{2}$, $x > 0$, one has

$$\begin{aligned}
\left| \varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right) - e^{-\frac{\xi^2}{2n}} \right| &\leq \left| \varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right) - 1 + \frac{\xi^2}{2n} \right| \\
&\quad + \left| 1 - \frac{\xi^2}{2n} - e^{-\frac{\xi^2}{2n}} \right| \\
&\leq \frac{1}{6\sigma^3 n^{\frac{3}{2}}} |\xi|^3 + \frac{|\xi|^4}{8n^2}. \tag{6.64}
\end{aligned}$$

Also from (6.63), $|\varphi(x)| \leq 1 - \frac{1}{2}\sigma^2 x^2 + \frac{\rho}{6}|x|^3$, for $\frac{1}{2}\sigma^2 x^2 \leq 1$. So for $|\xi| \leq T$ one has

$$\left| \varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right) \right| \leq 1 - \frac{1}{2n}\xi^2 + \frac{\rho}{6\sigma^3 n^{\frac{3}{2}}} |\xi|^3 \leq 1 - \frac{5}{18n}\xi^2 \leq e^{-\frac{5}{18n}\xi^2}.$$

Since $\sigma^3 < \rho$, assume $n \geq 10$; otherwise the theorem is clearly true for $\sqrt{n} \leq 3$. In this case, $|\varphi(\frac{\xi}{\sigma\sqrt{n}})|^{n-1} \leq e^{-\frac{1}{4}\xi^2}$. These estimates can be used to bound the integrand on the right side of (6.62) based on the simple inequality $|a^n - b^n| \leq n|a - b|c^{n-1}$, for $|a| \leq c, |b| \leq c$, with $a = \varphi(\frac{\xi}{\sigma\sqrt{n}}), b = e^{-\frac{\xi^2}{2n}}, c = e^{-\frac{\xi^2}{4}}$. In particular, for $\sqrt{n} > 3$, one obtains using this inequality that

$$\left| \varphi^n\left(\frac{\xi}{\sigma n}\right) - e^{-\frac{\xi^2}{2}} \right| \leq \frac{1}{T} \left(\frac{2}{9}\xi^2 + \frac{1}{18}|\xi|^3 \right) e^{-\frac{1}{4}\xi^2}.$$

Inserting this (integrable) bound on the integrand in (6.62) and integrating by parts, yields

$$\pi |F_n(x) - \Phi(x)| \leq \frac{8}{9}\sqrt{\pi} + \frac{98}{99}.$$

The assertion follows since $\sqrt{\pi} < \frac{9}{5}$ making the right side smaller than 4π . ■

Remark 6.5 After a rather long succession of careful estimates, the constant $c = 3$ in Feller's bound $c\rho/\sigma^3\sqrt{n}$ has been reduced⁶ to $c = 0.5600$ as best to date.

Definition 6.5 A nondegenerate distribution Q on \mathbb{R} , i.e., $Q \neq \delta_{\{c\}}$, is said to be *stable* if for every integer n there is a centering constant c_n and a scaling *index* $\alpha > 0$ such that $n^{-\frac{1}{\alpha}}(X_1 + \cdots + X_n - c_n)$ has distribution Q whenever $X_j, j \geq 1$, are i.i.d. with distribution Q .

It is straightforward to check that the normal distribution and Cauchy distribution are both stable with respective indices $\alpha = 2$ and $\alpha = 1$. Notice also that it follows directly from the definition that every stable distribution Q is **infinitely divisible** in the sense that for any integer $n \geq 1$, there is a probability distribution Q_n such that Q may be exposed as an n -fold convolution $Q = Q_n^{*n}$.

⁶See Shevtsova, I. G. (2010): An Improvement of Convergence Rate Estimates in the Lyapunov Theorem, *Doklady Math.* **82**(3), 862–864.

The following example⁷ illustrates a general framework in which symmetric stable laws arise naturally.

Example 3 (One-dimensional Holtzmark problem) Consider $2n$ points (eg., masses or charges) X_1, \dots, X_{2n} independently and uniformly distributed within an interval $[-n, n]$ so that the density of points is one. Suppose that there is a fixed point (mass, charge) at the origin that exerts an *inverse r th power force* on the randomly distributed points, where $r > 1/2$. That is, the force exerted by the point at the origin on a mass at location x is $-sgn(x)|x|^{-r}$. Let $F_n = -\sum_{j=1}^{2n} \frac{sgn(X_j)}{|X_j|^r}$ denote the total force exerted by the origin on the $2n$ points. The characteristic function of the limit distribution Q_r of F_n as $n \rightarrow \infty$ may be calculated as follows: For $\xi > 0$, using an indicated change of variable,

$$\begin{aligned} \mathbb{E}e^{i\xi F_n} &= \left(\mathbb{E} \cos\left(\frac{\xi sgn(X_1)}{|X_1|^r}\right) \right)^{2n} \\ &= \left(1 - \frac{\xi^{\frac{1}{r}}}{nr} \int_{\xi(\frac{1}{n})^r}^{\infty} (1 - \cos(y))y^{-\frac{r+1}{r}} dy \right)^{2n} \\ &\rightarrow e^{-a\xi^\alpha}, \end{aligned}$$

where $\alpha = 1/r$. This calculation uses the fact that $|1 - \cos(y)| \leq 2$ to obtain integrability on $[1, \infty)$. Also $\frac{1 - \cos(y)}{y^2} \rightarrow 1$ as $y \downarrow 0$ on $(0, 1)$. So one has $0 < a < \infty$ for $0 < \frac{1}{r} < 2$. Similar calculation holds for $\xi < 0$ to obtain $e^{-a|\xi|^\alpha}$. In particular Q_r is a so-called **symmetric stable distribution** with index $\alpha = \frac{1}{r} \in (0, 2)$ in the following sense: If $F_1^{(\infty)}, F_2^{(\infty)}, \dots$ are i.i.d. with distribution Q_r , then $m^{-r}(F_1^{(\infty)} + \dots + F_m^{(\infty)})$ has distribution Q_r . This example includes all such one-dimensional symmetric stable distributions with the notable exception of $\alpha = 2$, corresponding to the normal distribution. The case $\alpha = 2$ represents a different phenomena covered by the central limit theorem in Chapter IV and to be expanded upon in the next chapter.

Exercise Set VI

1. Prove that given $f \in L^2[-\pi, \pi]$ and $\varepsilon > 0$, there exists a continuous function g on $[-\pi, \pi]$ such that $g(-\pi) = g(\pi)$ and $\|f - g\| < \varepsilon$, where $\|\cdot\|$ is the L^2 -norm defined by (6.10). [Hint: By Proposition 2.6 in Appendix A, there exists a continuous function h on $[-\pi, \pi]$ such that $\|f - h\| < \frac{\varepsilon}{2}$. If $h(-\pi) \neq h(\pi)$, modify it on $[\pi - \delta, \pi]$ by a linear interpolation with a value $h(\pi - \delta)$ at $\pi - \delta$ and a value $h(-\pi)$ at π , where $\delta > 0$ is suitably small.]
2. (a) Prove that if $\mathbb{E}|X|^r < \infty$ for some positive integer r , then the characteristic function $\varphi(\xi)$ of X has a continuous r th order derivative $\varphi^{(r)}(\xi) =$

⁷For a more elaborate treatment of the physics of the Holtzmark distribution in higher dimensions see S. Chandreskar (1943): Stochastic problems in physics and astronomy, *Reviews of Modern Physics*, 15(3), reprinted in Wax (1954). The treatment provided here was inspired by Lamperti (1996).

$i^r \int_{\mathbb{R}} x^r e^{i\xi x} P_X(dx)$, where P_X is the distribution of X . In particular, $\varphi^{(r)}(0) = i^r \mathbb{E}X^r$. (b) Prove (6.47) assuming that f and $f^{(j)}$, $1 \leq j \leq r$, are integrable. [Hint: Prove (6.46) and use induction.] (c) If $r \geq 2$ in (b), prove that \hat{f} is integrable.

3. This exercise concerns the *normal* (or *Gaussian*) distribution.

- (i) Prove that for every $\sigma \neq 0$, $\varphi_{\sigma^2, \mu}(x) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < \infty$, is a probability density function (pdf). The probability on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with this pdf is called the *normal* (or *Gaussian*) *distribution with mean μ variance σ^2* , denoted by $\Phi_{\sigma^2, \mu}$. [Hint: Let $c = \int_{-\infty}^{\infty} e^{-x^2/2} dx$. Then $c^2 = \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy = \int_0^{\infty} \int_0^{2\pi} r e^{-r^2/2} d\theta dr = 2\pi$.]
- (ii) Show that $\int_{-\infty}^{\infty} x \varphi_{\sigma^2, \mu}(x) dx = \mu$, $\int_{-\infty}^{\infty} (x - \mu)^2 \varphi_{\sigma^2, \mu}(x) dx = \sigma^2$. [Hint: $\int_{-\infty}^{\infty} (x - \mu) \varphi_{\sigma^2, \mu}(x) dx = 0$, $\int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = 2 \int_0^{\infty} x (-de^{-x^2/2}) = 2 \int_0^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$.]
- (iii) Write $\varphi = \varphi_{1,0}$, the *standard normal density*. Show that its odd-order moments vanish and the even-order moments are given by $\mu_{2n} = \int_{-\infty}^{\infty} x^{2n} \varphi(x) dx = (2n-1) \cdot (2n-3) \cdots 3 \cdot 1$ for $n = 1, 2, \dots$ [Hint: Use integration by parts to prove the recursive relation $\mu_{2n} = (2n-1)\mu_{2n-2}$, $n = 1, 2, \dots$, with $\mu_0 = 1$.]
- (iv) Show $\hat{\Phi}_{\sigma^2, \mu}(\xi) = e^{i\xi\mu - \sigma^2\xi^2/2}$, $\hat{\varphi}(\xi) = e^{-\xi^2/2}$. [Hint: $\hat{\varphi}(\xi) = \int_{-\infty}^{\infty} (\cos(\xi x)) \varphi(x) dx$. Expand $\cos(\xi x)$ in a power series and integrate term by term using (iii).]
- (v) (*Fourier Inversion for $\varphi_{\sigma^2} \equiv \varphi_{\sigma^2, 0}$*). Show $\varphi_{\sigma^2}(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-i\xi x} \hat{\varphi}_{\sigma^2}(\xi) d\xi$. [Hint: $\hat{\varphi}_{\sigma^2}(\xi) = \sqrt{\frac{2\pi}{\sigma^2}} \varphi_{\frac{1}{\sigma^2}}(\xi)$. Now use (iv).]
- (vi) Let $\mathbf{Z} = (Z_1, \dots, Z_k)$ be a random vector where Z_1, Z_2, \dots, Z_k are i.i.d. random variables with standard normal density φ . Then \mathbf{Z} is said to have the *k -dimensional standard normal distribution*. Its pdf (with respect to Lebesgue measure on \mathbb{R}^k) is $\varphi_I(\mathbf{x}) = \varphi(x_1) \cdots \varphi(x_k) = (2\pi)^{-\frac{k}{2}} e^{-\frac{|\mathbf{x}|^2}{2}}$, for $\mathbf{x} = (x_1, \dots, x_k)$. If Σ is a $k \times k$ positive-definite symmetric matrix and $\mu \in \mathbb{R}^k$, then the *normal* (or *Gaussian*) *distribution $\Phi_{\Sigma, \mu}$ with mean $\bar{\mu}$ and dispersion* (or *covariance*) *matrix Σ* has pdf $\varphi_{\Sigma, \mu}(x) = (2\pi)^{-\frac{k}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\{-\frac{1}{2}(x - \mu) \cdot \Sigma^{-1}(x - \mu)\}$, where \cdot denotes the inner (dot) product on \mathbb{R}^k . (a) Show that $\hat{\varphi}_{\Sigma, \mu}(\xi) = \exp\{i\xi \cdot \mu - \frac{1}{2}\xi \cdot \Sigma\xi\}$, $\xi \in \mathbb{R}^k$. (Note that the characteristic function of any absolutely continuous distribution is the Fourier transform of its pdf). (b) If A is a $k \times k$ matrix such that $AA' = \Sigma$, show that for standard normal \mathbf{Z} , $A\mathbf{Z} + \mu$ has the distribution $\Phi_{\Sigma, \mu}$. (c) Prove the *inversion formula* $\varphi_{\Sigma, \mu}(x) = (2\pi)^{-k} \int_{\mathbb{R}^k} \hat{\varphi}_{\Sigma, \mu}(\xi) e^{-i\xi \cdot x} d\xi$, $x \in \mathbb{R}^k$.
- (vii) If (X_1, \dots, X_k) has a k -dimensional Gaussian distribution, show $\{X_1, \dots, X_k\}$ is a collection of independent random variables if and only if they are uncorrelated.

4. Suppose that $\{P_n\}_{n=1}^{\infty}$ is a sequence of Gaussian probability distributions on $(\mathbb{R}^k, \mathcal{B}^k)$ with respective mean vectors $m^{(n)} = (m_1^{(n)}, \dots, m_k^{(n)})$ and variance-

covariance matrices $\Gamma^{(n)} = ((\gamma_{i,j}^{(n)}))_{1 \leq i, j \leq k}$. (i) Show that if $m^{(n)} \rightarrow m$ and $\Gamma^{(n)} \rightarrow \Gamma$ (componentwise) as $n \rightarrow \infty$, then $P_n \Rightarrow P$, where P is Gaussian with mean vector m and variance–covariance matrix Γ . [Hint: Apply the continuity theorem for characteristic functions. Note that in the case of nonsingular Γ one may apply Scheffé’s theorem, or apply Fatou’s lemma to $P_n(G)$, G open.] (ii) Show that if $P_n \Rightarrow P$, then P must be Gaussian. [Hint: Consider the case $k = 1$, $m_n = 0$, $\sigma_n^2 = \int_{\mathbb{R}} x^2 P_n(dx)$. Use the continuity theorem and observe that if σ_n^2 ($n \geq 1$) is unbounded, then $\hat{P}_n(\xi) \equiv e^{-\frac{\sigma_n^2}{2} \xi^2}$ does not converge to a continuous limit at $\xi = 0$.]

5. (Change of Location/Scale/Orientation) Let \mathbf{X} be a k -dimensional random vector and compute the characteristic function of $\mathbf{Y} = A\mathbf{X} + \mathbf{b}$, where A is a $k \times k$ matrix and $\mathbf{b} \in \mathbb{R}^k$.

6. (Fourier Transform, Fourier Series, Inversion, and Plancherel) Suppose f is differentiable and vanishes outside a finite interval, and f' is square-integrable. Derive the inversion formula (6.35) by justifying the following steps. Define $g_N(x) := f(Nx)$, vanishing outside $(-\pi, \pi)$. Let $\sum c_{n,N} e^{inx}$, $\sum c_{n,N}^{(1)} e^{inx}$ be the Fourier series of g_N and its derivative g'_N , respectively.

- (i) Show that $c_{n,N} = \frac{1}{2N\pi} \hat{f}\left(-\frac{n}{N}\right)$.
- (ii) Show that $\sum_{n=-\infty}^{\infty} |c_{n,N}| \leq \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} g_N(x) dx \right| + A \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |g'_N(x)|^2 dx \right)^{1/2} < \infty$, where $A = (2 \sum_{n=1}^{\infty} n^{-2})^{1/2}$. [Hint: Split off $|c_{0,N}|$ and apply Cauchy–Schwarz inequality to $\sum_{n \neq 0} \frac{1}{|n|} |nc_{n,N}|$. Also note that $|c_{n,N}^{(1)}|^2 = |nc_{n,N}|^2$.]
- (iii) Show that for all sufficiently large N , the following convergence is uniform: $f(z) = g_N\left(\frac{z}{N}\right) = \sum_{n=-\infty}^{\infty} c_{n,N} e^{inz/N} = \sum_{n=-\infty}^{\infty} \frac{1}{2N\pi} \hat{f}\left(-\frac{n}{N}\right) e^{inz/N}$.
- (iv) Show that (6.35) follows by letting $N \rightarrow \infty$ in the previous step if $\hat{f} \in L^1(\mathbb{R}, dx)$.
- (v) Show that for any f that vanishes outside a finite interval and is square-integrable, hence integrable, one has, for all sufficiently large N , $\frac{1}{N} \sum_{n=-\infty}^{\infty} \left| \hat{f}\left(\frac{n}{N}\right) \right|^2 = 2\pi \int_{-\infty}^{\infty} |f(y)|^2 dy$. [Hint: Check that $\frac{1}{2\pi} \int_{-\pi}^{\pi} |g_N(x)|^2 dx = \frac{1}{2N\pi} \int_{-\infty}^{\infty} |f(y)|^2 dy$, and $\frac{1}{2\pi} \int_{-\pi}^{\pi} |g'_N(x)|^2 dx = \sum_{n=-\infty}^{\infty} |c_{n,N}|^2 = \frac{1}{4N^2\pi^2} \sum_{n=-\infty}^{\infty} \left| \hat{f}\left(\frac{n}{N}\right) \right|^2$.] Show that the Plancherel identity (6.36) follows in the limit as $N \rightarrow \infty$.

7. (General Inversion Formula and Plancherel Identity)

- (i) Prove (6.35) assuming only that f, \hat{f} are integrable. [Hint: Step 1. Continuous functions with compact support are dense in $L^1 \equiv L^1(\mathbb{R}, dx)$. Step 2. Show that translation $y \rightarrow g(\cdot + y) (\equiv g(x + y), x \in \mathbb{R})$, is continuous on \mathbb{R} into L^1 , for any $g \in L^1$. For this, given $\delta > 0$, find continuous h with compact support such that $\|g - h\|_1 < \delta/3$. Then find $\varepsilon > 0$ such that $\|h(\cdot + y) - h(\cdot + y')\|_1 < \delta/3$ if $|y - y'| < \varepsilon$. Then use $\|g(\cdot + y) - g(\cdot + y')\|_1 \leq \|g(\cdot + y) - h(\cdot + y)\|_1 + \|h(\cdot + y) - h(\cdot + y')\|_1 + \|h(\cdot + y') - g(\cdot + y')\|_1 < \delta$, noting that the Lebesgue integral (measure) is translation invariant. Step 3.

Use Step 2 to prove that $\mathbb{E}f(x + \varepsilon Z) \rightarrow f(x)$ in L^1 as $\varepsilon \rightarrow 0$, where Z is standard normal. *Step 4.* Use (6.45), which does not require f to be continuous, and Step 3, to show that the limit in (6.45) is equal a.e. to f .]

- (ii) (*Plancherel Identity*). Let $f \in L^1 \cap L^2$. Prove (6.36). [Hint: Let $\tilde{f}(x) := f(-x)$, $g = f * \tilde{f}$. Then $g \in L^1$, $|g(x)| \leq \|f\|_2^2$, $g(0) = \|f\|_2^2$. Also $g(x) = \langle f_x, f \rangle$, where $f_x(y) = f(x + y)$. Since $x \rightarrow f_x$ is continuous on \mathbb{R} into L^2 (using arguments similar to those in Step 2 of part (i) above), and $\langle \cdot, \cdot \rangle$ is continuous on $L^2 \times L^2$ into \mathbb{R} , g is continuous on \mathbb{R} . Apply the inversion formula (in part (i)) to get $\|f\|_2^2 = g(0) = \frac{1}{2\pi} \int \hat{g}(\xi) d\xi \equiv \frac{1}{2\pi} \int |\hat{f}(\xi)|^2 d\xi$.]
8. (*Smoothing Property of Convolution*) (a) Suppose μ, ν are probabilities on \mathbb{R}^k , with ν absolutely continuous with pdf f ; $\nu(dx) = f(x)dx$. Show that $\mu * \nu$ is absolutely continuous and calculate its pdf. (b) If $f, g \in L^1(\mathbb{R}^k, dx)$ and if g is bounded and continuous, show that $f * g$ is continuous. (c) If $f, g \in L^1(\mathbb{R}^k, dx)$, and if g and its first r derivatives $g^{(j)}$, $j = 1, \dots, r$ are bounded and continuous, show that $f * g$ is r times continuously differentiable. [Hint: Use induction.]
9. Suppose f, \tilde{f} are integrable on (\mathbb{R}, dx) . Show $\tilde{f}(x) = 2\pi f(-x)$.
10. Let $Q(dx) = \frac{1}{2} \mathbf{1}_{[-1, 1]}(x) dx$ be the uniform distribution on $[-1, 1]$.
- (i) Find the characteristic functions of Q and $Q^{*2} \equiv Q * Q$.
 - (ii) Show that the probability with pdf $c \sin^2 x / x^2$, for appropriate normalizing constant c , has a characteristic function with compact support and compute this characteristic function. [Hint: Use Fourier inversion for $f = \hat{Q}^2$.]
11. Derive the multidimensional extension of the Fourier inversion formula.
12. Show that if Q is a stable distribution symmetric about 0 with exponent α , then $c_n = 0$ and $0 < \alpha \leq 2$. [Hint: $\hat{Q}(\xi)$ must be real by symmetry, and positivity follows from the case $n = 2$ in the definition.]
13. Show that
- (i) The Cauchy distribution with pdf $(\pi(1 + x^2))^{-1}$, $x \in \mathbb{R}$, has characteristic function $e^{-|\xi|}$.
 - (ii) The characteristic function of the double-sided exponential distribution $\frac{1}{2} e^{-|x|} dx$ is $(1 + \xi^2)^{-1}$. [Hint: Use integration by parts twice to show $\int_{-\infty}^{\infty} e^{i\xi x} (\frac{1}{2} e^{-|x|}) dx \equiv \int_0^{\infty} e^{-x} \cos(\xi x) dx = (1 + \xi^2)^{-1}$.]
14. (i) Give an example of a pair of *dependent* random variables X, Y such that the distribution of their sum is the convolution of their distributions. [Hint: Consider the Cauchy distribution with $X = Y$.] (ii) Give an example of a non-Gaussian bivariate distribution such that the marginals are Gaussian. [Hint: Extend the proof of Theorem 6.7.]
15. Show that if φ is the characteristic function of a probability then φ must be uniformly continuous on \mathbb{R} .
16. (*Symmetric Distributions*) (i) Show that the characteristic function of \mathbf{X} is real-valued if and only if \mathbf{X} and $-\mathbf{X}$ have the same distribution. (ii) A *symmetrization* of (the distribution of) a random variable \mathbf{X} may be defined by (the distribution of) $\mathbf{X} - \mathbf{X}'$, where \mathbf{X}' is an independent copy of \mathbf{X} , i.e., independent of \mathbf{X} and

having the same distribution as \mathbf{X} . Express symmetrization of a random variable in terms of its characteristic function.

17. (*Multidimensional Gaussian characterization*) Suppose that $\mathbf{X} = (X_1, \dots, X_k)$ is a k -dimensional random vector having a positive pdf $f(x_1, \dots, x_k)$ on \mathbb{R}^k ($k \geq 2$). Assume that (a) f is differentiable, (b) X_1, \dots, X_k are independent, and (c) have an *isotropic density*, i.e., $f(x_1, \dots, x_k)$ is a function of $\|x\|^2 = x_1^2 + \dots + x_k^2$, $(x_1, \dots, x_k) \in \mathbb{R}^k$. Show that X_1, \dots, X_k are i.i.d. normal with mean zero and common variance. [*Hint*: Let f_j denote the marginal pdf of X_j and argue that $\frac{f_j}{2x_j f_j}$ must be a constant.]
18. (i) Show that the functions $\{e_\xi : \xi \in \mathbb{R}^k\}$ defined by $e_\xi(x) := \exp(i\xi \cdot \mathbf{x})$, $\mathbf{x} \in \mathbb{R}^k$ constitute a measure-determining class for probabilities on $(\mathbb{R}^k, \mathcal{B}^k)$. [*Hint*: Given two probabilities P, Q for which the integrals of the indicated functions agree, construct a sequence by $P_n = P \forall n = 1, 2, \dots$ whose characteristic functions will obviously converge to that of Q .]
 (ii) Show that the closed half-spaces of \mathbb{R}^k defined by $F_a := \{x \in \mathbb{R}^k : x_j \leq a_j, 1 \leq j \leq k\}$, $a = (a_1, \dots, a_k)$ constitute a measure-determining collection of Borel subsets of \mathbb{R}^k . [*Hint*: Use a trick similar to that above.]
19. Compute the distribution with characteristic function $\varphi(\xi) = \cos^2(\xi)$, $\xi \in \mathbb{R}^1$.
20. (*Fourier Inversion for Lattice Random Variables*)
 (i) Let $p_j, j \in \mathbb{Z}$, be a probability mass function (pmf) of a probability distribution Q on the integer lattice \mathbb{Z} . Show that the Fourier transform \hat{Q} is periodic with period 2π , and derive the inversion formula $p_j = (2\pi)^{-1} \int_{(-\pi, \pi]} e^{-ij\xi} \hat{Q}(\xi) d\xi$.
 (ii) Let Q be a *lattice distribution of span* $h > 0$, i.e., for some a_0 , $Q(\{a_0 + jh : j = 0, \pm 1, \pm 2, \dots\}) = 1$. Show that \hat{Q} is periodic with period $2\pi/h$ and write down an inversion formula. (iii) Extend (i), (ii) to the multidimensional lattice distributions with \mathbb{Z}^k in place of \mathbb{Z} .
21. (*Parseval's Relation*) Let $f, g, \in L^2([-\pi, \pi])$, with Fourier coefficients $\{c_n\}, \{d_n\}$, respectively. Prove that $\sum_n c_n \bar{d}_n = \frac{1}{2\pi} \int_{(-\pi, \pi]} f(x) \bar{g}(x) dx \equiv \langle f, g \rangle$.
 (ii) Let $f, g \in L^2(\mathbb{R}^k, dx)$ with Fourier transforms \hat{f}, \hat{g} . Prove that $\langle \hat{f}, \hat{g} \rangle = 2\pi \langle f, g \rangle$. [*Hint*: Use (a) the Plancherel identity and (b) the *polar identity* $4\langle f, g \rangle = \|f + g\|^2 - \|f - g\|^2$.]
22. (i) Let φ be continuous and positive-definite on \mathbb{R} in the sense of Bochner, and $\varphi(0) = 1$. Show that the sequence $\{c_j \equiv \varphi(j) : j \in \mathbb{Z}\}$ is positive-definite in the sense of Herglotz (6.27). (ii) Show that there exist distinct probability measures on \mathbb{R} whose characteristic functions agree at all integer points.
23. Show that the "triangular function" $\hat{f}(\xi) = (1 - |\xi|)^+$ is the characteristic function of $f(x) = 2 \frac{1 - \cos(x)}{x^2}$, $x \in \mathbb{R}$. [*Hint*: Consider the characteristic function of the convolution of two uniform distributions on $[-1/2, 1/2]$ and Fourier inversion.] Also show that $1 - \cos(x) \geq x^2/4$ for $|x| \leq \pi/3$. [*Hint*: Use $\cos(y) \geq 1/2$ and $\sin(y) \geq y$ for $0 < y < \pi/3$ in the formula $1 - \cos(x) = \int_0^x \sin(y) dy$.]

24. (*Chung–Fuchs*) For the one-dimensional random walk show that if $\frac{S_n}{n} \rightarrow 0$ in probability as $n \rightarrow \infty$, i.e., WLLN holds, then 0 is neighborhood recurrent. [*Hint*: Using the lemma for the proof of Chung–Fuchs, for any positive integer m and $\delta, \varepsilon > 0$, $\sum_{n=0}^{\infty} P(S_n \in B_\varepsilon) \geq \frac{1}{2m} \sum_{n=0}^{\infty} P(S_n \in B_{m\varepsilon}) \geq \frac{1}{2m} \sum_{n=0}^{m\delta-1} P(S_n \in B_{\delta\varepsilon})$, using monotonicity of $r \rightarrow P(S_n \in B_r)$. Let $m \rightarrow \infty$ to obtain for the indicated Cesàro average, using $\lim_{n \rightarrow \infty} P(S_n \in B_{\delta\varepsilon}) = 1$ from the WLLN hypothesis, that $\sum_{n=0}^{\infty} P(S_n \in B_\varepsilon) \geq \frac{1}{2\delta}$. Let $\delta \rightarrow 0$ and apply the Lemma 2.]
25. This exercise provides a version of the Chung–Fuchs Fourier analysis criteria for the case of random walks on the integer lattice. Show that
- $P(S_n = 0) = \frac{1}{(2\pi)^k} \int_{[-\pi, \pi]^k} \varphi(\xi) d\xi$, where $\varphi(\xi) = \mathbb{E}e^{i\xi \cdot X_1}$. [*Hint*: Use Fourier inversion formula.]
 - $\sum_{n=0}^{\infty} r^n P(S_n = 0) = \frac{1}{(2\pi)^k} \int_{[-\pi, \pi]^k} \operatorname{Re}\left(\frac{1}{1-r\varphi(\xi)}\right) d\xi$.
 - The lattice random walk $\{S_n : n = 0, 1, 2, \dots\}$ is recurrent if and only if $\lim_{r \uparrow 1} \int_{[-\pi, \pi]^k} \operatorname{Re}\left(\frac{1}{1-r\varphi(\xi)}\right) d\xi = \infty$. [*Hint*: Justify passage to the limit $r \uparrow 1$ and use Borel–Cantelli lemma.]
26. (i) Use the Chung–Fuchs criteria, in particular Corollary 6.15 and its converse, to determine whether the random walk with symmetric Cauchy displacement distribution is recurrent or transient. (ii) Extend this to symmetric stable displacement distributions with exponent $0 < \alpha \leq 2$.⁸
27. Show that 0 is neighborhood recurrent for the random walk if and only if $\sum_{n=0}^{\infty} P(S_n \in B_1) = \infty$.
28. Prove that the set of trigonometric polynomials is dense in $L^2([-\pi, \pi], \mu)$, where μ is a finite measure on $[-\pi, \pi]$.
29. Establish the formula $\int_{\mathbb{R}} g(x) \mu * \nu(dx) = \int_{\mathbb{R}} \int_{\mathbb{R}} g(x+y) \mu(dx) \nu(dy)$ for any bounded measurable function g .

⁸Surprisingly, recurrence and heavy tails may coexist, see Shepp, L (1964): Recurrent random walks with arbitrarily large steps, *Bull. Amer. Math. Soc.*, v. 70, 540–542; Grey, D.R. (1989): Persistent random walks may have arbitrarily large tails, *Adv. Appld. Probab.* **21**, 229–230.