Integration and Modern Analysis
John Benedetto and Wojtek Czaja

B Fourier analysis

B.1 Fourier transforms

Throughout the text of this book there are many examples and problems dealing with the notion of Fourier series and Fourier transforms, see, e.g., Examples 3.3.4, 3.6.6, 4.4.7, 4.5.7, and Problems 3.14, 3.28, 4.5, 4.17, 4.44, 5.20. In this appendix we give a brief outline of the basic elementary theory of classical Fourier analysis. There are many excellent texts and expositions including [39], [66], [268], [289], [437], [509], [151], [262], [455], [27].

Definition B.1.1. Fourier transform on $L^1_m(\mathbb{R})$
a. The *Fourier transform* of $f \in L^1_m(\mathbb{R})$ is the function F defined as

$$F(\xi) = \int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\, dx, \quad \xi \in \widehat{\mathbb{R}} = \mathbb{R}. \tag{B.1}$$

Notationally, we write the pairing between the functions f and F in the following way:

$$\hat{f} = F.$$

The space of Fourier transforms of $L^1_m(\mathbb{R})$ functions is denoted by $A(\widehat{\mathbb{R}})$, i.e.,

$$A(\widehat{\mathbb{R}}) = \{F : \widehat{\mathbb{R}} \to \mathbb{C} :\ \exists\, f \in L^1_m(\mathbb{R}) \text{ such that } \hat{f} = F\}.$$

b. Let $f \in L^1_m(\mathbb{R})$. The *Fourier transform inversion formula* is

$$f(x) = \int_{\widehat{\mathbb{R}}} F(\xi)e^{2\pi i x\xi}\, d\xi.$$

We use the notation, $\check{F} = f$, to denote this inversion. See [39], pages 2–3, for a formal intuitive derivation of the Fourier transform inversion formula by means of a form of the uncertainty principle. The *Jordan pointwise inversion formula* gives an explicit theorem. Also, see Theorem B.3.7.

Theorem B.1.2. Jordan inversion formula
Let $f \in L^1_m(\mathbb{R})$. Assume that $f \in BV([x_0 - \varepsilon, x_0 + \varepsilon])$, for some $x_0 \in \mathbb{R}$ and $\varepsilon > 0$. Then,

$$\frac{f(x_0+) + f(x_0-)}{2} = \lim_{M\to\infty} \int_{-M}^{M} \hat{f}(\xi)e^{2\pi i x\xi}\, d\xi.$$

Example B.1.3. $f \in L^1_m(\mathbb{R})$ **does not imply** $\hat{f} \in L^1_m(\widehat{\mathbb{R}})$
Let
$$f(x) = H(x)e^{-2\pi i r x},$$
where $r > 0$ and H is the *Heaviside function*, i.e., $H = \mathbb{1}_{[0,\infty)}$. Then,
$$\hat{f}(\xi) = \frac{1}{2\pi(r + i\xi)} \notin L^1_m(\widehat{\mathbb{R}}).$$

Theorem B.1.4. Algebraic properties of Fourier transforms
a. *Let* $f_1, f_2 \in L^1_m(\mathbb{R})$*, and assume* $c_1, c_2 \in \mathbb{C}$*. Then,*
$$\forall\, \xi \in \widehat{\mathbb{R}}, \quad (c_1 f_1 + c_2 f_2)\hat{}(\xi) = c_1 \hat{f}_1(\xi) + c_2 \hat{f}_2(\xi).$$

b. *Let* $f \in L^1_m(\mathbb{R})$ *and assume* $F = \hat{f} \in L^1_m(\widehat{\mathbb{R}})$*. Then,*
$$\forall\, x \in \mathbb{R}, \quad \widehat{F}(x) = f(-x).$$

c. *Let* $f \in L^1_m(\mathbb{R})$*. Then,*
$$\forall\, \xi \in \widehat{\mathbb{R}}, \quad \widehat{\overline{f}}(\xi) = \overline{\hat{f}(\xi)}.$$

For a fixed $\gamma \in \mathbb{R}$ we set
$$e_\gamma(x) = e^{2\pi i x \gamma}.$$

For a fixed $t \in \mathbb{R}$ and for a given function $f : \mathbb{R} \to \mathbb{C}$, the *translation operator* τ_t is defined as
$$\tau_t(f)(x) = f(x - t),$$
and, for a fixed $\lambda \in \mathbb{R} \setminus \{0\}$ and for a given function $f : \mathbb{R} \to \mathbb{C}$, the *dilation operator* is defined by the dilation formula,
$$f_\lambda(x) = \lambda f(\lambda x).$$

Example B.1.5. Dilation and the Poisson function
If $f(x) = e^{-2\pi r|x|}$, $r > 0$, then
$$\hat{f}(\xi) = \frac{1}{r} P_{1/r}(\xi) = \frac{1}{r}\frac{1}{\pi(1 + \xi^2/r^2)} \in L^1_m(\widehat{\mathbb{R}}),$$
where $P(\xi) = 1/(\pi(1 + \xi^2))$ is the *Poisson function.*

Proposition B.1.6. *Let* $f \in L^1_m(\mathbb{R})$*, let* $t \in \mathbb{R}$*, let* $\gamma \in \widehat{\mathbb{R}}$*, and let* $\lambda \in \widehat{\mathbb{R}} \setminus \{0\}$*. Then,*

i. $(e_\gamma f)\hat{}(\xi) = \tau_\gamma(\hat{f})(\xi)$,
ii. $(\tau_t(f))\hat{}(\xi) = e_{-t}(\xi)\hat{f}(\xi)$,
iii. $(f_\lambda)\hat{}(\xi) = (\lambda/|\lambda|)\hat{f}(\xi/\lambda)$.

In the following result we shall assume the existence of a pair (f, F) of functions, such that

$$F = \hat{f} \quad \text{and} \quad f = \check{F}. \tag{B.2}$$

We shall not be concerned with correct mathematical hypotheses for asserting the existence of the Fourier transform or of the inversion formula.

Proposition B.1.7. *Assume there exists a pair (f, F) of functions which satisfies (B.2).*

a. *f is real if and only if*

$$\overline{F(\xi)} = F(-\xi).$$

In this case,

$$F(\xi) = \int_{\mathbb{R}} f(x)\cos(2\pi x\xi)\,dx - i\int_{\mathbb{R}} f(x)\sin(2\pi x\xi)\,dx$$

and

$$f(x) = 2\mathrm{Re}\int_0^\infty F(\xi)e^{2\pi i x\xi}\,d\xi.$$

b. *f is real and even if and only if F is real and even. In this case,*

$$F(\xi) = 2\int_0^\infty f(x)\cos(2\pi x\xi)\,dx$$

and

$$f(x) = 2\int_0^\infty F(\xi)\cos(2\pi x\xi)\,d\xi.$$

c. *f is real and odd if and only if F is odd and imaginary. In this case,*

$$F(\xi) = -2i\int_0^\infty f(x)\sin(2\pi x\xi)\,dx$$

and

$$f(x) = 2i\int_0^\infty F(\xi)\sin(2\pi x\xi)\,d\xi.$$

Example B.1.8. The Gaussian

Let $f(x) = e^{-\pi r x^2}$, $r > 0$. We could calculate $\hat{f}$ by means of contour integrals, but we choose a real, and by now classical approach [168], page 476. By definition of $\hat{f}$, which is real and even, we have

$$(\hat{f})'(\xi) = -2\pi i\int_{\mathbb{R}} te^{-\pi r x^2}e^{-2\pi i x\xi}\,dx. \tag{B.3}$$

Noting that

$$\frac{d}{dx}(e^{-\pi r x^2}) = -2\pi r x\,e^{-\pi r x^2},$$

we rewrite (B.3) as

$$\begin{aligned}(\hat{f})'(\xi) &= -2\pi i \int \frac{-1}{2\pi r}(e^{-\pi r x^2})' e^{-2\pi i x\xi}\,dx \\ &= \frac{i}{r}\left[e^{-\pi r x^2} e^{-2\pi i x\xi}\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} e^{-\pi r x^2}(-2\pi i\xi)e^{-2\pi i x\xi}\,dx\right] \\ &= \frac{-2\pi\xi}{r}\hat{f}(\xi).\end{aligned}$$

Thus, $\hat{f}$ is a solution of the differential equation,

$$F'(\xi) = -\frac{2\pi\xi}{r}F(\xi); \tag{B.4}$$

and (B.4) is solved by elementary means with solution

$$F(\xi) = Ce^{-\pi\xi^2/r}.$$

Taking $\xi = 0$ and using the definition of the Fourier transform, we see that

$$C = \int_{\mathbb{R}} e^{-\pi r x^2}\,dx.$$

In order to calculate C we first evaluate $a = \int_0^\infty e^{-u^2}\,du$.

$$\begin{aligned} a^2 &= \int_0^\infty e^{-s^2}\,ds \int_0^\infty e^{-t^2}\,dt \\ &= \int_0^\infty \int_0^\infty e^{-(s^2+t^2)}\,dsdt = \int_0^{\pi/2}\int_0^\infty e^{-r^2}\,rdrd\theta \\ &= \frac{\pi}{4}\int_0^\infty e^{-u}\,du = \frac{\pi}{4}.\end{aligned}$$

Thus, $a = \frac{\sqrt{\pi}}{2}$ and so

$$\int_{\mathbb{R}} e^{-u^2}\,du = \sqrt{\pi}.$$

Consequently,

$$C = \int_{\mathbb{R}} e^{-\pi r x^2}\,dt = \frac{1}{\sqrt{\pi r}}\int_{\mathbb{R}} e^{-u^2}\,du = \frac{1}{\sqrt{r}}.$$

Therefore, we have shown that

$$\hat{f}(\xi) = \frac{1}{\sqrt{r}}e^{-\pi\xi^2/r}.$$

We write

$$G(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$$

so that if $\lambda > 0$, then

$$(G_\lambda)\hat{}(\xi) = e^{-(\pi\xi/\lambda)^2}.$$

In particular,

$$\frac{1}{\sqrt{r}}(G_{\sqrt{\pi r}})\hat{} = G_{\sqrt{\pi/r}}$$

and hence $(G_{\sqrt{\pi}})\hat{} = G_{\sqrt{\pi}}$. We refer to G as the *Gauss function* or *Gaussian*, and note that $\int_{\mathbb{R}} G(x)\, dx = 1$.

B.2 Analytic properties of Fourier transforms

Theorem B.2.1. Boundedness and continuity of Fourier transforms
Assume $f \in L^1_m(\mathbb{R})$.

a. $\forall\, \xi \in \widehat{\mathbb{R}},\ |\hat{f}(\xi)| \leq \|f\|_1$.

b. $\forall\, \varepsilon > 0,\ \exists\, \delta > 0$ *such that* $\forall\, \xi$ *and* $\forall\, \zeta$, *for which* $|\zeta| < \delta$, *we have* $|\hat{f}(\xi + \zeta) - \hat{f}(\xi)| < \varepsilon$, *i.e.,* $\hat{f}$ *is uniformly continuous.*

Proof. ***a.*** Part a is immediate from the definition of Fourier transform.

b. First, we note that

$$|\hat{f}(\xi + \zeta) - \hat{f}(\xi)| \leq \int_{\mathbb{R}} |f(x)| \left|e^{-2\pi i x\zeta} - 1\right|\, dx.$$

Let $g_\zeta(x) = |f(x)| \left|e^{-2\pi i x\zeta} - 1\right|$. Since $\lim_{\zeta \to 0} g_\zeta(x) = 0$ for all $x \in \mathbb{R}$, and since $|g_\zeta(x)| \leq 2|f(x)|$, we can use LDC to obtain

$$\lim_{\zeta \to 0} \int_{\mathbb{R}} g_\zeta(x)\, dx = 0.$$

This limit holds independently of ξ. Consequently, we have

$$\forall \varepsilon > 0,\ \exists \zeta_0 > 0 \text{ such that } \forall\, \zeta \in (-\zeta_0, \zeta_0) \text{ and } \forall\, \xi \in \widehat{\mathbb{R}},\quad |\hat{f}(\xi+\zeta) - \hat{f}(\xi)| < \varepsilon.$$

This is the desired uniform continuity.

□

The next result has essentially the same proof as the *Riemann-Lebesgue lemma for* $L^1_m(\mathbb{T})$, see Theorem 3.6.4.

Theorem B.2.2. Riemann–Lebesgue lemma for $L^1_m(\mathbb{R})$
Assume $f \in L^1_m(\mathbb{R})$.

$$\lim_{|\xi| \to \infty} \hat{f}(\xi) = 0.$$

Theorem B.2.3. Differentiation of Fourier transforms
a. *Assume that $f^{(n)}$, $n \geq 1$, exists everywhere and that*

$$f(\pm\infty) = \ldots = f^{(n-1)}(\pm\infty) = 0.$$

Then,

$$(f^{(n)})\hat{}(\xi) = (2\pi i\xi)^n \hat{f}(\xi).$$

b. *Assume that $x^n f(x) \in L^1_m(\mathbb{R})$, for some $n \geq 1$. Then, $x^k f(x) \in L^1_m(\mathbb{R})$, $k = 1, \ldots, n-1$, $(\hat{f})', \ldots, (\hat{f})^{(n)}$ exist everywhere, and*

$$\forall\, k = 0, \ldots, n, \quad \left((-2\pi i\cdot)^k f(\cdot)\right)\hat{}(\xi) = \hat{f}^{(k)}(\xi).$$

Proof. ***a.*** In the statement of the theorem, $f(\pm\infty) = 0$ denotes the facts that $\lim_{x\to\pm\infty} f(x) = 0$. Using integration by parts (Theorem 4.6.3) we compute

$$\int_{-S}^{T} f^{(n)}(x)e^{-2\pi i x\xi}\,dx = f^{(n-1)}(x)e^{-2\pi i x\xi}\Big|_{-S}^{T} + 2\pi i\xi \int_{-S}^{T} f^{(n-1)}(x)e^{-2\pi i x\xi}\,dx.$$

Iterating this procedure we obtain that

$$\int_{-S}^{T} f^{(n)}(x)e^{-2\pi i x\xi}\,dx = \sum_{j=0}^{n-1}(2\pi i\xi)^j\, f^{(n-(j+1))}(x)e^{-2\pi i x\xi}\Big|_{-S}^{T} + (2\pi i\xi)^n \int_{-S}^{T} f(x)e^{-2\pi i x\xi}\,dx.$$

Letting $S, T \to \infty$, the right hand side converges to $(2\pi i\xi)^n \hat{f}(\xi)$ and the result is proved.

b. Without loss of generality we assume that $n = 1$ and we fix $\xi \in \widehat{\mathbb{R}}$. Then,

$$\frac{\hat{f}(\xi+\zeta) - \hat{f}(\xi)}{\zeta} = \int_{\mathbb{R}} f(x)e^{-2\pi i x\xi}\left(\frac{e^{-2\pi i x\zeta} - 1}{\zeta}\right)\,dx.$$

If we denote the integrand on the right hand side by $f(x,\zeta)$, then we have

$$|f(x,\zeta)| \leq 4\pi|xf(x)|,$$

which follows from the inequality

$$\left|\frac{e^{-2\pi i x\zeta} - 1}{\zeta}\right| \leq 4\pi|x|.$$

Moreover, we have that

$$\lim_{\zeta\to 0} f(x,\zeta) = -2\pi i x f(x)e^{-2\pi i x\xi}.$$

Thus, we can use LDC to assert that

$$\lim_{\zeta \to 0} \frac{\hat{f}(\xi + \zeta) - \hat{f}(\xi)}{\zeta} = \int_{\mathbb{R}} (-2\pi i x) f(x) e^{-2\pi i x \xi} \, dx.$$

□

Example B.2.4. The role of absolute continuity

a. We recall from Section 4.6 that the facts that f is differentiable m-*a.e.* and that $f, f' \in L^1_m(\mathbb{R})$ do not imply that f' is absolutely continuous. Therefore, it is *not* necessarily true that

$$\forall\, \xi \in \widehat{\mathbb{R}}, \quad (f')\hat{}(\xi) = 2\pi i \xi \hat{f}(\xi). \tag{B.5}$$

Consider the Cantor function C_C associated with the usual Cantor set C, see Example 1.3.17. Let

$$f(x) = C_C(x+1)\mathbb{1}_{[-1,0]} + (1 - C_C(x))\mathbb{1}_{[0,1]}.$$

Clearly, f is a continuous compactly supported function of bounded variation on $\mathbb{R}$ for which $f' = 0$ m-*a.e.*. In particular, $f, f' \in L^1_m(\mathbb{R})$ and (B.5) fails, cf., Theorem 4.6.8.

b. In obtaining the formula

$$\int_{\mathbb{R}} f^{(n)}(x) e^{-2\pi i x \xi} \, dx = (2\pi i \xi)^n \hat{f}(\xi), \tag{B.6}$$

we use a subtle fact that everywhere differentiability of $f^{(n-1)}$ allows us to deduce it is absolutely continuous, see Theorem 4.6.7.

Equation (B.6) is also valid, without the aforementioned subtlety, if the hypotheses, that $f^{(n-1)}$ is everywhere differentiable and $f^{(n)} \in L^1_m(\mathbb{R})$, are replaced by the hypothesis that $f^{(n)}$ be piecewise continuous.

c. The assumption in Theorem B.2.3 that $f(\pm\infty) = \ldots = f^{(n-1)}(\pm\infty) = 0$ is not necessary. For simplicity, let $n = 1$, and assume that $f, f' \in L^1_m(\mathbb{R})$ and that f is absolutely continuous. For fixed $a \in \mathbb{R}$ and $c \in \mathbb{C}$, set $F(x) = c + \int_a^x f'(t)\, dt$. By FTC-I, F is absolutely continuous and $F' = f'$ m-*a.e.* Since f is absolutely continuous, we have $f = F + C$ on $[a, \infty)$, for some $C \in \mathbb{C}$, and consequently,

$$\forall\, x \in [a, \infty), \quad f(x) = F(a) + C + \int_a^x f'(t)\, dt.$$

Therefore, $f(a) = F(a) + C$ and

$$f(x) - f(a) = \int_a^x f'(t)\, dt.$$

This observation, combined with the fact that $f' \in L^1_m(\mathbb{R})$, implies that $\lim_{x \to \pm\infty} f(x)$ exist. Moreover, since $f \in L^1_m(\mathbb{R})$, $\lim_{x \to \pm\infty} f(x) = 0$.

Example B.2.5. $C_0(\widehat{\mathbb{R}}) \setminus A(\widehat{\mathbb{R}}) \neq \emptyset$
Theorems B.2.1 and B.2.2 allow us to conclude that $A(\widehat{\mathbb{R}}) \subseteq C_0(\widehat{\mathbb{R}})$. It is not difficult to see that the inclusion is proper. Indeed, let F be defined as

$$F(\xi) = \begin{cases} \frac{1}{\log(\xi)}, & \text{if } \xi > e, \\ \frac{\xi}{e}, & \text{if } 0 \leq \xi \leq e, \end{cases}$$

on $[0, \infty)$ and as $-F(-\xi)$ on $(-\infty, 0]$. Then, $F \in C_0(\widehat{\mathbb{R}})$. The fact that $F \notin A(\widehat{\mathbb{R}})$ depends on the divergence of

$$\int_e^\infty \frac{1}{\xi \log(\xi)}\, d\xi,$$

cf., [39], [195].

This function F is not an isolated example. In fact, $A(\widehat{\mathbb{R}})$ is a set of first category in $C_0(\widehat{\mathbb{R}})$. Even more, a Baire category argument can also be used to show the existence of $F \in C_c(\widehat{\mathbb{R}})$ for which $F \notin A(\widehat{\mathbb{R}})$. Explicit examples of such functions are more difficult to construct, but it is possible to do, e.g., define the *butterfly function*,

$$B(\xi) = \begin{cases} \frac{1}{n}\sin(2\pi 4^n \xi), & \text{if } \frac{1}{2^{n+1}} < |\xi| \leq \frac{1}{2^n}, \\ 0, & \text{if } \xi = 0 \text{ or } |\xi| > \frac{1}{2}, \end{cases}$$

see [224].

B.3 Approximate identities

In Problem 3.5 we defined the convolution $f * g$ of $f, g \in L^1_m(\mathbb{R})$ to be

$$f * g(x) = \int_{\mathbb{R}} f(t)g(x-t)\, dt = \int_{\mathbb{R}} f(x-t)g(t)\, dt.$$

Proposition B.3.1. *Let $f, g \in L^1_m(\mathbb{R})$. Then, $f * g \in L^1_m(\mathbb{R})$ and*

$$(f * g)\hat{}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

Proof. The assertion that $f * g \in L^1_m(\mathbb{R})$ is part of Problem 3.5. Thus, we can use the Fubini–Tonelli theorems (Theorem 3.7.5 and Theorem 3.7.8) to compute

$$\begin{aligned} (f * g)\hat{}(\xi) &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-t)g(t)e^{-2\pi i x\xi}\, dt\, dx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-t)g(t)e^{-2\pi i(x-t)\xi}e^{-2\pi i t\xi}\, dt\, dx \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-t)e^{-2\pi i(x-t)\xi}\, dx \right) g(t)e^{-2\pi i t\xi}\, dt \\ &= \int_{\mathbb{R}} \hat{f}(\xi)g(t)e^{-2\pi i t\xi}\, dt = \hat{f}(\xi)\hat{g}(\xi). \end{aligned}$$

□

This innocent proposition is actually a raison d'être for transform methods, generally, and for the Fourier transform, in particular, see [39].

The following notion is critical in approximating $\delta \in M_b(\mathbb{R})$, and for providing examples in applications including signal processing and spectral estimation.

Definition B.3.2. Approximate identity
An *approximate identity* is a family $\{K_{(\lambda)} : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$ of functions with the properties:

i. $\forall \lambda > 0, \quad \int_{\mathbb{R}} K_{(\lambda)}(x)\, dx = 1,$

ii. $\exists\, M > 0$ such that $\forall\, \lambda > 0, \quad \|K_{(\lambda)}\|_1 \leq M,$

iii. $\forall\, \delta > 0, \quad \lim_{\lambda \to \infty} \int_{|x| \geq \delta} |K_{(\lambda)}(x)|\, dx = 0,$

cf., Problem 3.28*c*.

Proposition B.3.3. *Let $K \in L^1_m(\mathbb{R})$ have the property that $\int_{\mathbb{R}} K(x)\, dx = 1$. Then, the family, $\{K_\lambda : K_\lambda(x) = \lambda K(\lambda x),\ \lambda > 0\} \subseteq L^1_m(\mathbb{R})$, of dilations of K is an approximate identity.*

Example B.3.4. Approximate identities
a. The *Fejér function* W is defined as

$$W(x) = \frac{1}{2\pi} \left(\frac{\sin(x/2)}{x/2} \right)^2,$$

cf., Problem 3.28. The *Fejér function* W is non-negative and $\int_{\mathbb{R}} W(x)\, dx = 1$. Thus, the *Fejér kernel* $\{W_\lambda : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$, defined as the family of dilations of W, is an approximate identity. LIPÓT FEJÉR's name at birth was Weisz, whence W. (In Hungarian, "white" is "fehér".)

b. The *Dirichlet function* D is the function

$$D(x) = \frac{\sin(x)}{\pi x}.$$

Although $\int_{\mathbb{R}} D(t)\, dt = 1$, we have $D \notin L^1_m(\mathbb{R})$, and so the *Dirichlet kernel* $\{D_\lambda : \lambda > 0\}$ is not an approximate identity.

c. The *Poisson function*, defined in Example B.1.5, is

$$P(x) = \frac{1}{\pi(1 + x^2)},$$

and it satisfies $\int_{\mathbb{R}} P(x)\, dx = 1$. Thus, the *Poisson kernel* $\{P_\lambda : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$ is an approximate identity.

d. The *Gauss function* defined in Example B.1.8 is

$$G(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}.$$

G is positive and $\int_{\mathbb{R}} G(x)\, dx = 1$; and, thus, the *Gauss kernel* $\{G_\lambda : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$ is an approximate identity.

Theorem B.3.5. Approximation and uniqueness
Let $f \in L^1_m(\mathbb{R})$.

a. *If $\{K_{(\lambda)} : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$ is an approximate identity, then*

$$\lim_{\lambda\to\infty} \|f - f * K_{(\lambda)}\|_1 = 0.$$

b. *We have*

$$\lim_{\lambda\to\infty} \int_{\mathbb{R}} \left| f(x) - \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\xi|}{\lambda}\right) \hat{f}(\xi) e^{2\pi i x\xi}\, d\xi \right| dx = 0.$$

c. *If $\hat{f} = 0$ on $\widehat{\mathbb{R}}$, then f is the 0 function.*

Proof. **a.** We use the fact that $\int_{\mathbb{R}} K_{(\lambda)}(x)\, dx = 1$ to compute

$$\begin{aligned} \|f - f * K_{(\lambda)}\|_1 &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} K_{(\lambda)}(t) f(x)\, dt - \int_{\mathbb{R}} K_{(\lambda)}(t) f(x-t)\, dt \right| dx \\ &\le \int_{\mathbb{R}} |K_{(\lambda)}(t)| \left(\int_{\mathbb{R}} |f(x) - f(x-t)|\, dx \right) dt. \end{aligned}$$

Let $\varepsilon > 0$. Using the result from Problem 3.14*b*, there is $\delta > 0$ with the property that

$$\forall\, |t| < \delta, \quad \int_{\mathbb{R}} |f(x) - f(x-t)|\, dx < \frac{\varepsilon}{M},$$

where $\|K_{(\lambda)}\|_1 \le M$. Therefore, we have the estimate

$$\begin{aligned} \|f - f * K_{(\lambda)}\|_1 &\le 2\|f\|_1 \int_{|t|\ge\delta} |K_{(\lambda)}(t)|\, dt + \frac{\varepsilon}{M} \int_{|t|\le\delta} |K_{(\lambda)}(t)|\, dt \\ &\le 2\|f\|_1 \int_{|t|\ge\delta} |K_{(\lambda)}(t)|\, dt + \varepsilon. \end{aligned}$$

Consequently, by the definition of approximate identity, we have

$$\overline{\lim_{\lambda\to\infty}} \|f - f * K_{(\lambda)}\|_1 \le \varepsilon.$$

Since ε was arbitrary, the proof of part *a* is complete.

b. It is not difficult to calculate that the Fejér kernel satisfies

$$W_\lambda(x) = \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\xi|}{\lambda}\right) e^{2\pi i x\xi}\, d\xi,$$

see, e.g., [39]. Then, by the definition of convolution and an application of the Fubini–Tonelli theorems, we compute

$$f * W_\lambda(x) = \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\xi|}{\lambda}\right) \hat{f}(\xi) e^{2\pi i x \xi}\, d\xi.$$

Since $\{W_\lambda\}$ is an approximate identity, part b follows from part a.

c. Part c is immediate from part b.

□

Proposition B.3.6. *Let $f \in L^\infty_m(\mathbb{R})$ be continuous on $\mathbb{R}$. If $\{K_{(\lambda)} : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$ is an approximate identity, then*

$$\forall\, x \in \mathbb{R}, \quad \lim_{\lambda \to \infty} f * K_{(\lambda)}(x) = f(x).$$

If $f \in L^1_m(\mathbb{R})$ and $\hat{f} \in L^1_m(\widehat{\mathbb{R}})$, we can use Theorem B.3.5 to obtain the following pointwise inversion theorem. What we explicitly mean in its statement is that if $f \in L^1_m(\mathbb{R})$ and $\hat{f} \in L^1_m(\widehat{\mathbb{R}})$, then the formula in (B.7) is true m-*a.e.*; and that if f is continuous then (B.7) is true for all $x \in \mathbb{R}$. Compare the proof of Theorem B.3.7 with that in [39], pages 38–39.

Theorem B.3.7. Inversion formula for $L^1_m(\mathbb{R}) \cap A(\mathbb{R})$
Let $f \in L^1_m(\mathbb{R}) \cap A(\mathbb{R})$. Then,

$$\forall\, x \in \mathbb{R}, \quad f(x) = \int_{\widehat{\mathbb{R}}} \hat{f}(\xi) e^{2\pi i x \xi}\, d\xi. \tag{B.7}$$

Proof. The statement of the theorem follows from two observations. First, if $\{K_{(\lambda)} : \lambda > 0\} \subseteq L^1_m(\mathbb{R})$ is an approximate identity, then there exists a subsequence $\{\lambda_n : n = 1, \ldots\}$ such that

$$\lim_{n \to \infty} f * K_{(\lambda_n)} = f \quad m\text{-}a.e.$$

This fact is a consequence of Theorem B.3.5*a*. Second, assume that $\hat{f} \in L^1_m(\widehat{\mathbb{R}})$, that $(K_{(\lambda)})\hat{} \in L^1_m(\widehat{\mathbb{R}})$, and

$$\forall\, x \in \mathbb{R}, \quad K_{(\lambda)}(x) = \int_{\widehat{\mathbb{R}}} (K_{(\lambda)})\hat{}(\xi) e^{2\pi i x \xi}\, d\xi.$$

Then,

$$\lim_{\lambda \to \infty} \left\| \int_{\widehat{\mathbb{R}}} \hat{f}(\xi) e^{2\pi i x \xi}\, d\xi - f * K_{(\lambda)}(x) \right\|_\infty = 0.$$

□

Remark. It follows from the first observation in the proof of Theorem B.3.7 that, if $f \in L^1_m(\mathbb{R})$, then there exists a sequence $\{\lambda_n : n = 1, \ldots\} \subseteq (0, \infty)$ such that

$$\lim_{\lambda_n \to \infty} \int_{-\lambda_n/2\pi}^{\lambda_n/2\pi} \left(1 - \frac{2\pi|\xi|}{\lambda_n}\right) \hat{f}(\xi) e^{2\pi i x \xi}\, d\xi = f(x) \quad m\text{-}a.e.$$

It turns out that λ_n can be replaced by λ in the above, and that the convergence m-*a.e.* can be enlarged to include all x in the *Lebesgue set* for f, see, e.g., [195], as well as Example 4.4.7, Definition 8.4.8, and Section 8.8.4.

B.4 The $L^2_m(\mathbb{R})$ theory of Fourier transforms

We have defined the Fourier transform for functions $f \in L^1_m(\mathbb{R})$. Now our goal is to extend this transform to the space $L^2_m(\mathbb{R})$. Clearly, $L^2_m(\mathbb{R}) \not\subseteq L^1_m(\mathbb{R})$ and so we cannot use the formula (B.1), as the function under the integral sign may not be integrable.

Lemma B.4.1. *Let $f \in C_c(\mathbb{R})$. Then, $\hat{f} \in L^2_m(\widehat{\mathbb{R}})$ and*

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

Proof. Let $\tilde{f}(t) = \overline{f(-t)}$. Clearly, $\hat{f} \in A(\widehat{\mathbb{R}})$, since $C_c(\mathbb{R}) \subseteq L^1_m(\mathbb{R})$. Define $g = f * \tilde{f}$. Thus, g is continuous, $g \in L^1_m(\mathbb{R}) \cap L^\infty_m(\mathbb{R})$, and

$$g(0) = \|f\|_2^2.$$

Moreover, using the Fubini theorem and the translation invariance of Lebesgue measure, we have

$$\forall \xi \in \widehat{\mathbb{R}}, \quad \hat{g}(\xi) = |\hat{f}(\xi)|^2.$$

By Proposition B.3.6 and since g is continuous, we deduce that

$$g(0) = \lim_{\lambda \to \infty} \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\xi|}{\lambda}\right) |\hat{f}(\xi)|^2 \, d\xi.$$

Finally, the Levi–Lebesgue theorem allows us to assert that $\hat{f} \in L^2_m(\widehat{\mathbb{R}})$ and that

$$\|\hat{f}\|_2^2 = g(0) = \|f\|_2^2.$$

□

Theorem B.4.2. Plancherel theorem
There is a unique linear bijection $\mathcal{F} : L^2_m(\mathbb{R}) \to L^2_m(\widehat{\mathbb{R}})$ with the properties:
- ***a.*** $\forall\, f \in L^1_m(\mathbb{R}) \cap L^2_m(\mathbb{R})$ *and* $\forall\, \xi \in \widehat{\mathbb{R}}$, $\hat{f}(\xi) = \mathcal{F}(f)(\xi)$;
- ***b.*** $\forall\, f \in L^2_m(\mathbb{R})$, $\|f\|_2 = \|\mathcal{F}(f)\|_2$.

Proof. i. We first define the action of the operator $\mathcal{F}$ on $C_c(\mathbb{R})$ by

$$\mathcal{F}(f) = \hat{f}.$$

It follows from Lemma B.4.1 that, for $f \in C_c(\mathbb{R})$, $\mathcal{F}(f) \in L^2_m(\widehat{\mathbb{R}})$.

ii. Next, we shall prove that $\mathcal{F}(C_c(\mathbb{R})) \subseteq A(\widehat{\mathbb{R}}) \cap L^2_m(\widehat{\mathbb{R}})$ is a dense subspace of $L^2_m(\widehat{\mathbb{R}})$. Indeed, let $g \in L^2_m(\widehat{\mathbb{R}})$ and suppose that

$$\forall\, f \in C_c(\mathbb{R}), \quad \int_{\mathbb{R}} \hat{f}(\xi)\overline{g(\xi)}\, d\xi = 0. \tag{B.8}$$

If $f \in C_c(\mathbb{R})$, then the function $\tau_u(f)(x) = f(x-u)$ is also an element of $C_c(\mathbb{R})$, and so (B.8) implies that

$$\forall\, f \in C_c(\mathbb{R}) \text{ and } \forall\, u \in \mathbb{R}, \quad \int_{\mathbb{R}} \hat{f}(\xi)\overline{g(\xi)}e^{-2\pi i u\xi}\, d\xi = 0. \tag{B.9}$$

By the Hölder inequality, $\hat{f}\overline{g} \in L^1_m(\widehat{\mathbb{R}})$, and so (B.9) allows us to invoke the uniqueness theorem (Theorem B.3.5*c*) to conclude that $\hat{f}\overline{g} = 0$ *m-a.e.* for each $f \in C_c(\mathbb{R})$.

Note that

$$\forall\, f \in C_c(\mathbb{R}) \text{ and } \forall\, \xi \in \widehat{\mathbb{R}}, \quad e^{2\pi i x\xi} f(x) \in C_c(\mathbb{R}).$$

Thus, $\mathcal{F}(C_c(\mathbb{R}))$ is translation invariant, i.e.,

$$\forall\, f \in C_c(\mathbb{R}) \text{ and } \forall\, \xi \in \widehat{\mathbb{R}}, \quad \tau_u(\hat{f}) \in \mathcal{F}(C_c(\mathbb{R})).$$

From this we conclude that, for each $\xi_0 \in \widehat{\mathbb{R}}$, there is $f \in C_c(\mathbb{R})$ for which $|\hat{f}| > 0$ on some interval centered about ξ_0. To verify this claim, suppose there is ξ_0 such that for each $f \in C_c(\mathbb{R})$ and for each interval I centered at ξ_0, $\hat{f}$ has a zero in I. Consequently, $\hat{f}(\xi_0) = 0$ for each $f \in C_c(\mathbb{R})$. By the translation invariance of $\mathcal{F}(C_c(\mathbb{R}))$, $\tau_\eta(\hat{f}) \in \mathcal{F}(C_c(\mathbb{R}))$ for each $\eta \in \widehat{\mathbb{R}}$, and so

$$\forall\, f \in C_c(\mathbb{R}) \text{ and } \forall\, \eta \in \widehat{\mathbb{R}}, \quad \tau_\eta(\hat{f})(\xi_0) = 0,$$

i.e., $\hat{f} = 0$ on $\widehat{\mathbb{R}}$ for each $f \in C_c(\mathbb{R})$. This contradicts Theorem B.3.5*c*, and the claim is proved.

Therefore, if we assume (B.8) we can conclude that $g = 0$ *m-a.e.* Consequently, by the Hahn–Banach theorem (Theorem A.8.3) and by the fact that $L^2_m(\widehat{\mathbb{R}})$ is its own dual, we have that $\mathcal{F}(C_c(\mathbb{R}))$ is dense in $L^2_m(\widehat{\mathbb{R}})$.

iii. We have shown that $\mathcal{F}$ is a continuous linear injection $C_c(\mathbb{R}) \to L^2_m(\widehat{\mathbb{R}})$, when $C_c(\mathbb{R})$ is endowed with the $L^2_m(\mathbb{R})$ norm, and so $\mathcal{F}$ has a unique linear injective extension to $L^2_m(\mathbb{R})$. Also, $\mathcal{F}(C_c(\mathbb{R}))$ is closed and dense in $L^2_m(\widehat{\mathbb{R}})$ by Lemma B.4.1 and by part *ii*. Thus, $\mathcal{F}$ is also surjective.

Property *a* now follows since $C_c(\mathbb{R})$ is dense in $L^1_m(\mathbb{R})$, when equipped with $L^1_m(\mathbb{R})$ norm; and property *b* is an immediate consequence of the continuity of $\mathcal{F}$.

□

Notationally, because of Plancherel theorem, we refer to $\mathcal{F}(f)$ as the *Fourier transform of* $f \in L^2_m(\mathbb{R})$. We often write

$$\hat{f} = \mathcal{F}(f).$$

Theorem B.4.3. Parseval formula
Let $f, g \in L^2_m(\mathbb{R})$. *Then, the following formulas hold:*

$$\int_{\mathbb{R}} f(x)\overline{g(x)}\, dx = \int_{\widehat{\mathbb{R}}} \hat{f}(\xi)\overline{\hat{g}(\xi)}\, d\xi \tag{B.10}$$

and

$$\int_{\mathbb{R}} f(x)g(x)\, dx = \int_{\widehat{\mathbb{R}}} \hat{f}(\xi)\hat{g}(-\xi)\, d\xi. \tag{B.11}$$

Proof. Equation (B.10) is a consequence of Theorem B.4.2*b* and the fact that

$$4f\overline{g} = |f+g|^2 - |f-g|^2 + i|f+ig|^2 - i|f-ig|^2.$$

Equation (B.11) can be proved similarly.

□

We shall refer to Theorems B.4.2 and B.4.3 as the *Parseval–Plancherel theorem.* PARSEVAL was a French engineer who gave a formal verification of the Fourier series version of Theorem B.4.2*b* in 1799; his publication is dated 1805.

Example B.4.4. An idempotent problem in $L^1_m(\mathbb{R})$ and $L^2_m(\mathbb{R})$
Consider the equation

$$f = f * f. \tag{B.12}$$

a. If we ask whether (B.12) has a solution $f \in L^1_m(\mathbb{R}) \setminus \{0\}$, the answer is "no". Indeed, if there were such an f, then $\hat{f} = (\hat{f})^2$ so that $\hat{f}$ only takes the values 0 or 1. If $\hat{f} = 0$ on $\widehat{\mathbb{R}}$, then $f = 0$ by Theorem B.3.5*c*. If $\hat{f} = 1$ on $\widehat{\mathbb{R}}$, then $f \notin L^1_m(\mathbb{R})$ since $A(\widehat{\mathbb{R}}) \subseteq C_0(\widehat{\mathbb{R}})$. If $\hat{f}$ takes both 0 and 1 values we contradict the continuity of $\hat{f}$.

b. If we ask whether (B.12) has a solution $f \in L^2_m(\mathbb{R}) \setminus \{0\}$, the answer is "yes". In fact, let $\hat{f} = \mathbb{1}_A$ where $m(A) < \infty$. We are using here the Parseval–Plancherel theorem to assert the existence of $f \in L^2_m(\mathbb{R})$ for which $\hat{f} = \mathbb{1}_A$.

See the related discussion in [39], Example 1.10.6 and Remark 3.10.13.

Example B.4.5. The Hilbert transform
a. *Formally,* the *Hilbert transform* $\mathcal{H}(f)$ of $f : \mathbb{R} \to \mathbb{C}$ is the convolution,

$$\mathcal{H}(f)(t) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|t-x| \geq \varepsilon} \frac{f(x)}{t-x}\, dx.$$

The Hilbert transform opens the door to a large and profound area of harmonic analysis associated with the theory, relevance, and importance of *singular integrals*, e.g., [435], [184], cf., [347] for a magnificent introduction.

b. As mentioned in Example A.13.5, $\mathcal{H} \in \mathcal{L}(L^2_m(\mathbb{R}))$, and $\mathcal{H}$ is a unitary operator on $L^2_m(\mathbb{R})$. Further, $\mathcal{H} \circ \mathcal{H} = -Id$ on $L^2_m(\mathbb{R})$, and

$$\mathcal{H} = \mathcal{F}^{-1}\sigma(\mathcal{H})\mathcal{F},$$

where $\sigma(\mathcal{H})(\gamma) = -i\text{sgn}\,(\gamma)$.

c. Let $f : \mathbb{R} \to \mathbb{C}$ satisfy supp $(f) \subseteq [0, \infty)$, and define the unilateral *Laplace transform* of f as $\mathcal{L}(f)(t) = \int_0^\infty f(x)e^{-tx}\, dx$. A formal calculation, which is valid under mild hypotheses, shows that

$$\forall\, t > 0, \quad \mathcal{L}(\mathcal{L}(f))(t) = -\pi\mathcal{H}(f)(-t).$$

See [39], Problem 2.57 for a role of $\mathcal{H}$ in signal processing as related to the Paley–Wiener logarithmic integral theorem [353].

B.5 Fourier series

In Section 3.3 we defined the Fourier series of a function $f \in L^1_m(\mathbb{T})$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We now elaborate.

Definition B.5.1. Fourier series
a. Let $f \in L^1_m([-\Omega, \Omega))$. The *Fourier series* of f is the series

$$\forall\, x \in [-\Omega, \Omega), \quad S(f)(x) = \sum_{n\in\mathbb{Z}} c_n e^{-\pi i n x/\Omega}, \tag{B.13}$$

where the coefficients are defined as

$$c_n = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} f(x)e^{\pi i n x/\Omega}\, dx.$$

The c_n are the *Fourier coefficients* of f.

b. If the sequence $c = \{c_n : n \in \mathbb{Z}\}$ satisfies $\sum_{n\in\mathbb{Z}} |c_n| < \infty$, then the right hand side of (B.13) is well defined, and we say it is the *Fourier transform* $\hat{c}$ of $\{c_n : n \in \mathbb{Z}\}$.

In Definition B.5.1, instead of thinking about functions defined on a finite interval, we can think of functions f which are 2Ω-periodic, Lebesgue measurable functions which are integrable on all compact subsets of $\mathbb{R}$, i.e., locally integrable functions. The set of all locally integrable functions is denoted by $L^1_{\text{loc}}(\mathbb{R})$, e.g., Section 5.5. If $\Omega > 0$ and $f \in L^1_{loc}(\mathbb{R})$ is 2Ω-periodic, then we write $f \in L^1(\mathbb{T}_{2\Omega})$. $\mathbb{T}_{2\Omega}$ is identified with $\mathbb{R}/2\Omega\mathbb{Z}$. The L^1-*norm* of $f \in L^1(\mathbb{T}_{2\Omega})$ is

$$\|f\|_{L^1(\mathbb{T}_{2\Omega})} = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} |f(x)|\, dx.$$

If f is a 2Ω-periodic, Lebesgue measurable function, and $f^2 \in L^1(\mathbb{T}_{2\Omega})$, then we write $f \in L^2(\mathbb{T}_{2\Omega})$. The L^2-*norm* of $f \in L^2(\mathbb{T}_{2\Omega})$ is

$$\|f\|_{L^2(\mathbb{T}_{2\Omega})} = \left(\frac{1}{2\Omega}\int_{-\Omega}^{\Omega} |f(x)|^2\, dx\right)^{1/2}.$$

By the Hölder inequality, $L^2(\mathbb{T}_{2\Omega}) \subseteq L^1(\mathbb{T}_{2\Omega})$, and we have

$$\forall\, f \in L^2(\mathbb{T}_{2\Omega}), \quad \|f\|_{L^1(\mathbb{T}_{2\Omega})} \le \|f\|_{L^2(\mathbb{T}_{2\Omega})}.$$

Definition B.5.2. Fourier transform on $L^1(\mathbb{T}_{2\Omega})$
If $f \in L^1(\mathbb{T}_{2\Omega})$, its *Fourier transform* is the sequence $\hat{f} = \{\hat{f}(n) : n \in \mathbb{Z}\}$, where

$$\forall\, n \in \mathbb{Z}, \quad \hat{f}(n) = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} f(x) e^{-\pi i n x/\Omega}\, dx.$$

In Appendix B.1, we defined the Fourier transform of $f \in L^1_m(\mathbb{R})$ to be a function on $\widehat{\mathbb{R}} = \mathbb{R}$. Here we have two dual settings. First, for a sequence $c = \{c_n : n \in \mathbb{Z}\}$ such that $\sum |c_n| < \infty$, the Fourier transform of c is defined on $\mathbb{T}_{2\Omega}$. Second, for $f \in L^1(\mathbb{T}_{2\Omega})$, the Fourier transform of f is defined on $\mathbb{Z}$. Mathematically, $\mathbb{R}$ and $\widehat{\mathbb{R}}$ are *locally compact Abelian groups* (LCAGs) that are *dual*, in a technical sense, to each other, see Appendix B.9 and Appendix B.10. Similarly, the discrete LCAG $\mathbb{Z}$ is the dual group of the compact LCAG $\mathbb{T}_{2\Omega}$, and vice versa, e.g., [391], [151], [33].

We shall use the Riemann–Lebesgue lemma, Theorem 3.6.4, to verify DIRICHLET's fundamental theorem, which provides sufficient conditions on a function $f \in L^1(\mathbb{T}_{2\Omega})$ so that $S(f)(x_0) = f(x_0)$ for a given point x_0. The following ingenious proof is due to PAUL CHERNOFF [100], cf., [313] and the classical proof as found in [509]. The Dirichlet theorem for Fourier series naturally preceded the analogous inversion theorem for Fourier transforms, as formulated in Theorem B.1.2 and Section B.3.

Theorem B.5.3. Dirichlet theorem
If $f \in L^1(\mathbb{T}_{2\Omega})$ and f is differentiable at x_0, then $S(f)(x_0) = f(x_0)$ in the sense that

$$\lim_{M,N\to\infty} \sum_{n=-M}^{N} c_n e^{-\pi i n x_0/\Omega} = f(x_0),$$

where $c = \{c_n : n \in \mathbb{Z}\}$ is the sequence of Fourier coefficients of f.

Proof. i. Without loss of generality, assume $x_0 = 0$ and $f(x_0) = 0$. In fact, if $f(x_0) \neq 0$, then consider the function $f - f(x_0)$ instead of f, which is also an element of $L^1(\mathbb{T}_{2\Omega})$, and then translate this function to the origin.

ii. Since $f(0) = 0$ and $f'(0)$ exists, we can verify that

$$g(x) = \frac{f(x)}{e^{-\pi i x/\Omega} - 1}$$

is bounded in some interval centered at the origin. To see this note that

$$g(x) = \frac{f(x)}{x} \frac{1}{\sum_{j=1}^{\infty} (-\pi i/\Omega)^j (1/j!) x^j},$$

and, hence, $g(x)$ is close to $-\Omega f'(0)/(\pi i)$ in a neighborhood of the origin.

The boundedness near the origin, coupled with integrability of f on $\mathbb{T}_{2\Omega}$, yields the integrability of g on $\mathbb{T}_{2\Omega}$. Therefore, since $f(x) = g(x)(e^{-\pi i x/\Omega} - 1)$, we compute $c_n = d_{n+1} - d_n$, where $d = \{d_n : n \in \mathbb{Z}\}$ is the sequence of Fourier coefficients of g. Thus, the partial sum $\sum_{n=-M}^{N} c_n e^{-\pi i n x_0/\Omega}$ is the telescoping series

$$\sum_{n=-M}^{N} (d_{n+1} - d_n) = d_{N+1} - d_{-M}.$$

Consequently, we can apply the Riemann–Lebesgue lemma to to the sequence of Fourier coefficients of g to obtain

$$\lim_{M,N \to \infty} \sum_{n=-M}^{N} c_n e^{-\pi i n x/\Omega} = 0.$$

□

With regard to Theorem B.5.3, we can further assert that *if* $f \in BV_{loc}(\mathbb{R})$, f *is* 2Ω*-periodic, and* f *is continuous on a closed subinterval* $I \subseteq \mathbb{T}_{2\Omega}$*, then*

$$\sum_{n=-N}^{N} \hat{f}(n) e^{-\pi i n x/\Omega}$$

converges uniformly to f *on* I, cf., [509], Volume I, pages 57–58. The *Dirichlet theorem* and this version of it for intervals of continuity are often referred to as the *Dirichlet–Jordan test.*

B.6 The $L^1(\mathbb{T}_{2\Omega})$ theory of Fourier series

Definition B.6.1. $A(\mathbb{T}_{2\Omega})$ **and** $A(\mathbb{Z})$

a. If $c = \{c_n : n \in \mathbb{Z}\}$ is a sequence such that $\sum |c_n| < \infty$ and if $\Omega > 0$, then $\hat{c}$ is an *absolutely convergent Fourier series*, and the space of such series is denoted by $A(\mathbb{T}_{2\Omega})$. By definition, the norm of $\hat{c} \in A(\mathbb{T}_{2\Omega})$ is

$$\|\hat{c}\|_{A(\mathbb{T}_{2\Omega})} = \|c\|_1 = \sum_{n \in \mathbb{Z}} |c_n|.$$

We have the proper inclusions

$$A(\mathbb{T}_{2\Omega}) \subseteq C(\mathbb{T}_{2\Omega}) \subseteq L^\infty(\mathbb{T}_{2\Omega}) \subseteq L^2(\mathbb{T}_{2\Omega}) \subseteq L^1(\mathbb{T}_{2\Omega}).$$

b. Let $A(\mathbb{Z})$ be the space of all sequences $c = \{c_n : n \in \mathbb{Z}\}$ such that $\hat{c} \in L^1(\mathbb{T}_{2\Omega})$. The space $\ell^2(\mathbb{Z})$ of square-summable sequences (defined in Section 5.5) is a subset of $A(\mathbb{Z})$.

Example B.6.2. Trigonometric series

a. The Riemann–Lebesgue lemma asserts that if $f \in L^1(\mathbb{T}_{2\Omega})$, then

$$\lim_{n \to \pm\infty} c_n = 0,$$

where $c = \{c_n : n \in \mathbb{Z}\}$ is the sequence of Fourier coefficients of f. On the other hand, suppose we are give a trigonometric series $\sum c_n e^{-2\pi i n x}$ for which $\lim_{n \to \pm\infty} c_n = 0$. Then, it is not necessarily true that this series is the Fourier series of some function $f \in L^1(\mathbb{T}_{2\Omega})$. Indeed, the trigonometric series,

$$\sum_{n=2}^{\infty} \frac{\sin(\pi n x/\Omega)}{\log(n)},$$

converges pointwise for each $x \in \mathbb{R}$, but it is *not* the Fourier series of an element $f \in L^1(\mathbb{T}_{2\Omega})$. This is an analogue of Example B.2.5.

b. Let $\Omega > 0$. Then, the series,

$$\sum_{n=3}^{\infty} \frac{\sin(\pi n x/\Omega)}{n \log(n)},$$

converges uniformly on $\mathbb{R}$ to a function $f \in C(\mathbb{T}_{2\Omega}) \setminus A(\mathbb{T}_{2\Omega})$.

Remark. If $f \in L^1(\mathbb{T}_{2\Omega})$ and if we let

$$S_N(f)(x) = \sum_{n=-N}^{N} c_n e^{-\pi i n x/\Omega},$$

where $c = \{c_n : n \in \mathbb{Z}\}$ is the sequence of Fourier coefficients of f, then it is known that the desirable statement,

$$\lim_{N \to \infty} \|S_N(f) - f\|_{L^1(\mathbb{T}_{2\Omega})} = 0, \tag{B.14}$$

is *not* true for all $f \in L^1(\mathbb{T}_{2\Omega})$. On the other hand, a sequence $\{f_n : n = 1, \ldots\} \subseteq L^1(\mathbb{T}_{2\Omega})$ converges to $f \in L^1(\mathbb{T}_{2\Omega})$ *weakly*, i.e.,

$$\forall\, g \in L^\infty(\mathbb{T}_{2\Omega}), \quad \lim_{n \to \infty} \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} (f_n(x) - f(x)) g(x)\, dx = 0,$$

cf., Definition 6.3.1, if and only if

$$\lim_{n \to \infty} \int_A (f_n(x) - f(x))\, dx = 0 \tag{B.15}$$

for every Lebesgue measurable set $A \subseteq \mathbb{T}_{2\Omega}$. If we have weak convergence, or, equivalently, (B.15), then it follows from Theorem 6.5.1 that (B.14) is true for $S_n(f) = f_n$ if $\{f_n : n \in \mathbb{N}\}$ converges to f in measure.

Definition B.6.3. Convolution

a. Let $f, g \in L^1(\mathbb{T}_{2\Omega})$. The *convolution* of f and g, denoted by $f * g$, is

$$f * g(x) = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} f(x-y)g(y)\,dy = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} f(y)g(x-y)\,dy,$$

cf., Problem 3.28. As with $L^1_m(\mathbb{R})$, see Problem 3.5, it is not difficult to prove that $f * g \in L^1(\mathbb{T}_{2\Omega})$ and

$$\forall\, f, g \in L^1(\mathbb{T}_{2\Omega}), \quad \|f * g\|_{L^1(\mathbb{T}_{2\Omega})} \leq \|f\|_{L^1(\mathbb{T}_{2\Omega})}\|g\|_{L^1(\mathbb{T}_{2\Omega})}.$$

b. $L^1(\mathbb{T}_{2\Omega})$ is a *commutative algebra* taken with the operations of addition and convolution, i.e., it is a complex vector space under addition, and convolution is distributive with respect to addition, as well as associative and commutative. $L^1(\mathbb{T}_{2\Omega})$ is a commutative Banach algebra when normed by $\|...\|_{L^1(\mathbb{T}_{2\Omega})}$.

Proposition B.6.4. *Let $f, g \in L^1(\mathbb{T}_{2\Omega})$, with corresponding sequences $c = \{c_n : n \in \mathbb{Z}\}$, $d = \{d_n : n \in \mathbb{Z}\} \in A(\mathbb{Z})$ of Fourier coefficients. Then, the sequence $c \cdot d = \{c_n d_n : n \in \mathbb{Z}\} \in A(\mathbb{Z})$ is the sequence of Fourier coefficients of $f * g \in L^1(\mathbb{T}_{2\Omega})$, i.e.,*

$$\forall\, n \in \mathbb{Z}, \quad c_n d_n = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} f * g(x) e^{\pi i n x/\Omega}\,dx.$$

Definition B.6.5. Approximate identity

An *approximate identity* is a family $\{K_N : N = 1, \ldots\} \subseteq L^1(\mathbb{T}_{2\Omega})$ of functions with the properties:

i. $\forall\, N = 1, \ldots, \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} K_N(x)\,dx = 1$,

ii. $\exists\, M > 0$ such that $\forall\, N = 1, \ldots, \|K_N\|_{L^1(\mathbb{T}_{2\Omega})} \leq M$,

iii. $\forall\, \delta \in (0, \Omega]$, $\lim_{N\to\infty} \frac{1}{2\Omega}\int_{\delta \leq |x| \leq \Omega} |K_N(x)|\,dx = 0$.

The following theorem is the analogue for Fourier series of Theorem B.3.5 and Proposition B.3.6, see Problem 3.28.

Theorem B.6.6. Approximation and uniqueness

a. *Let $f \in C(\mathbb{T}_{2\Omega})$, and let $\{K_N : N = 1, \ldots\} \subseteq L^1(\mathbb{T}_{2\Omega})$ be an approximate identity. Then,*

$$\lim_{N\to\infty} \|f - f * K_N\|_{L^\infty(\mathbb{T}_{2\Omega})} = 0.$$

b. *Let $f \in L^1(\mathbb{T}_{2\Omega})$, and let $\{K_N : N = 1, \ldots\} \subseteq L^1(\mathbb{T}_{2\Omega})$ be an approximate identity. Then,*

$$\lim_{N\to\infty} \|f - f * K_N\|_{L^1(\mathbb{T}_{2\Omega})} = 0,$$

cf., Theorem B.3.5a.

c. *Let $f \in L^1(\mathbb{T}_{2\Omega})$, and let $c = \{c_n : n \in \mathbb{Z}\}$ be its sequence of Fourier coefficients. If $c_n = 0$ for each $n \in \mathbb{Z}$, then $f = 0$ m-a.e.*

B.7 The Stone–Weierstrass theorem

In 1904, FEJÉR proved the following fundamental approximation theorem using the fact that the Fejér kernel $\{W_N : N \geq 0\}$, defined in Problem 3.28, is an approximate identity, without explicitly defining the general concept of an approximate identity.

Theorem B.7.1. Fejér theorem
a. *Let* $f \in C(\mathbb{T}_{2\Omega})$*. Then,*

$$\lim_{N\to\infty} \|f - f * W_N\|_{L^\infty(\mathbb{T}_{2\Omega})} = 0. \tag{B.16}$$

b. *Let* $f \in L^p_m(\mathbb{T}_{2\Omega})$*,* $1 \leq p < \infty$*. Then,*

$$\lim_{N\to\infty} \|f - f * W_N\|_{L^p(\mathbb{T}_{2\Omega})} = 0.$$

Theorem B.7.1 and the Weierstrass approximation theorem are conceptually closely related. We note that WEIERSTRASS' original proof (1885) also used a convolution approximate identity argument [80], pages 269–273.

Theorem B.7.2. Weierstrass approximation theorem
Let $f : [\alpha, \beta] \to \mathbb{C}$ *be a continuous function. There is a sequence* $\{Q_N : N = 1, \ldots\}$ *of polynomials for which*

$$\lim_{N\to\infty} \|f - Q_N\|_{L^\infty_m([\alpha,\beta])} = 0. \tag{B.17}$$

Equation (B.17) can be derived from (B.16) in the following way. By translation we can take $f \in C([-\Omega, \Omega])$. Next choose c such that $g(-\Omega) = g(\Omega)$ where $g(x) = f(x) - c$ for $x \in [-\Omega, \Omega]$. In fact, let

$$c = \frac{f(\Omega) - f(-\Omega)}{2\Omega}.$$

Apply (B.16) to g considered as an element of $C(\mathbb{T}_{2\Omega})$. Finally, uniformly approximate the trigonometric polynomials $g * W_N$ on $[-\Omega, \Omega]$ by polynomial approximants of their Taylor series expansions.

A monumental journey in effective abstraction was undertaken by MARSHALL H. STONE in 1937 [440], and has resulted in the Stone–Weierstrass theorem.

STONE's own works [441], [442] are a readable paradigm of the creative inquiry required to formulate fundamental abstract ideas resulting from and embedded in classical results.

In order to state a useful version of STONE's theorem, let X be a locally compact Hausdorff space, and note that the complex Banach space $C_0(X)$ (defined in Section A.2) is also in *algebra.* This means that $C_0(X)$ is not only a vector space, but that it is closed under pointwise multiplication, and that the commutative, associative, and distributive laws hold. A subset $S \subseteq C_0(X)$ is *separating* if

$$\forall\, x_1, x_2 \in X,\ \exists\, f \in S \ \text{ such that } \quad f(x_1) \neq f(x_2).$$

Theorem B.7.3. Stone–Weierstrass theorem
Let X be a locally compact Hausdorff space, and let $S \subseteq C_0(X)$ be a separating subalgebra with the following properties:
i. $\forall\, f \in S,\ \overline{f} \in S$,
ii. $\forall\, x \in X,\ \exists f \in S$ *such that* $f(x) \neq 0$.
Then, S is dense in $C_0(X)$.

If X is compact we refer to [319], pages 11–12, for a brief proof based on [441] by STONE, and to [229], pages 90–99, for a more complete treatment.

B.8 The $L^2(\mathbb{T}_{2\Omega})$ theory of Fourier series

Recall that according to Definition A.12.1 an *orthonormal basis* for $L^2(\mathbb{T}_{2\Omega})$ is an orthonormal sequence $\{e_n : n \in \mathbb{Z}\} \subseteq L^2(\mathbb{T}_{2\Omega})$ such that

$$\forall\, f \in L^2(\mathbb{T}_{2\Omega}),\ \exists\, \{c_n : n \in \mathbb{Z}\} \subseteq \mathbb{C} \text{ such that } \quad f = \sum_{n\in\mathbb{Z}} c_n e_n \ \text{ in } \ L^2(\mathbb{T}_{2\Omega}).$$

In fact, $L^2(\mathbb{T}_{2\Omega})$ is a Hilbert space with inner product defined in Example A.12.3*a*, where the following result was also stated.

Proposition B.8.1. *The sequence $\{e_n(x) = e^{-\pi i n x/\Omega} : n \in \mathbb{Z}\} \subseteq L^2(\mathbb{T}_{2\Omega})$ is an orthonormal basis for $L^2(\mathbb{T}_{2\Omega})$.*

Theorem B.8.2 is a special case of Theorem A.12.6, once it is proved that $L^2(\mathbb{T}_{2\Omega})$ is a Hilbert space. The following proof is self-contained.

Theorem B.8.2. Parseval formula
Let $f, g \in L^2(\mathbb{T}_{2\Omega})$ with corresponding sequences $c = \{c_n : n \in \mathbb{Z}\}$, $d = \{d_n : n \in \mathbb{Z}\}$ of Fourier coefficients of f and g. Then, $c, d \in \ell^2(\mathbb{Z})$ and

$$\frac{1}{2\Omega}\int_{\Omega}^{\Omega} f(x)\overline{g(x)}\, dx = \sum_{n\in\mathbb{Z}} c_n \overline{d_n};$$

and, in particular,

$$\frac{1}{2\Omega}\int_{\Omega}^{\Omega} |f(x)|^2\, dx = \sum_{n\in\mathbb{Z}} |c_n|^2.$$

Proof. We first observe that for any $f \in L^2(\mathbb{T}_{2\Omega})$ and for any $N \in \mathbb{N}$,

$$0 \leq \|f - S_N(f)\|^2_{L^2(\mathbb{T}_{2\Omega})} = \|f\|^2_{L^2(\mathbb{T}_{2\Omega})} - \sum_{|n|\leq N} |c_n|^2,$$

which implies

$$\sum_{|n|\leq N} |c_n|^2 \leq \|f\|^2_{L^2(\mathbb{T}_{2\Omega})}. \tag{B.18}$$

Further, if $N > M$, then

$$\|S_N(f) - S_M(f)\|^2_{L^2(\mathbb{T}_{2\Omega})} = \sum_{M<|n|\leq N} |c_n|^2,$$

and so, by (B.18), $\{S_N(f) : N \in \mathbb{N}\}$ is a Cauchy sequence in $L^2(\mathbb{T}_{2\Omega})$. Thus, $\sum c_n e^{-\pi inx/\Omega}$ converges to some $h \in L^2(\mathbb{T}_{2\Omega})$ since $L^2(\mathbb{T}_{2\Omega})$ is complete. Now, for any $f \in L^2(\mathbb{T}_{2\Omega})$ and corresponding h we have, by Proposition B.8.1, that

$$\frac{1}{2\Omega}\int_{-\Omega}^{\Omega}(f(x) - h(x))e^{\pi inx/\Omega}\,dx$$
$$= c_n - \lim_{N\to\infty}\sum_{|m|\leq N} c_m \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} e^{-\pi i(m-n)x/\Omega}\,dx = 0.$$

Therefore, $f = h$ *m-a.e.* in $\mathbb{T}_{2\Omega}$, or, in other words,

$$\lim_{N\to\infty}\|f - S_N(f)\|_{L^2(\mathbb{T}_{2\Omega})} = 0.$$

Using this fact, we obtain

$$\frac{1}{2\Omega}\int_{\Omega}^{\Omega} f(x)\overline{g(x)}\,dx = \lim_{N\to\infty}\frac{1}{2\Omega}\int_{\Omega}^{\Omega} S_N(f)(x)\overline{S_N(g)(x)}\,dx$$
$$= \lim_{N\to\infty}\sum_{|m|,|n|\leq N} c_m\overline{d_n}\frac{1}{2\Omega}\int_{\Omega}^{\Omega} e^{-\pi i(m-n)x/\Omega}\,dx$$
$$= \sum_{n\in\mathbb{Z}} c_n\overline{d_n},$$

where the last equality again follows from Proposition B.8.1.

□

The inequality (B.18) is called the *Bessel inequality*, cf., Theorem A.12.5*a*. It implies that if $f \in L^2(\mathbb{T}_{2\Omega})$, then the sequence $c = \{c_n : n \in \mathbb{Z}\}$ of Fourier coefficients of f is square summable, i.e.,

$$\sum_{n\in\mathbb{Z}} |c_n|^2 < \infty.$$

F. RIESZ' formulation of the *Riesz–Fischer theorem* completes the picture as follows, see Section 5.6.2.

Theorem B.8.3. $l^2(\mathbb{Z}) \longrightarrow L^2(\mathbb{T}_{2\Omega})$
There is a unique linear bijection $\mathcal{F} : l^2(\mathbb{Z}) \longrightarrow L^2(\mathbb{T}_{2\Omega})$ *with the properties:*
a. $\forall\, c \in l^2(\mathbb{Z})$, $\|c\|_{l^2(\mathbb{Z})} = \|\mathcal{F}(c)\|_{L^2(\mathbb{T}_{2\Omega})}$,
b. $\forall\, f \in L^2(\mathbb{T}_{2\Omega})$, $\mathcal{F}^{-1}(f)$ *is the sequence of Fourier coefficients of* f.

B.9 Haar measure

An additive group G with a locally compact Hausdorff topology is a *locally compact group* if the function $G \times G \to G$, $(x,y) \mapsto x - y$, is continuous. For significant, classical treatises on locally compact groups, that were begun in the 1930s, see [363] and [482] by PONTRYAGIN and ANDRÉ WEIL, respectively.

A complex vector space X which is also a topological space is a *topological vector space* if the functions $\mathbb{C} \times X \to X$, $(c,x) \mapsto cx$, and $X \times X \to X$, $(x,y) \mapsto x+y$, are continuous. Let X be a Hausdorff topological vector space. $K \subseteq X$ is *absorbing* if

$$\forall\, x \in X,\ \exists\, \varepsilon > 0 \quad \text{such that} \quad 0 < |c| \leq \varepsilon \implies cx \in K.$$

$K \subseteq X$ is *balanced* if

$$|c| \leq 1 \implies cK \subseteq K.$$

In both cases, $c \in \mathbb{C}$.

A fundamental result in harmonic analysis is that *if G is a locally compact group then there is a Borel measure m_G on G such that*

$$\forall\, B \in \mathcal{B}(G) \text{ and } \forall\, x \in G, \quad m_G(B) = m_G(B+x),$$

where $B + x = \{y + x : y \in B\}$. In this case m_G is a *right Haar measure* on G; and, when $B + x$ is replaced by $x + B$, m_G is a *left Haar measure* on G. Thus, the crucial feature of translation invariance for Lebesgue measure on the line extends to locally compact groups. If every right Haar measure is a left Haar measure on a locally compact group G, and vice-versa, then G is *unimodular.* Every compact group and every locally compact Abelian group (LCAG) is unimodular.

We shall prove the existence of a Haar measure on G *compact* and *Abelian* using the Markov–Kakutani fixed point theorem, which we shall also prove. It is easy to show that there is only one such m_G which ensures that $m_G(G) = 1$ for G compact, see Theorem B.9.3.

Theorem B.9.1. Markov–Kakutani fixed point theorem
Let X be a Hausdorff topological vector space, take a compact and convex set $K \subseteq X$, and let $\{T_\alpha\}$ be a family of continuous linear maps $T_\alpha : X \to X$ which satisfies

$$\forall\, \alpha, \quad T_\alpha(K) \subseteq K$$

and

$$\forall\, \alpha, \beta \quad T_\alpha \circ T_\beta = T_\beta \circ T_\alpha. \tag{B.19}$$

Then, there is $k \in K$ such that

$$\forall\, \alpha, \quad T_\alpha(k) = k.$$

Proof. Let

$$T_\alpha^{(n)} = \frac{Id + T_\alpha + \ldots + T_\alpha^{n-1}}{n}$$

where $Id : X \to X$ is the identity map and T_α^j is the composition $T_\alpha \circ \ldots \circ T_\alpha$, j times. Clearly, each $T_\alpha^{(n)} : X \to X$ is continuous and linear. We write $T = \{T_\alpha^n\}$ and let $\tilde{T}$ be the set of finite products, under composition, of elements from T. Note that for each $u \in \tilde{T}$, $u(K) \subseteq K$. This follows by the convexity of K. Thus, $u \circ v(K) \subseteq u(K)$ for $u, v \in \tilde{T}$. Further, because of (B.19), $u \circ v = v \circ u$ for $u, v \in \tilde{T}$; and hence $v \circ u(K) \subseteq v(K)$ implies $u \circ v(K) \subseteq v(K)$.

Let $\tilde{K} = \bigcap\{u(K) : u \in \tilde{T}\}$ and note that

$$\forall\, u, v \in \tilde{T}, \quad u(K) \cap v(K) \neq \emptyset;$$

in fact,

$$u \circ v(K) = u \circ v(K) \cap u \circ v(K) \subseteq u(K) \cap u \circ v(K) \subseteq u(K) \cap v(K).$$

Since K is compact, $u(K)$ is compact for each $u \in \tilde{T}$. Consequently, because $u(K) \cap v(K) \neq \emptyset$ for $u, v \in \tilde{T}$, since each $u(K) \subseteq K$, and because K is compact, there is $k \in K$ such that for all $u \in \tilde{T}$, $k \in u(K)$. Therefore, $k \in \tilde{K}$.

In particular, for each α and n, $k \in T_\alpha^{(n)}(K)$, so that there is $t \in K$ (depending on α and n) such that $T_\alpha^{(n)}(t) = k$. Thus,

$$T_\alpha(k) = \frac{T_\alpha(t) + \ldots + T_\alpha^n(t)}{n},$$

and, hence,

$$T_\alpha(k) - k = -\frac{t}{n} + \frac{T_\alpha^n(t)}{n} \tag{B.20}$$

since $T_\alpha^{(n)}(t) = k$.

Note that the function $K \times K \to X$, $(s, t) \mapsto s - t$, takes $K \times K$ into a compact set E.

Because of (B.20) and the fact that $T_\alpha^n(t) \in K$ we have

$$T_\alpha(k) - k = \frac{T_\alpha^n(t) - t}{n} \in \frac{1}{n}E. \tag{B.21}$$

We shall show

$$\bigcap_{n=1}^{\infty} \left(\frac{1}{n}E\right) = \{0\}, \tag{B.22}$$

so that since (B.21) is true for each n, $T_\alpha(k) - k = 0$, and we are done.

Let $V \subseteq U$ be a balanced set for which $0 \in \operatorname{int} V$. By the definition of a topological vector space there is a balanced and absorbing open set $W \subseteq X$ containing 0, such that $W + W \subseteq V$. $\{x + W : x \in X\}$ is an open cover of E so that by the compactness there are points x_i, $i = 1, \ldots, m$, such that

$$E \subseteq \bigcup_{i=1}^{m} (x_i + W).$$

Since W is absorbing, we have $rx_i \in W$ for $i = 1, \ldots, m$ and for some $r \in (0, 1]$. Therefore,

$$r(E \cap (x_i + W)) \subseteq W + rW, \quad i = 1, \ldots, m. \tag{B.23}$$

Because the right-hand side of (B.23) is independent of i,

$$rE = \bigcup_{i=1}^{m} r(E \cap (x_i + W)) \subseteq W + rW.$$

Thus, $E \subseteq (1/r)V$, noting that $W + rW \subseteq W + W \subseteq V$; and so, if $1/n < r$, then

$$\frac{1}{n}E \subseteq \frac{1}{nr}V \subseteq \frac{r}{r}V = V.$$

Consequently, $\bigcap_{n=1}^{\infty}(1/n)E \subseteq V$; so that since X is Hausdorff and V is arbitrary we have (B.22).

□

Theorem B.9.2. Existence of Haar measure
Let G be a compact Abelian group. Then, there is a Haar measure m_G on G.

Proof. Let $M_1(G) = \{\mu \in M_b(G) : \|\mu\|_1 \leq 1\}$. By the Banach–Alaoglu theorem, $M_1(G)$ is weak * compact in $M_b(G)$. Let $M_1^+(G) = \{\mu \in M_1(G) : \mu(1) = 1\}$.

Note that μ is positive if $\mu \in M_1^+(G)$; to prove this we assume the opposite and obtain a contradiction using the fact that $\|\mu\|_1 = \mu(1)$, e.g., [69], page 101.

If $M_b(G)$ is taken with the weak * topology, then the map $M_b(G) \to \mathbb{C}$, $\mu \mapsto \mu(1)$, is continuous. Hence, $\{\mu \in M_b(G) : \mu(1) = 1\}$ is weak * closed. Thus, $M_1^+(G)$ is weak * compact. It is easy to check that $M_1^+(G)$ is convex. For $x \in G$ and $\mu \in M_b(G)$ we define the translation $\tau_x(\mu)$ as

$$\tau_x(\mu)(f) = \int f(y - x)\, d\mu(y),$$

where $f \in C(G)$. Then, for each $x \in G$ we define the map $T_x : M_b(G) \to M_b(G)$, $\mu \mapsto \tau_x(\mu)$. Note that T_x is continuous with the weak * topology on both domain and range, linear, and

$$\forall\, x, y \in G, \quad T_x \circ T_y = T_{x+y} = T_y \circ T_x,$$

since G is Abelian.

It is also elementary to check that, for each $x \in G$,

$$T_x(M_1^+(G)) \subseteq M_1^+(G).$$

Therefore, by Theorem B.9.1, there is $m_G \in M_1^+(G)$ such that $\tau_x(m_G) = m_G$, for all $x \in G$, the required translation invariance. Further, $\|m_G\|_1 = 1$, $m_G(1) = 1$, and m_G is positive.

□

Remark. The question of existence of Haar measures goes back to SOPHUS LIE. The Hungarian mathematician ALFRÉD HAAR (1933) [203] proved the existence of translation invariant measures on separable compact groups. As a matter of fact, HAAR credits ADOLF HURWITZ for a remark in [242], which is essential for proving the existence of a Haar measure on a Lie group, see [203]. Existence of a Haar measure on a general locally compact group was first proved by ANDRÉ WEIL [482] and, later the same year, by HENRI CARTAN [93].

Besides the existence, it is natural to ask about the uniqueness of Haar measure on locally compact groups. This question was first answered by VON NEUMANN for compact groups [476]. VON NEUMANN later extended his own result to second countable locally compact groups [477] (employing a different technique). Here we prove the uniqueness of Haar measure in the simple context of a LCAG. We follow the proof of [391], see also [70], [318], [344], [403], and [482]. For a short proof in the non-Abelian case, which uses a notion of an approximate identity, we refer the reader to [255].

Theorem B.9.3. Uniqueness of Haar measure
Let G be a LCAG. Let m_G^1 and m_G^2 be two Haar measures on G. Then, there exists $C > 0$ such that $m_G^1 = C m_G^2$.

Proof. Let $g_1 \in C_c^+(G)$ be chosen so that $\int_G g_1 \, dm_G^1 = 1$, and let $C = \int_G g_1(-x)\, dm_G^2(x)$. Then, for all $g_2 \in C_c^+(G)$, we have

$$\begin{aligned}
\int_G g_2 \, dm_G^2 &= \int_G g_1(x_1)\, dm_G^1(x_1) \int_G g_2(x_2)\, dm_G^2(x_2) \\
&= \int_G \left(\int_G g_2(x_2)\, dm_G^2(x_2) \right) g_1(x_1)\, dm_G^1(x_1) \\
&= \int_G \left(\int_G g_2(x_1 + x_2)\, dm_G^2(x_2) \right) g_1(x_1)\, dm_G^1(x_1) \\
&= \int_G \int_G g_1(x_1) g_2(x_1 + x_2)\, dm_G^2(x_2)\, dm_G^1(x_1) \\
&= \int_G \int_G g_1(x_1) g_2(x_1 + x_2)\, dm_G^1(x_1)\, dm_G^2(x_2) \\
&= \int_G \int_G g_1(y_1 - y_2) g_2(y_1)\, dm_G^1(y_1)\, dm_G^2(y_2) \\
&= \int_G \int_G g_1(y_1 - y_2) g_2(y_1)\, dm_G^2(y_2)\, dm_G^1(y_1)
\end{aligned}$$

$$= \int_G \left(\int_G g_1(y_1 - y_2)\, dm_G^2(y_2) \right) g_2(y_1)\, dm_G^1(y_1)$$
$$= \left(\int_G g_1(-y_2)\, dm_G^2(y_2) \right) \int_G g_2(y_1)\, dm_G^1(y_1) = C \int_G g_2\, dm_G^1.$$

□

Example B.9.4. Examples of Haar measures
a. Let $G = \mathbb{R}^d$ be considered as an additive group. With topology defined by the Euclidean norm, G is a LCAG. Then, the Lebesgue measure m^d is a Haar measure on G.

b. Let $G = \mathbb{R} \setminus \{0\}$ be the multiplicative group of real numbers taken with the topology induced from the Euclidean norm topology from $\mathbb{R}$. Then, G is a LCAG. Further, $m_G(A) = \int_A (1/|x|)\, dm(x)$ defines a Haar measure on G.

c. Let $G = \mathbb{C}$ be the additive group of complex numbers with the usual topology of the complex plane. Then, G is a LCAG, and $m_G(A) = \int_A dxdy = \int_A dm^2(z)$, $z = x + iy$, is a Haar measure. This is the product measure on $\mathbb{R} \times \mathbb{R}$ which is isomorphic to $\mathbb{C}$.

d. Let $G = \mathbb{C} \setminus \{0\}$, $0 \in \mathbb{C}$, be the multiplicative group of complex numbers with the induced topology from $\mathbb{C}$. G is a LCAG, and $m_G(A) = \int_A (1/|z|^2)\, dm^2(z)$ defines a Haar measure on G.

$\mathbb{T} = \mathbb{R}/\mathbb{Z}$, identified with $\{z \in \mathbb{C} : |z| = 1\}$, is a compact subgroup of G. Note that $m_G(\mathbb{T}) = 0$. However, there is a natural locally compact topology on $\mathbb{T}$ and corresponding Haar measure $m_{\mathbb{T}}$ when $\mathbb{T}$ is identified with the additive group $[0, 1)$ with addition defined mod 1. In this case, $\mathbb{T}$ and $m_{\mathbb{T}}$ can be identified with $([0, 1), \mathcal{M}, m)$. We mention this because of the discussion of the *dual group* in Appendix B.10.

e. Let $G = (\mathbb{R}^+ \setminus \{0\}) \times \mathbb{R}$ be a group with group action defined by $(a, b)\cdot(a', b') = (aa', ab'+b)$. Note that, in general, $(aa', ab'+b) \neq (a'a, a'b+b')$. G is a non-Abelian locally compact group, and it is called the *affine group*. In this case, a left Haar measure is defined by $m_G^L(A) = \int_A (1/a^2)\, dm(a)\, dm(b)$; and a right Haar measure on G is defined by $m_G^R(A) = \int_A (1/a)\, dm(a)\, dm(b)$. They are distinct.

f. Let G be a discrete group. Then, $m_G = \sum_{g \in G} \delta_g$.

g. The general linear group $G = GL(d, \mathbb{R})$ with matrix multiplication is non-Abelian locally compact group, whose topology is induced from the product topology on the d^2-direct product $\mathbb{R} \times \ldots \times \mathbb{R}$. G played a role in Section 8.7. In this case, $m_G(A) = \int_A |\det(X)|^{-d} dm^{d^2}(X)$ defines both a left and a right Haar measure on G.

It is well known that *every locally compact topological group is complete with its right uniform structure.* We shall verify this general statement for the metrizable setting.

Theorem B.9.5. Completeness of locally compact groups
Let G be a metrizable locally compact group with metric $\rho : G \times G \to [0, \infty)$. G is complete.

Proof. Since G is a metrizable topological space, it is completely regular and has a countable open basis at the origin $0 \in G$. Therefore, since G is a T_0 topological group the metric can be taken to be right translation invariant, i.e,

$$\forall\, x, y, z \in G, \quad \rho(x, y) = \rho(x + z, y + z),$$

see [227], pages 68–70, for a proof of translation invariance using a countable basis at the origin.

To prove that G is complete, let $\{x_n : n = 1, \ldots\} \subseteq G$ be Cauchy and find $x_0 \in G$ such that

$$\forall\, \varepsilon > 0,\ \exists\, N_\varepsilon > 0 \quad \text{such that} \quad \forall\, n \geq N_\varepsilon, \quad \rho(x_n, x_0) < \varepsilon. \tag{B.24}$$

By local compactness let C be a compact neighborhood of the origin and choose $r > 0$ such that $B(0, r) = \{x \in G : \rho(x, 0) < r\}$. Since $\{x_n : n = 1, \ldots\}$ is Cauchy,

$$\exists\, N_r \text{ such that } \forall\, n \geq N_r, \quad \rho(x_n, x_{N_r}) = \rho(x_n - x_{N_r}, 0) < r,$$

and so $\{y_n = x_n - x_{N_r} : n \geq N_r\} \subseteq B(0, r) \subseteq C$. Thus, there is a subsequence $\{x_{m_n} : n = 1, \ldots\}$ and $y_r \in G$ such that $\lim_{n\to\infty} \rho(y_{m_n}, y_r) = 0$. We set $x_0 = y_r + x_{N_r}$, and, hence,

$$0 = \lim_{n\to\infty} \rho(x_{m_n} - x_{N_r}, x_0 - x_{N_r}) = \lim_{n\to\infty} \rho(x_{m_n}, x_0). \tag{B.25}$$

We now verify (B.24). By (B.25),

$$\exists\, N_1 \text{ such that } \forall\, m_n > N_1, \quad \rho(x_{m_n}, x_0) < \frac{\varepsilon}{2}. \tag{B.26}$$

Also,

$$\exists\, N_2 \text{ such that } \forall\, m, n > N_2, \quad \rho(x_n, x_m) < \frac{\varepsilon}{2} \tag{B.27}$$

since $\{x_n : n = 1, \ldots\}$ is Cauchy; and, hence, if $N_\varepsilon = \max(N_1, N_2)$, then (B.27) implies that

$$\forall\, n, m_k > N_\varepsilon, \quad \rho(x_n, x_{m_k}) < \frac{\varepsilon}{2}. \tag{B.28}$$

Therefore, by (B.26) and (B.28), if $n > N_\varepsilon$, then

$$\rho(x_n, x_0) \leq \rho(x_n, x_{m_k}) + \rho(x_{m_k}, x_0) < \varepsilon,$$

where we have chosen $x_{m_k} > N_\varepsilon$.

□

Section B.9 is the beginning of the story of abstract harmonic analysis. RUDIN's [391] and HEWITT and ROSS' [227] treatises are a superb next step.

B.10 Dual groups and the Fourier analysis of measures

Let G be a LCAG. The collection Γ of all continuous group homomorphisms $\gamma : G \to \mathbb{T}$ is the *dual group* of G. Γ with pointwise multiplication is an Abelian group. The fact that Γ is a group follows from the definition, $\gamma_1 + \gamma_2$, by the rule $(\gamma_1 + \gamma_2)(x) = \gamma_1(x)\gamma_2(x)$, $x \in G$. Then, if Γ is equipped with the "weak * topology" $\sigma(\Gamma, G)$, it becomes a LCAG. In fact, we have not defined the weak * topology on groups so, to be precise, let us define a basis $\mathcal{B}$ for such a topology. For each compact $K \subseteq G$ and $r > 0$, let $N(K, r) = \{\gamma \in \Gamma : \forall\, x \in K,\ |1 - \gamma(x)| < r\}$. Set $\mathcal{B} = \{N(K, r) + \gamma : \gamma \in \Gamma,\ K \subseteq G \text{ is compact, and } r > 0\}$.

Given a Haar measure m_G on G. We define the *Fourier transform* of $f \in L^1_{m_G}(\mathcal{G})$ as

$$\forall\, \gamma \in \Gamma, \quad \mathcal{F}(f)(\gamma) = \int_G f(x)\overline{\gamma(x)}\, dm_G(x). \tag{B.29}$$

$\mathcal{F} : L^1_{m_G}(G) \to C_0(\Gamma)$ is a homomorphism, where multiplication is convolution in $L^1_{m_G}(G)$ and pointwise multiplication of functions in $C_0(\Gamma)$. This notion of Fourier transform can be extended to bounded Radon measures on G by means of the formula,

$$\forall\, \mu \in M_b(G), \quad \mathcal{F}(\mu)(\gamma) = \int_G \overline{\gamma(x)}\, d\mu(x). \tag{B.30}$$

$\mathcal{F}(\mu)$ is a bounded uniformly continuous function on Γ.

Example B.10.1. Homomorphisms and transforms
a. If $G = \mathbb{R}$, then $L^1_m R) \subseteq M_b(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R})$ with analogous inclusions for $G = \mathbb{R}^d$, see Definition 7.5.7. In this case, the dual group Γ can be identified with $\mathbb{R}$, which we denoted by $\widehat{\mathbb{R}}$ in Definition B.1.1. The identification can be made by proving that if $\gamma \in \Gamma$ is a continuous homomorphism $G \to \mathbb{T}$, then there is a real number $\xi = \xi_\gamma \in \widehat{\mathbb{R}}$ such that $\gamma(x) = e^{2\pi i x \xi}$. Clearly, if γ is of this form, then $\gamma \in \Gamma$. Thus, (B.29) is the same definition of Fourier transform as given in Definition B.1.1.

b. Let $\mathcal{A}$ be an algebra, such as $L^1_m(\mathbb{R})$ with multiplication defined by convolution. If $\mathcal{A}$ is a topological algebra such as the Banach algebra $L^1_m(\mathbb{R})$ with topology defined by the norm $\|...\|_1$, then we can consider the subspace $\mathcal{H} \subseteq \mathcal{A}'$ of continuous linear homomorphisms $\mathcal{A} \to \mathbb{C}$. This setting gives rise to a *transform* $\mathcal{T}$ for which there is an all important *exchange formula*, $\mathcal{T}(f * g) = \mathcal{T}(f)\mathcal{T}(g)$, where $f, g \in \mathcal{A}$.

For example, if $\mathcal{A} = L^1_m(\mathbb{R})$ as in part *a*, then $\mathcal{H}$ can be defined by $\{e^{2\pi i x \xi} : \xi = \xi_\gamma \text{ and } \gamma \in \Gamma = \widehat{\mathbb{R}}\}$ and $\mathcal{T}$ is the usual Fourier transform $\mathcal{F}$ as given in Definition B.1.1.

This is the basis of the *Gelfand theory*, and this particular point of view is expanded in [33].

c. Another example of part *b* is the case of $\mathcal{E}'(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R})$, the topological algebra of distributions having compact support with multiplication defined by convolution. When properly topologized, the dual of $\mathcal{E}'(\mathbb{R})$ is $C^\infty(\mathbb{R})$ [415], and $\mathcal{H}$ is defined by $\{e^{sx} : s \in \mathbb{C}\}$. In this case, the transform $\mathcal{T}$ is the classical bilateral *Laplace transform.*

d. Let $\mathbb{Z}_N$ be the set of integers $0, 1, \ldots, N-1$ under addition modulo N. $\mathbb{Z}_N$ is a commutative group with this definition of addition, and it is a compact Abelian group when it is taken with the discrete topology, i.e., every element is defined to be an open set. Then, the dual group $\widehat{\mathbb{Z}_N}$ is $\mathbb{Z}_N$. Further, the Fourier transform defined by (B.29) is the classical *discrete Fourier transform* (DFT) $\mathcal{F}_N$ defined for $f \in L^1_c(\mathbb{Z}_N)$ by

$$\forall\, n = 0, 1, \ldots, N-1, \quad \mathcal{F}_N(f)[n] = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} f[m] e^{-2\pi i mn/N}.$$

In Example A.13.5, we noted that $\mathcal{F}_N : L^2_c(\mathbb{Z}_N) \to L^2_c(\mathbb{Z}_N)$ is a unitary operator. Of course, in the setting of $\mathbb{Z}_N$, $L^1_c(\mathbb{Z}_N) = L^2_c(\mathbb{Z}_N) = \mathbb{C}^N$.

For a useful introduction to the DFT and its fast algorithm, the fast Fourier transform (FFT), see [39], Sections 3.8–3.10. For comprehensive treatments, see [73], [453], and [481].

Remark. One of the landmarks of SCHWARTZ' theory of distributions is that the Fourier transform can be defined in a meaningful and productive way. In fact, if $T \in \mathcal{S}'(\mathbb{R})$, then the Fourier transform $\widehat{T}$ of T is defined by the Parseval duality formula,

$$\forall\, f \in \mathcal{S}(\mathbb{R}), \quad \widehat{T}(\overline{\widehat{f}}) = T(\overline{f}), \tag{B.31}$$

cf., Theorem 8.7.6. In this case, $\widehat{T} \in \mathcal{S}'(\mathbb{R})$. If $T = \mu \in M_b(\mathbb{R})$, then $\hat{\mu}$ defined by (B.31 is the same as $\mathcal{F}(\mu)$ defined by (B.30).

It should be pointed out that the theory of distributions and corresponding harmonic analysis is highly developed on locally compact groups.

Definition B.10.2. Positive-definite functions

a. Let G be a LCAG with the dual group Γ. $R : \Gamma \to \mathbb{C}$ is *positive-definite* if

$$\forall\, c_1, \ldots, c_n \in \mathbb{C} \text{ and } \forall\, \gamma_1, \ldots, \gamma_n \in \Gamma, \quad \sum_{j,k=1}^{n} c_j \overline{c_k} R(\gamma_j - \gamma_k) \geq 0.$$

b. Let $F \in L^2_{m_\Gamma}(\Gamma)$, and define the involution $\tilde{F}(\gamma) = \overline{F(-\gamma)}$. Then, a straightforward calculation shows that $R = F * \tilde{F}$ is a continuous positive-definite function.

c. Another elementary calculation shows that *if $\mu \in M_b(G)$ is positive, then $\mathcal{F}(\mu)$ is a continuous positive-definite function.* The Bochner theorem, which we now state, is the converse. It was first published by GUSTAV HERGLOTZ (1911) for $\Gamma = \mathbb{Z}$. BOCHNER (1933) proved it for $\Gamma = \widehat{\mathbb{R}}$ and WEIL published the proof for any LCAG. For positive-definite distributions and their extensions, see [183].

Theorem B.10.3. Bochner theorem
Let R be a continuous positive-definite function on Γ. Then, there exists a unique positive bounded Radon measure $\mu \in M_b(G)$ such that $R = \mathcal{F}(\mu)$ on Γ.

There are several accessible conceptually different proofs, e.g., [139], [268], [391], cf., [415].

Remark. Because of our decomposition theorems for measures, the Bochner theorem allows us to assert that $F : \Gamma \to \mathbb{C}$ is of the form $\mathcal{F}(\mu)$, for some $\mu \in M_b(G)$, if and only if F is a finite linear combination of positive-definite functions. This does not give a useful, implementable, *intrinsic* characterization of $\mathcal{F}(M_b(G))$. For perspective, such a characterization does exist for $\mathcal{F}(L^2_{m_G}(G))$ because of the Plancherel theorem which is valid for LCAGs. This characterization for $\mathcal{F}(L^2_{m_G}(G))$ is that $F \in \mathcal{F}(L^2_{m_G}(G))$ *if and only if* $F \in L^2_{m_\Gamma}(\Gamma)$.

The problem for such an intrinsic characterization of $\mathcal{F}(M_b(G))$ or even $\mathcal{F}(L^1_m(\mathbb{R}))$ is *unsolved*. We know that $\mathcal{F}(L^1_m(\mathbb{R})) \subseteq C_0(\widehat{\mathbb{R}})$, but we do not have implementable conditions on $F \in C_0(\widehat{\mathbb{R}})$, as a function on $\widehat{\mathbb{R}}$, which are necessary and sufficient to assert that $F \in \mathcal{F}(L^1_m(\mathbb{R}))$.

Example B.10.4. Lévy continuity theorem
a. In Section 6.6.5 we defined a sequence $\{\mu_n : n = 1, \ldots\} \subseteq M_1^+(\mathbb{R}^d)$ to be *tight* if

$$\forall\, \varepsilon > 0,\ \exists\, K_\varepsilon \subseteq \mathbb{R}^d, \text{ compact, such that } \forall\, n = 1, \ldots, \quad \mu_n(K_\varepsilon) > 1 - \varepsilon.$$

(Besides being defined in Section 6.6.5, the space $M_1^+(\mathbb{R}^d)$ was also defined, in a slightly different way, in Theorem B.9.2.) We then stated the Prohorov theorem which describes the relation between convergence in the sense of Bernoulli and tight sequences. These ideas have far reaching consequences when dealing with the convergence of probability distributions, for example, when dealing with probability measures on certain infinite dimensional spaces. See [325] for the "down to earth" theory on $\mathbb{R}^d$. An important result in this area is the Lévy continuity theorem proved by PAUL LÉVY in 1925, e.g., [61].

b. A central part of the Lévy continuity theorem, which is also appropriate for this section, is the following result: *Given $\{\mu_n, \mu : n = 1, \ldots\} \subseteq M_1^+(\mathbb{R}^d)$; then, $\{\mu_n : n = 1, \ldots\}$ converges to μ in the sense of Bernoulli if and only if*

$$\forall\, \gamma \in \widehat{\mathbb{R}}^d, \quad \lim_{n\to\infty} \hat{\mu}_n(\gamma) = \hat{\mu}(\gamma). \tag{B.32}$$

For each fixed $\gamma \in \widehat{\mathbb{R}}^d$, $e^{-2\pi i x\cdot\gamma} \in C_b(\mathbb{R}^d)$ as a function of x, and so (B.32) is a consequence of Bernoulli convergence. The proof that (B.32) implies convergence in the sense of Bernoulli is more substantial, and we give an outline.

c. Assume (B.32). We know that $\hat{\mu}$ is continuous on $\widehat{\mathbb{R}}^d$, although for this proof we only need it to be continuous at $0 \in \widehat{\mathbb{R}}^d$. By (B.32), $\hat{\mu}(0) = \lim_{n\to\infty} \hat{\mu}_n(0) = 1$. Let $\varepsilon > 0$. By the continuity,

$$\exists\, \delta > 0 \text{ such that } \forall\, \xi \in Q_\delta, \quad |1 - \hat{\mu}(\xi)| < \frac{\varepsilon}{2}, \tag{B.33}$$

where $Q_\delta = \{\xi \in \widehat{\mathbb{R}}^d : \forall\, j = 1, \ldots, d,\ |\xi_j| \leq \delta\}$. Let $A = \{\xi \in \widehat{\mathbb{R}}^d : \forall\, j = 1, \ldots, d,\ |x_j| \leq \frac{1}{\pi\delta}\}$. Thus, if $x = (x_1, \ldots, x_d) \in A^\sim$, then some $|x_k| > 1/(\pi\delta)$; and so,

$$\left| \prod_{j=1}^{d} \frac{\sin(2\pi x_j \delta)}{2\pi x_j \delta} \right| < \frac{1}{2}.$$

Further,

$$\forall\, x \in \mathbb{R}^d, \quad \frac{1}{m^d(Q_\delta)} \int_{Q_\delta} e^{-2\pi i x\cdot\xi}\, d\xi = \prod_{j=1}^{d} \frac{\sin(2\pi x_j \delta)}{2\pi x_j \delta}.$$

Therefore, for each $n \geq 1$,

$$\begin{aligned} \frac{1}{m^d(Q_\delta)} \int_{Q_\delta} (1 - \hat{\mu}(\xi))\, d\xi &= \int_{\mathbb{R}^d} \left(1 - \frac{1}{m^d(Q_\delta)} \int_{Q_\delta} e^{-2\pi i x\cdot\xi}\, d\xi\right)\, d\mu_n(x) \\ &\geq \int_{A^\sim} \left(1 - \frac{1}{m^d(Q_\delta)} \int_{Q_\delta} e^{-2\pi i x\cdot\xi}\, d\xi\right)\, d\mu_n(x) \\ &> \frac{1}{2}\mu_n(A^\sim). \end{aligned} \tag{B.34}$$

Because of (B.32) and the fact that $\|\hat{\mu}_n\|_\infty \leq \|\mu\|_1 = 1$, we can use (B.33), (B.34), and LDC to assert that

$$\begin{aligned} \overline{\lim_{n\to\infty}}\, \mu_n(A^\sim) &\leq \overline{\lim_{n\to\infty}}\, \frac{2}{m^d(Q_\delta)} \int_{Q_\delta} (1 - \hat{\mu}_n(\xi))\, d\xi \\ &= \frac{2}{m^d(Q_\delta)} \int_{Q_\delta} (1 - \hat{\mu}(\xi))\, d\xi < \varepsilon. \end{aligned}$$

Consequently, we choose N such that $\mu_n(A^\sim) < \varepsilon$ for all $n > N$. Next, choose compact $K_\varepsilon \supseteq A$ for which $\mu_n(K_\varepsilon) \geq 1 - \varepsilon$ for $1 \leq n \leq N$. Thus, $\mu_n(K_\varepsilon) \geq 1 - \varepsilon$ for all $n \geq 1$, and $\{\mu_n : n = 1, \ldots\}$ is tight.

Thus, by part *b* of the Prohorov theorem (Theorem 6.6.5), there is $\nu \in M_1^+(\mathbb{R}^d)$ and a subsequence $\{\mu_{n_k} : k = 1, \ldots\} \subseteq \{\mu_n : n = 1, \ldots\}$ for which $\{\mu_{n_k} : k = 1, \ldots\}$ converges to ν in the sense of Bernoulli. Now, by the

Fourier uniqueness theorem for measures and another application (perhaps "corollary" is more accurate) of the Prohorov theorem, it is elementary to prove that $\{\mu_n : n = 1, \ldots\}$ converges to ν in the sense of Bernoulli and that $\nu = \mu$. This completes the proof.

B.11 Radial Fourier transforms

To outline Fourier analysis on $\mathbb{R}$ (or $\mathbb{Z}$ or $\mathbb{T}_{2\Omega}$) and then to define an essentially qualitative theory on LCAGs, as we have done, may be said to have missed a very big point. There remains the quantitative Fourier analysis of $\mathbb{R}^d$, with unresolved geometric intricacies in topics such as spectral synthesis and with analytic mysteries beyond extensions of the Paley–Littlewood theory and the singular integral operators which generalize the Hilbert transform. Some difficulties are implied in Sections 8.8.2–8.8.5, but, notwithstanding 20th century accomplishments, e.g., [196], so much is yet to be fathomed, see [163].

On the other hand, the radial theory, with its one-degree of freedom, is highly developed. A function $f : \mathbb{R}^d \to \mathbb{C}$ is *radial* if $f(x) = g(\|x\|)$, for some $g : [0, \infty) \to \mathbb{C}$, where $\|x\| = (x_1^2 + \ldots + x_d^2)^{1/2}$. An equivalent definition was given in Example 8.6.2*c*. Thus, f is radial if and only if $f(S(x)) = f(x)$ for all $S \in SO(d, \mathbb{R})$. We record some useful facts in this area. To whet one's appetite we give two useful examples, compute the Fourier transform of a radial function, and state a fascinating theorem.

Definition B.11.1. Fourier transform on $L^1_{m^d}(\mathbb{R}^d)$
The *Fourier transform* of $f \in L^1_{m^d}(\mathbb{R}^d)$ is the function F defined as

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

Also, the *dilation* f_λ, $\lambda > 0$, is $f_\lambda(x) = \lambda^d f(\lambda x)$. In particular, $\int_{\mathbb{R}^d} f_\lambda(x)\, dx = \int_{\mathbb{R}^d} f(x)\, dx$.

Example B.11.2. The Gaussian in $\mathbb{R}^d$
Let $f(x) = e^{-\pi r \|x\|^2}$, $r > 0$, $x \in \mathbb{R}^d$. By Example B.1.8 and the Fubini theorem, the Fourier transform F of f is

$$F(\xi) = r^{-d/2} e^{-\pi \|\xi\|^2 / r}.$$

Setting $G^d(x) = G(x_1) \ldots G(x_d)$, we have $\int_{\mathbb{R}^d} G^d(x)\, dx = 1$ and

$$(G^d_\lambda)\hat{}(\xi) = e^{-(\pi \|\xi\| / \lambda)^2}.$$

Example B.11.3. The Poisson function
We shall define and compute the Fourier transform of the natural generalization to $\mathbb{R}^d$ of the Poisson function on $\mathbb{R}$, as defined in Example B.1.5 and Example B.3.4.

a. We begin on $\mathbb{R}$, using either Theorem B.1.2 or Theorem B.3.7, to compute

$$e^{-2\pi r|\gamma|} = \frac{2r}{\pi}\int_0^\infty \frac{\cos(2\pi t\gamma)}{r^2+t^2}\,dt. \tag{B.35}$$

If $r = 1$, we use (B.35) to obtain

$$\begin{aligned} e^{-2\pi|\gamma|} &= \frac{2}{\pi}\int_0^\infty \cos(2\pi t\gamma)\left(\int_0^\infty e^{-u}e^{-ut^2}\,du\right)dt \\ &= \frac{2}{\pi}\int_0^\infty e^{-u}\left(\frac{1}{2}\int_{-\infty}^\infty e^{-ut^2}e^{2\pi i t\gamma}\,dt\right)du. \end{aligned}$$

We apply Example B.1.8 to the right side, so that

$$e^{-2\pi|\gamma|} = \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{e^{-u}}{\sqrt{u}}e^{-(\pi\gamma)^2/u}\,du.$$

Thus, for $r > 0$ and $\gamma \in \mathbb{R}$,

$$e^{-2\pi r|\gamma|} = \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{e^{-u}}{\sqrt{u}}e^{-(\pi r\gamma)^2/u}\,du. \tag{B.36}$$

b. Next, we recall the definition of the *gamma function* Γ:

$$\forall\, u > 0, \quad \Gamma(u) = \int_0^\infty x^{u-1}e^{-x}\,dx.$$

Using the Laplace transform $\mathcal{L}$ in Example B.4.5*c*, and for fixed $u > -1$, we have

$$\forall\, s \text{ such that } \operatorname{Re}(s) > 0, \quad \mathcal{L}(x^u)(s) = \frac{1}{s^{u+1}}\Gamma(u+1),$$

as a function of s. It is easy to see that if $u = m \in \mathbb{N}$ then $\mathcal{L}(x^m)(s) = m!/s^{m+1}$ and so $\Gamma(m+1) = m!$. Further, $\Gamma(u+1) = u\Gamma(u)$ and $\Gamma(1/2) = \sqrt{\pi}$.

c. Now consider $e^{-2\pi r\|\xi\|}$, $\xi \in \widehat{\mathbb{R}}^d$, and take its inverse Fourier transform using (B.36) and Example B.11.2:

$$\begin{aligned} \int_{\widehat{\mathbb{R}}^d} e^{-2\pi r\|\xi\|}e^{2\pi i x\cdot\xi}\,d\xi &= \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{e^{-u}}{\sqrt{u}}\left(\int_{\widehat{\mathbb{R}}^d} e^{-(\pi r\|\xi\|)^2/u}e^{2\pi i x\cdot\xi}\,d\xi\right)du \\ &= \frac{1}{r^d\pi^{(d+1)/2}}\int_0^\infty e^{-u(1+\|x\|^2/r^2)}u^{(d-1)/2}\,du \\ &= \frac{r}{\pi^{(d+1)/2}}\frac{1}{(r^2+\|x\|^2)^{(d+1)/2}}\int_0^\infty e^{-v}v^{(d-1)/2}\,dv. \end{aligned}$$

Therefore,

$$\int_{\widehat{\mathbb{R}}^d} e^{-2\pi r\|\xi\|}e^{2\pi i x\cdot\xi}\,d\xi = \frac{r\Gamma((d+1)/2)}{\pi^{(d+1)/2}}\frac{1}{(r^2+\|x\|^2)^{(d+1)/2}}. \tag{B.37}$$

The right side of (B.37) is the *Poisson function* on $\mathbb{R}^d$ and $e^{-2\pi r\|\xi\|}$ is its Fourier transform on $\widehat{\mathbb{R}}^d$.

Example B.11.4. Bessel functions on $\mathbb{R}$
We shall use *Bessel functions of the first kind,*

$$J_\nu(x) = \frac{(x/2)^\nu}{\sqrt{\pi}\Gamma(\nu+1/2)} \int_0^\pi e^{-ix\cos(\theta)} \sin^{2\nu}(\theta)\, d\theta,$$

for the case $\nu = (d-2)/2$. The functions J_ν arose as solutions of the differential equation,

$$z^2 \frac{d^2 w}{dz^2} + z\frac{dw}{dz} + (z^2 - \nu^2)w = 0,$$

see [350].

Theorem B.11.5. The Fourier transform of radial functions
Let $f \in L^1_{m^d}(\mathbb{R}^d)$ be radial. Then, its Fourier transform F is radial. Further,

$$F(\|\xi\|) = \frac{2\pi}{\|\xi\|^{(d-2)/2}} \int_0^\infty r^{d/2} J_{(d-2)/2}(2\pi r\|\xi\|)\phi(r)\, dr, \tag{B.38}$$

where $r = \|x\|$ and $\phi(r) = f(\|x\|)$.

The proof that F is radial comes down to checking that $F(\xi)$ is invariant with respect to rotations about the origin. The formula (B.38) results from Theorem 8.7.10, Example 8.7.11, and Example B.11.4, e.g., [413], pages 222–227.

Example B.11.6. Radial Fourier transforms for $d = 1, 2, 3$
Given the notation of Theorem B.11.5. If $d = 1$, then

$$F(|\xi|) = 2\int_0^\infty \cos(2\pi r|\xi|)\phi(r)\, dr$$

as predicted from Proposition B.1.7. If $d = 2$, then

$$F(\|\xi\|) = 2\pi \int_0^\infty r J_0(2\pi r\|\xi\|)\phi(r)\, dr.$$

If $d = 3$, then

$$F(\|\xi\|) = \frac{2}{\|\xi\|}\int_0^\infty r\sin(2\pi r\|\xi\|)\phi(r)\, dr.$$

We close this section with SCHOENBERG's surprising result (1938), see [407], [139], pages 201–206.

Theorem B.11.7. Schoenberg theorem
Let $F : [0,\infty) \to \mathbb{C}$ be continuous, and, for all $d \geq 1$, define $R(\xi) = F(\|\xi\|)$, $\xi \in \widehat{\mathbb{R}}^d$. R is positive-definite on $\widehat{\mathbb{R}}^d$ for all $d \geq 1$ if and only if

$$\exists\, \mu \in M_b([0,\infty)),\ \mu \geq 0, \text{ such that } \forall\, \gamma \geq 0, \quad F(\gamma) = \int_0^\infty e^{-x\gamma^2}\, d\mu(x),$$

noting this integral is the Laplace transform $\mathcal{L}(\mu)(\gamma^2)$.

B.12 Wiener's Generalized Harmonic Analysis (GHA)

In this section we not only have a chance to advertise WIENER's beautiful theory of Generalized Harmonic Analysis (GHA), but the theory itself combines many of the ideas we have introduced including the Bochner theorem and distribution theory. It is particularly exciting because of the continuing applicability of these notions, e.g., power spectra and spectrograms, in physics and signal processing. See the penultimate bullet in Section 9.6.4.

In 1930, NORBERT WIENER [496], Volume II, pages 183–324, proved an analogue of the Parseval–Plancherel formula, $\|f\|_2 = \|\hat{f}\|_2$, for functions which are *not* elements of $L^2(\mathbb{R})$. We refer to his formula as the *Wiener–Plancherel formula*, e.g., (B.40). It became a beacon in his perception and formulation of the statistical theory of communication, e.g., [493], [308]. WIENER [496] even chose to have the formula appear on the cover of his autobiography, *I Am a Mathematician.* (This is a 20th century analogue of ARCHIMEDES' tombstone, which had a carving of a sphere inscribed in a cylinder to commemorate his "1:2:3" theorem, see Section 3.9.1 for details concerning the mathematical results, CICERO's role, and a recent update.)

Besides the use of GHA as an explanation of the polychromatic nature of sunlight, WIENER discussed the background for GHA in [496], Volume II, pages 183–324; and this background has been explained scientifically and historically in a virtuoso display of scholarship by MASANI, e.g., MASANI's remarkable commentaries in [496], Volume II, pages 333–379, as well as [331]. Two precursors, whose work Wiener studied and who should be mentioned vis a vis GHA, were Sir ARTHUR SCHUSTER and Sir GEOFFREY I. TAYLOR. SCHUSTER pointed out analogies between the harmonic analysis of light and the statistical analysis of hidden periods associated with meteorological and astronomical data. TAYLOR conducted experiments in fluid mechanics dealing with the *onset to turbulence*, and formulated a special case of correlation. A third scientist, whose work (1914) vis a vis GHA was not known to WIENER, was ALBERT EINSTEIN. EINSTEIN writes: "Suppose the quantity y (for example, the number of sun spots) is determined empirically as a function of time, for a very large interval, T. How can one represent the statistical behavior of y?" In his heuristic answer to this question he came close to the notions of autocorrelation and power spectrum, e.g., B.12.5, cf., [331], pages 112–113, EINSTEIN's paper [153], and commentaries by MASANI [330] and YAGLOM [501].

The Fourier analysis of $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ or the theory of Fourier series were inadequate tools to analyze the issues confronting SCHUSTER, TAYLOR, and EINSTEIN. On the other hand, GHA became a successful device to gain some insight into their problems, as well as other problems where the data and/or noises can not be modelled by the Fourier transform decay, finite energy, or periodicity inherent in the above classical theories, e.g., [9], Chapter II, [29], [373].

The material in this Section outlines GHA and is due to WIENER [496], Volume II, pages 183–324 and pages 519–619, [490], cf., [33] Chapter 2, [52]. The higher dimensional theory, with its geometrical ramifications, is found in [41], [37], cf., [6].

Definition B.12.1. Bounded quadratic means
The space $BQM(\mathbb{R})$ of functions having *bounded quadratic means* is the set of all functions $f \in L^2_{\text{loc}}(\mathbb{R})$ for which

$$\sup_{T>0} \frac{1}{2T} \int_{-T}^{T} |f(t)|^2 \, dt < \infty.$$

The *Wiener space* $W(\mathbb{R})$ is the set of all functions $f \in L^2_{\text{loc}}(\mathbb{R})$ for which

$$\int \frac{|f(t)|^2}{1+t^2} \, dt < \infty.$$

Theorem B.12.2. Inclusions for GHA
The following inclusions hold:

$$L^\infty_m(\mathbb{R}) \subseteq BQM(\mathbb{R}) \subseteq W(\mathbb{R}) \subseteq S'(\mathbb{R}).$$

Moreover, all the inclusions are proper.

Definition B.12.3. The Wiener s-function
The *Wiener s-function* associated with $f \in BQM(\mathbb{R})$ is defined as the sum $s = s_1 + s_2$, where

$$s_1(\xi) = \int_{-1}^{1} f(t) \frac{e^{-2\pi i t\xi} - 1}{-2\pi i t} \, dt$$

and

$$s_2(\xi) = \int_{|t|\geq 1} f(t) \frac{e^{-2\pi i t\xi}}{-2\pi i t} \, dt.$$

Since $f \in L^1[-1,1]$, we have $s_1 \in C(\mathbb{R})$ and $|s_1(\xi)| \leq 2|\xi| \|f\|_{L^1[-1,1]}$. Since $f \in BQM(\mathbb{R})$, Theorem B.12.2 and the Parseval–Plancherel theorem allow us to conclude that $s_2 \in L^2_m(\widehat{\mathbb{R}})$. In particular, $s \in L^2_{\text{loc}}(\widehat{\mathbb{R}}) \cap S'(\widehat{\mathbb{R}})$.

Theorem B.12.4. The derivative of the Wiener s-function
Let $f \in BQM(\mathbb{R})$. Then, $f \in S'(\mathbb{R})$ and

$$s' = \hat{f},$$

where $s \in L^2_{\text{loc}}(\widehat{\mathbb{R}}) \cap S'(\widehat{\mathbb{R}})$ is the Wiener s-function associated with f.

Definition B.12.5. Deterministic autocorrelation
The *deterministic autocorrelation* R of a function $f : \mathbb{R} \to \mathbb{C}$ is formally defined as

$$R(t) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(u+t)\overline{f(u)}\,du.$$

To fix ideas, suppose R exists for each $t \in \mathbb{R}$. It is easy to prove that R is positive-definite, and so $R = \widehat{S}$ for some positive bounded Radon measure $S \in M_b(\widehat{\mathbb{R}})$ by Theorem B.10.3. There is also a notion of *stochastic autocorrelation* arising in the study of stationary stochastic processes, see [284], [283], [140], [61]. Deterministic and stochastic autocorrelation are the same if the process has the property of being *correlation ergodic*, see [356]. This notion is not unrelated to the ergodic theorem discussed in Section 8.8.6. S is called the *power spectrum* of f, and, in applications, letters such as "S" are used instead of "μ".

The *Wiener–Plancherel formula* is equation (B.40) in the following result.

Theorem B.12.6. Wiener–Plancherel formula
Let $f \in BQM(\mathbb{R})$, and suppose its deterministic autocorrelation $R = \widehat{S}$ exists for each $t \in \mathbb{R}$.

a. *Then,*

$$\forall\, t \in \mathbb{R}, \quad R(t) = \lim_{\varepsilon\to 0} \frac{2}{\varepsilon}\int |\Delta_\varepsilon s(\xi)|^2 e^{-2\pi i t\xi}\,d\xi, \tag{B.39}$$

where $\Delta_\varepsilon s(\xi) = \frac{1}{2}(s(\xi+\varepsilon) - s(\xi-\varepsilon))$.

b. *In particular,*

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2\,dt = \lim_{\varepsilon\to 0} \frac{2}{\varepsilon}\int |\Delta_\varepsilon s(\xi)|^2\,d\xi. \tag{B.40}$$

Example B.12.7. Related formulas and spectral estimation

a. Because of (B.39) and assuming the setup of Theorem B.12.6, the following formulas are true under the proper hypotheses, e.g., [33], page 90, [36], page 847:

$$\lim_{\varepsilon\to 0} \frac{2}{\varepsilon}|\Delta_\varepsilon s(\xi)|^2 = S, \tag{B.41}$$

and

$$\begin{aligned}\int |\hat{k}(\xi)|^2\,dS(\xi) &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |k * f(t)|^2\,dt \\ &= \lim_{\varepsilon\to 0} \frac{2}{\varepsilon}\int |\hat{k}(\xi)\Delta_\varepsilon s(\xi)|^2\,d\xi.\end{aligned} \tag{B.42}$$

b. Formally, (B.42) is (B.40) for the case $k = \delta$. For $k \in C_c(\mathbb{R})$ the first equality of (B.42) is not difficult, e.g., [36], pages 847–848. The second equality, or, equivalently, Theorem B.12.6, requires the *Wiener Tauberian theorem* (Theorem B.12.9).

c. The following diagram illustrates the action and "levels" of the functions and measure in Theorem B.12.6 for a given signal f.

$$\begin{array}{ccc} f & \longleftrightarrow \hat{f} = s' & s \\ \downarrow & & \downarrow \\ R = \widehat{S} \longleftrightarrow & S & \left\{\frac{2}{\varepsilon}|\Delta_\varepsilon s|^2|\right\} \end{array}$$

d. Since S is the "power" spectrum, (B.40) and (B.41) allow us to assert that

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |f(t)|^2 \, dt$$

is a measure of the total power of f, cf., WIENER's comparison of energy and power in [493], pages 39–40 and 42. In light of the spectral estimation problem, see, e.g., [39], Definition 2.8.6, the middle term of (B.42) is a measure of the power in a frequency band $[\alpha, \beta]$ if $\hat{k} = \mathbb{1}_{[\alpha,\beta]}$ in the first term of (B.42), cf., [36], Theorem 5.2.

Remark. The Parseval–Plancherel formula, $\|f\|_2 = \|\hat{f}\|_2$, allowed us to define the Fourier transform of a square integrable function (Theorem B.4.2), and, at certain levels of abstraction, it is considered to characterize what is meant by an harmonic analysis of f. On the other hand, for most applications in $\mathbb{R}$, the formula assumes the workaday role of an effective tool used to obtain quantitative results. It is this latter role that was envisaged for the Wiener–Plancherel formula in dealing with the non-square-integrable case. After all, distribution theory gives the proper definition of the Fourier transform of tempered distributions. The real issue is to obtain quantitative results for problems where an harmonic analysis of a non-square-integrable function is desired. As mentioned above, a host of such problems comes under the heading of an harmonic (spectral) analysis of signals containing non-square-integrable noise and/or random components, whether it be speech recognition, image processing, geophysical modeling, or turbulence in fluid mechanics. Such problems can be attacked by BEURLING's profound theory of spectral synthesis, e.g., [33], as well as by the extensive multifaceted theory of time series, e.g., [366]. BEURLING's spectral synthesis does not deal with energy and power considerations, i.e., quadratic criteria, and time series relies on a stochastic point of view. The Wiener–Plancherel formula deals with these problems deterministically, and, hence, with potential for real implementation.

Example B.12.8. Elementary power spectra
a. The value of an autocorrelation R is that it can be measured in many cases where the underlying signal f can not be quantified. This is the basis of the Michelson interferometer. Also, the discrete part of the power spectrum S characterizes periodicities in f, e.g., [492], Chapter X. This can be illustrated by taking $f(t) = \sum_{k=1}^{n} r_k e^{-2\pi i t \lambda_k}$, $r_k \in \mathbb{C}$, $\lambda_k \in \widehat{\mathbb{R}}$. The

L^2-autocorrelation is not defined, but the deterministic autocorrelation is $R(t) = \sum_{k=1}^{n} |r_k|^2 e^{-2\pi i t \lambda_k}$ (by direct calculation); and hence the power spectrum is

$$S = \sum_{k=1}^{n} |r_k|^2 \, \delta_{\lambda_k}.$$

b. If $f : \mathbb{R} \to \mathbb{C}$ has the property that $\lim_{|t|\to\pm\infty} f(t) = 0$, then $S = 0$. It is elementary to construct examples f for which $S = 0$ whereas $\overline{\lim}_{|t|\to\pm\infty} |f(t)| > 0$, cf., [490], pages 151–154, [29], pages 99–100, [33], pages 84 and 87, [35], Section IV.

As mentioned in Example B.12.7, the following result is required to prove the Wiener–Plancherel formula. It was first stated in the Remark after Problem 5.8.

Theorem B.12.9. Wiener Tauberian theorem
Let $f \in L^1_m(\mathbb{R})$ have a non-vanishing Fourier transform and let $g \in L^\infty_m(\mathbb{R})$. If

$$\lim_{t\to\infty} f * g(t) = r \int_{-\infty}^{\infty} f(u)\, du, \tag{B.43}$$

then

$$\forall h \in L^1_m(\mathbb{R}), \quad \lim_{t\to\infty} h * g(t) = r \int_{-\infty}^{\infty} h(u)\, du. \tag{B.44}$$

Remark. ***a.*** Theorem B.12.9 has the format of *classical Tauberian theorems*: A boundedness (or related) condition and "summability" by a certain method yield "summability" by other methods. In Theorem B.12.9, the boundedness or "Tauberian" condition is the hypothesis that $g \in L^\infty_m(\mathbb{R})$. The given summability is (B.43), where f represents a so-called "summability method". The conclusion (B.44) of the theorem is summability for a whole class of summability methods, viz., for all $h \in L^1_m(\mathbb{R})$. A classical and masterful treatment of summability methods is due to HARDY [210].

If G is the Gaussian defined in Example B.1.8, then $\widehat{G}$ never vanishes. Thus, in this case, if $g \in L^\infty_m(\mathbb{R})$ has the property that

$$\lim_{t\to\infty} G * g(t) = r,$$

then

$$\forall\, \lambda, \quad \lim_{t\to\infty} W_\lambda * g(t) = r,$$

where $\{W_\lambda\}$ is the Fejér kernel.

The particular functions used by WIENER to prove his Wiener Tauberian formulas are found in [490], [33], pages 91-92.

b. *Modern* Tauberian theorems have an algebraic and/or functional analytic flavor to them. For example, the Wiener Tauberian theorem is a special case of the fact that *if $\hat{f} \in A(\mathbb{R})$, $T \in A'(\mathbb{R})$, and $T\hat{f} = 0$, then $\hat{f} = 0$ on* supp T. In fact, the generalizations of Theorem B.12.9 are much more far reaching than this. [33] gives an extensive treatment of both classical and modern Tauberian theory, as well as the history of the subject, and applications to spectral synthesis and analytic number theory, see also, e.g., [239].

Because of the importance of translation invariant systems and the theory of multipliers, we define the *closed translation invariant subspace* V_f *generated by* $f \in X$, where X is $L^1_m(\mathbb{R})$ or $L^2_m(\mathbb{R})$, to be the closure in X of the linear span of translations of f by $t \in \mathbb{R}$, i.e.,

$$V_f = \overline{\text{span}}\{\tau_t(f) : t \in \mathbb{R}\}. \tag{B.45}$$

Theorem B.12.10. Zero sets and dense subspaces
a. *If* $f \in L^1_m(\mathbb{R})$ *and* $\hat{f}$ *never vanishes, then* $V_f = L^1_m(\mathbb{R})$.
b. *If* $f \in L^2_m(\mathbb{R})$ *and* $|\hat{f}| > 0$ *a.e., then* $V_f = L^2_m(\mathbb{R})$.

Proof. Part a is the Wiener Tauberian theorem, and we refer to [490], [33], pages 25–26, 49–50, 94–95, and Section 2.3 for proofs.

The proof of part b is much simpler than that of part a, and so we shall give it. Suppose $V_f \neq L^2_m(\mathbb{R})$. Then, there is $h \in L^2_m(\mathbb{R}) \setminus \{0\}$ such that

$$\forall\, t \in \mathbb{R}, \quad \int (\tau_t(f))(u)\overline{h(u)}\, du = 0. \tag{B.46}$$

Equation (B.46) is a consequence of the Hahn–Banach theorem and the fact that $L^2_m(\mathbb{R})' = L^2_m(\mathbb{R})$. By the Parseval–Plancherel theorem,

$$\forall\, t \in \mathbb{R}, \quad \int \hat{f}(\xi)\overline{\hat{h}(\xi)} e^{-2\pi i t\xi}\, d\xi = 0.$$

$\hat{f}\overline{\hat{h}} \in L^1_m(\widehat{\mathbb{R}})$ by the Hölder inequality, and so, by the L^1-uniqueness theorem (Theorem B.3.5c), $\hat{f}\overline{\hat{h}} = 0$ a.e. Since $|\hat{f}| > 0$ a.e., we conclude that $\hat{h} = 0$ a.e., and this contradicts the hypothesis on h. Thus, $V_f = L^2(\mathbb{R})$.

□

Subspaces such as V_f in (B.45) play an important role in Gabor and wavelet decompositions in the case that the set of translates $\tau_t(f)$ is reduced to $\{\tau_r(f) : r \in D\}$ where D is a discrete subset of $\mathbb{R}$, e.g., [338], [114], [45], [301].

Remark. ***a.*** Let $M_b^+(\widehat{\mathbb{R}})$ be the space of positive bounded Radon measures on $\widehat{\mathbb{R}}$. In GHA, a function f is analyzed for its frequency information by computing its autocorrelation R and its power spectrum $S = R^\vee \in M_b^+(\widehat{\mathbb{R}})$. Mathematically, this is a mapping between a class of functions f and a class of measures $S \in M_b^+(\widehat{\mathbb{R}})$. A natural question to ask is the following: For any $\mu \in M_b^+(\widehat{\mathbb{R}})$, does there exist f whose autocorrelation R exists, and for which $\check{R} = \mu$?

b. The question of part a is answered affirmatively in the case of weakly stationary stochastic processes (WSSPs) by the Wiener–Khinchin theorem: *A necessary and sufficient condition for* R *to be the stochastic autocorrelation of some* WSSP X *is that there exist* $S \in M_b^+(\widehat{\mathbb{R}})$ *for which* $\widehat{S} = R$. In one direction, if R is the stochastic autocorrelation of a WSSP X, then $S = R^\vee \in M_b^+(\widehat{\mathbb{R}})$ by Theorem B.10.3. The question in part a deals with the opposite direction, and the positive answer is not difficult to prove, e.g., [366], pages 221-222, [148], pages 62–63 and 72–73. KHINCHIN's proof dates from 1934, and there were further probabilistic contributions by WOLD (1938), CRAMÉR (1940), and KOLMOGOROV [283], cf., [38].

c. The deterministic and *constructive* affirmative answer to the question in part a is the *Wiener–Wintner theorem* (1939) [496]. JEAN BASS and JEAN-PAUL BERTRANDIAS made significant contributions to this result, e.g., [29]; and the multidimensional version is found in [36], [273].

Theorem B.12.11. Wiener–Wintner theorem
Let $\mu \in M_b^+(\widehat{\mathbb{R}})$. *There is a constructible function* $f \in L^\infty_{loc}(\mathbb{R})$ *such that its deterministic autocorrelation* R *exists for all* $t \in \mathbb{R}$, *and* $\widehat{R} = \mu$.

B.13 Epilogue

This appendix serves as a handmaiden to the book, but the material is really a preface to harmonic analysis as one of the goddesses of mathematics. There are magnificent, profound edifices from classical Fourier series to representation theory, from non-harmonic Fourier series to sampling, wavelets, and time-frequency analysis, from Fourier methods in classical partial differential equations to pseudodifferential operators, from the computation of Gauss sums to the role of Fourier analysis at all levels of analytic number theory, and from Fast Fourier Transforms to an ever expanding litany of genuine applications. We have referenced introductory texts and groundbreaking treatises, and encyclopedic works of scholarship.