which are dense subspaces of $L^p(\mathbb{R})$. On these domains, $P\colon D_P \to L^p(\mathbb{R})$ and $M\colon D_M \to L^p(\mathbb{R})$. Show, however, that P and M are unbounded even when restricted to these domains, i.e.,

$$\sup_{\substack{f\in D_P,\\ \|f\|_p=1}} \|Pf\|_p \;=\; \infty \;=\; \sup_{\substack{f\in D_M,\\ \|f\|_p=1}} \|Mf\|_p.$$

1.5 Approximate Identities

Although $L^1(\mathbb{R})$ is closed under convolution, we have seen that it has no identity element. In this section we will show that there are functions in $L^1(\mathbb{R})$ that are "almost" identity elements for convolution. We will construct families of functions $\{k_\lambda\}_{\lambda>0}$ which have the property that $f * k_\lambda$ converges to f in $L^1(\mathbb{R})$ (and in fact in many other senses, depending on what space f belongs to). If k_λ is a "nice" function, then $f * k_\lambda$ will inherit that niceness, and so we will have a nice function that is arbitrarily close to f. Using this procedure we will be able to show that many spaces of nice functions are dense in $L^1(\mathbb{R})$, including $C_c(\mathbb{R})$, $C_c^m(\mathbb{R})$ for each $m \in \mathbb{N}$, and even $C_c^\infty(\mathbb{R})$.

1.5.1 Definition and Existence of Approximate Identities

The properties that a family $\{k_\lambda\}_{\lambda>0}$ will need to possess in order to be an approximate identity for convolution are listed in the next definition.

Definition 1.51. An *approximate identity* or a *summability kernel* is a family $\{k_\lambda\}_{\lambda>0}$ of functions in $L^1(\mathbb{R})$ such that

(a) $\displaystyle\int k_\lambda(x)\,dx \;=\; 1$ for every $\lambda > 0$,

(b) $\displaystyle\sup_\lambda \|k_\lambda\|_1 \;<\; \infty$, and

(c) for every $\delta > 0$,

$$\lim_{\lambda\to\infty} \int_{|x|\ge\delta} |k_\lambda(x)|\,dx \;=\; 0.$$

Note that, by definition, an approximate identity is a family of integrable functions. If it is the case that $k_\lambda \ge 0$ for each λ, then requirement (b) in Definition 1.51 follows from requirement (a). However, in general the elements of an approximate identity need not be nonnegative functions.

The "easy" way to create an approximate identity is through dilation of a single function.

Exercise 1.52. Let $k \in L^1(\mathbb{R})$ be any function that satisfies

$$\int k(x)\,dx \;=\; 1.$$

Define k_λ by an L^1-normalized dilation:

$$k_\lambda(x) \;=\; \lambda\, k(\lambda x), \qquad \lambda > 0,$$

and show that the resulting family $\{k_\lambda\}_{\lambda>0}$ forms an approximate identity.

Note that there is an inherent ambiguity in our notation: We may use $\{k_\lambda\}_{\lambda>0}$ to indicate a generic family of functions indexed by λ, or, as introduced in Notation 1.5, we may use k_λ to denote the L^1-normalized dilation of a function k. The intended meaning is usually clear from context.

In any case, if we define k_λ by dilation, then, as λ increases, k_λ becomes more and more similar to our intuition of what a "δ-function" (a function that is an identity for convolution) would look like (see the illustration in Figure 1.7 and the related discussion in Section 1.3.5). While there is no such identity for convolution in $L^1(\mathbb{R})$, the collection of functions $\{k_\lambda\}_{\lambda>0}$ in some sense forms an approximation to this nonexistent δ-function, for requirement (c) implies that k_λ becomes more and more concentrated near the origin as λ increases, with $\int k_\lambda = 1$ for every λ.

Consider also the appearance of $k_\lambda(x) = \lambda k(\lambda x)$ in the frequency domain. By Exercise 1.13, we have

$$\widehat{k_\lambda}(\xi) \;=\; \widehat{k}\biggl(\frac{\xi}{\lambda}\biggr).$$

Since k is integrable with $\int k = 1$, we know that $\widehat{k}$ is continuous, and

$$\widehat{k}(0) \;=\; \int k \;=\; 1.$$

The continuity of $\widehat{k}$ therefore implies that, for each $\xi \in \mathbb{R}$,

$$\lim_{\lambda\to\infty} \widehat{k_\lambda}(\xi) \;=\; \lim_{\lambda\to\infty} \widehat{k}\biggl(\frac{\xi}{\lambda}\biggr) \;=\; \widehat{k}(0) \;=\; 1 \tag{1.22}$$

(see the illustration in Figure 1.9 using the Gaussian function $g(x) = e^{-\pi x^2}$, which has the interesting property that $\widehat{g}(\xi) = e^{-\pi\xi^2}$, see Exercise 1.88). Thus $\widehat{k_\lambda}$ converges pointwise everywhere to the constant function 1. This again matches our intuition for what the Fourier transform of a "δ-function" would be if there was one, for if there was a function δ that satisfied both $\delta(x) = 0$ for $x \neq 0$ *and* $\int \delta = 1$, then $\widehat{\delta}$ would be identically constant:

$$\widehat{\delta}(\xi) \;=\; \int \delta(x)\, e^{-2\pi i\xi x}\, dx \;=\; \int \delta(x)\, dx \;=\; 1.$$

This contradicts the Riemann–Lebesgue Lemma and so no such function δ can exist. However, when we define δ not as a function but rather as a distribution

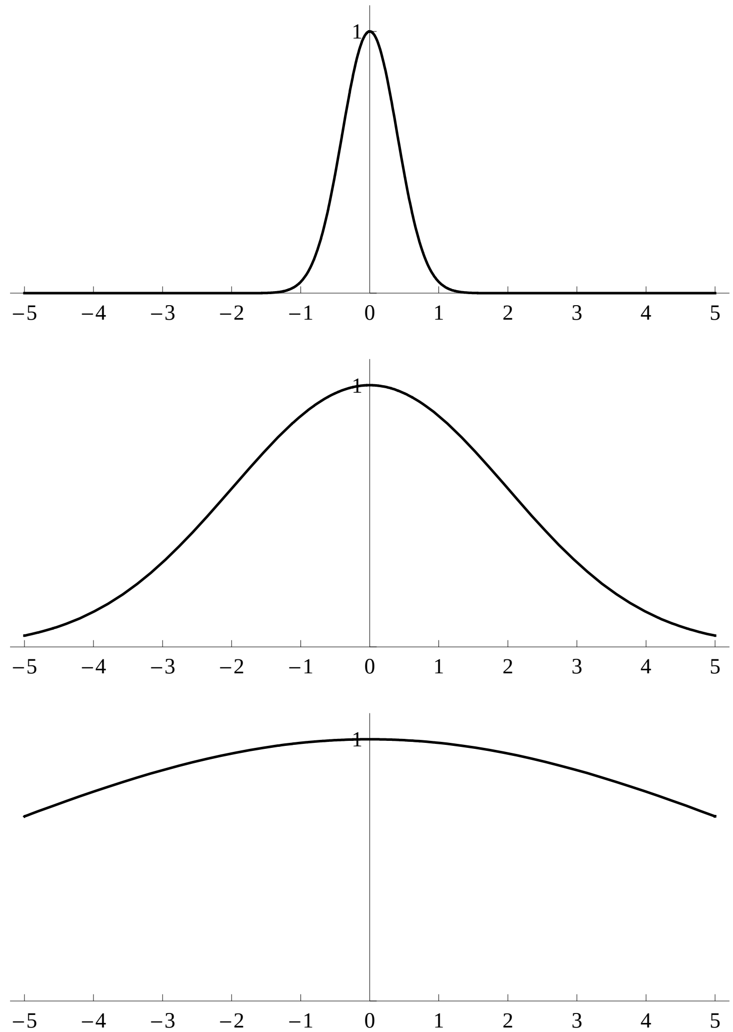

Fig. 1.9. Top: The Fourier transform $\widehat{g}(\xi) = e^{-\pi\xi^2}$ of the function $g(x) = e^{-\pi x^2}$. Middle: The Fourier transform $\widehat{g_5}(\xi) = \widehat{g}(\xi/5)$ of the dilated function $g_5(x) = 5g(5x)$. Bottom: The Fourier transform $\widehat{g_{15}}(\xi) = \widehat{g}(\xi/15)$ of the dilated function $g_{15}(x) = 15g(15x)$.

in Chapter 3 or as a measure in Chapter 4, we will see that $\widehat{\delta}$ does exist and is precisely the constant function.

In any case, given any $f \in L^1(\mathbb{R})$ and an approximate identity $\{k_\lambda\}_{\lambda>0}$ of the form $k_\lambda(x) = \lambda k(\lambda x)$, we have that

$$(f * k_\lambda)^\wedge(\xi) \;=\; \widehat{f}(\xi)\,\widehat{k_\lambda}(\xi) \;\to\; \widehat{f}(\xi), \qquad \xi \in \mathbb{R}.$$

So, we at least have that $(f * k_\lambda)^\wedge(\xi)$ converges pointwise to $\widehat{f}(\xi)$, and this gives us hope that $f * k_\lambda$ should converge to f in other senses as well. Our goal in this section is understand in what sense this hope holds true.

1.5.2 Approximation in $L^p(\mathbb{R})$ by an Approximate Identity

We begin by quantifying the notion that an approximate identity is approximately an identity for convolution in $L^1(\mathbb{R})$. The proof of this theorem illustrates the "standard trick" of introducing k_λ into an equation by virtue of the fact that $\int k_\lambda = 1$, and also the division of the integral into small and large parts in order to make use of the defining properties of an approximate identity.

Theorem 1.53. *Let $\{k_\lambda\}_{\lambda>0}$ be an approximate identity. Then*

$$\forall\, f \in L^1(\mathbb{R}), \qquad \lim_{\lambda\to\infty} \|f - f * k_\lambda\|_1 \;=\; 0.$$

That is, $f * k_\lambda \to f$ in L^1-norm as $\lambda \to \infty$.

Proof. Fix any $f \in L^1(\mathbb{R})$. Since $k_\lambda \in L^1(\mathbb{R})$, we know that $f * k_\lambda \in L^1(\mathbb{R})$, and we wish to show that it approximates f well in L^1-norm. Using the fact that $\int k_\lambda = 1$, we compute that

$$\begin{aligned}
\|f - f * k_\lambda\|_1 \;&=\; \int |f(x) - (f * k_\lambda)(x)|\, dx \\
&=\; \int \left| f(x) \int k_\lambda(t)\, dt \;-\; \int f(x-t)\, k_\lambda(t)\, dt \right| dx \\
&\le\; \iint |f(x) - f(x-t)|\, |k_\lambda(t)|\, dt\, dx \\
&=\; \iint |f(x) - f(x-t)|\, |k_\lambda(t)|\, dx\, dt \\
&=\; \int |k_\lambda(t)| \int |f(x) - T_t f(x)|\, dx\, dt \\
&=\; \int |k_\lambda(t)|\, \|f - T_t f\|_1\, dt, \qquad\qquad (1.23)
\end{aligned}$$

where the interchange in the order of integration is permitted by Tonelli's Theorem, since the integrands are nonnegative. We want to show that the quantity in equation (1.23) is small when λ is large.

Choose any $\varepsilon > 0$. Since translation is strongly continuous on $L^1(\mathbb{R})$, there exists a $\delta > 0$ such that

$$|t| < \delta \quad\Longrightarrow\quad \|f - T_t f\|_1 < \varepsilon.$$

Also, by definition of approximate identity, we know that

$$K \;=\; \sup_{\lambda} \|k_\lambda\|_1 \;<\; \infty,$$

and that there exists some λ_0 such that

$$\lambda > \lambda_0 \quad\Longrightarrow\quad \int_{|t|\geq\delta} |k_\lambda(t)|\,dt < \varepsilon.$$

Therefore, for $\lambda > \lambda_0$ we can continue equation (1.23) as follows:

$$\begin{aligned}
(1.23) \;&=\; \int_{|t|<\delta} |k_\lambda(t)|\,\|f - T_t f\|_1\,dt \;+\; \int_{|t|\geq\delta} |k_\lambda(t)|\,\|f - T_t f\|_1\,dt \\
&\leq\; \int_{|t|<\delta} |k_\lambda(t)|\,\varepsilon\,dt \;+\; \int_{|t|\geq\delta} |k_\lambda(t)|\,\big(\|f\|_1 + \|T_t f\|_1\big)\,dt \\
&\leq\; \varepsilon \int |k_\lambda(t)| \;+\; 2\|f\|_1 \int_{|t|\geq\delta} |k_\lambda(t)|\,dt \\
&\leq\; \varepsilon K \;+\; 2\|f\|_1\,\varepsilon.
\end{aligned}$$

Thus $\|f - f * k_\lambda\|_1 \to 0$ as $\lambda \to \infty$. $\square$

To illustrate the convergence proved in the preceding theorem, consider the particular function $\chi = \chi_{[0,1]}$ and a particular approximate identity that will be of considerable use to us later. This is the *Fejér kernel* $\{w_\lambda\}_{\lambda>0}$, which is produced by dilating the *Fejér function* $w(x) = \big(\frac{\sin \pi x}{\pi x}\big)^2$ depicted in Figure 1.12. In Figure 1.10, we see the convolutions $\chi * w$, $\chi * w_5$, and $\chi * w_{25}$.

In addition to the convergence apparent in these figures, note that the convolved functions appear to be continuous, while χ is discontinuous. This is due to the smoothing effect of convolution, which was discussed in Section 1.3.7.

There are many, many variations on the theme of Theorem 1.53. To begin, since $L^p * L^1 \subseteq L^p$, we expect that we may be able to extend to other values of p, and indeed for p finite we have the following result.

Exercise 1.54. Let $\{k_\lambda\}_{\lambda>0}$ be an approximate identity. Prove that if $1 \leq p < \infty$, then

$$\forall\, f \in L^p(\mathbb{R}), \qquad \lim_{\lambda\to\infty} \|f - f * k_\lambda\|_p \;=\; 0.$$

That is, $f * k_\lambda \to f$ in L^p-norm as $\lambda \to \infty$.

As usual, we do need to be careful when $p = \infty$, and we consider this issue next.

1.5.3 Uniform Convergence

Now we turn to approximation by approximate identities in spaces such as $L^\infty(\mathbb{R})$, $C_0(\mathbb{R})$, and $C_b(\mathbb{R})$. Considering that the elements k_λ of an approximate identity are integrable functions, if f belongs to $L^\infty(\mathbb{R})$ then $f * k_\lambda$

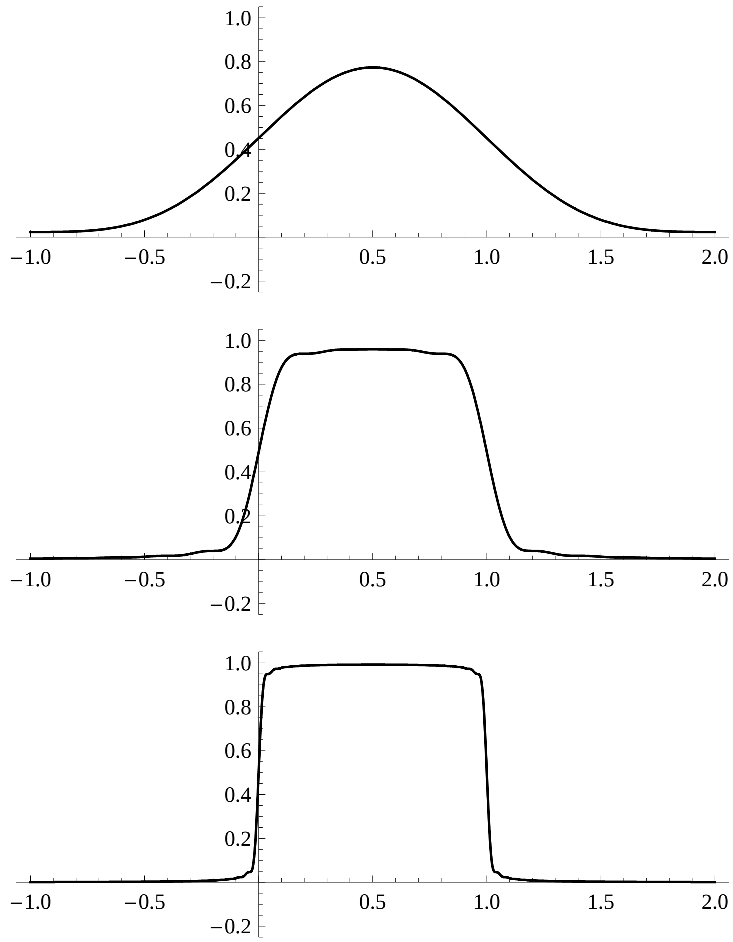

Fig. 1.10. Convolution with an approximate identity. Top: $\chi * w$. Middle: $\chi * w_5$. Bottom: $\chi * w_{25}$.

belongs to $C_b(\mathbb{R})$ by Exercise 1.34. Hence it is simply too much to expect that $f * k_\lambda$ will converge to f in L^∞-norm for all $f \in L^\infty(\mathbb{R})$. However, if we impose some smoothness on f, then we can obtain convergence in senses that are appropriate to f.

If we assume that f belongs to $C_0(\mathbb{R})$, then $f * k_\lambda$ also belongs to $C_0(\mathbb{R})$ by Theorem 1.31. By replacing $L^\infty(\mathbb{R})$ with $C_0(\mathbb{R})$, we therefore obtain for $p = \infty$ an exact analogue of the convergence result of Exercise 1.54.

Exercise 1.55. Let $\{k_\lambda\}_{\lambda>0}$ be an approximate identity. Prove that

$$\forall\, f \in C_0(\mathbb{R}), \qquad \lim_{\lambda\to\infty} \|f - f * k_\lambda\|_\infty \;=\; 0.$$

That is, $f * k_\lambda$ converges uniformly to f as $\lambda \to \infty$.

Suppose that we relax the decay hypothesis on f in Exercise 1.55, and assume only that $f \in C_b(\mathbb{R})$ instead of in $C_0(\mathbb{R})$. In this case we do have that $f * k_\lambda \in C_b(\mathbb{R})$ by Exercise 1.34, so we can hope that $f * k_\lambda$ will still converge to f. This is true, but now the convergence will be uniform only when restricted to compact sets.

Exercise 1.56. Let $\{k_\lambda\}_{\lambda>0}$ be an approximate identity. Prove that if $f \in C_b(\mathbb{R})$, then $f * k_\lambda$ converges uniformly to f on compact sets, i.e.,

$$\forall \text{compact } K \subseteq \mathbb{R}, \quad \lim_{\lambda\to\infty} \big\|(f - f * k_\lambda)\, \chi_K\big\|_\infty \;=\; 0.$$

Give an example showing that $f * k_\lambda$ need not converge uniformly to f on $\mathbb{R}$.

On the other hand, if we impose even a slight amount of "extra smoothness" on $f \in C_b(\mathbb{R})$, then we can restore the uniform convergence of $f*k_\lambda$ to f on the entire real line. This extra smoothness is quantified in terms of Hölder continuity, which is defined and discussed in Section B.10 of the Appendices.

Exercise 1.57. Let $\{k_\lambda\}_{\lambda>0}$ be an approximate identity. Prove that if $f \in C_b(\mathbb{R})$ is Hölder continuous for some exponent $0 < \alpha \le 1$, then $f*k_\lambda$ converges uniformly to f on $\mathbb{R}$, i.e.,

$$\lim_{\lambda\to\infty} \|f - f * k_\lambda\|_\infty \;=\; 0.$$

1.5.4 Pointwise Convergence

As we have seen, if $f \in L^p(\mathbb{R})$ with p finite, then $f * k_\lambda \to f$ in L^p-norm. If we impose some restrictions on k, then we can also show that $(f * k_\lambda)(x)$ converges to $f(x)$ at each Lebesgue point x of f. And since almost every point in $\mathbb{R}$ is a Lebesgue point of f, this says that $f * k_\lambda$ converges pointwise a.e. to f. For background on Lebesgue points, see Section B.9 of the Appendices.

Theorem 1.58. *Let k be a bounded, compactly supported function such that $\int k = 1$, and define $k_\lambda(x) = \lambda k(\lambda x)$. If $f \in L^1(\mathbb{R})$, then $(f * k_\lambda)(x) \to f(x)$ as $\lambda \to \infty$ for each point x in the Lebesgue set of f. In particular, $f * k_\lambda$ converges to f pointwise a.e.*

Proof. By hypothesis, $\operatorname{supp}(k) \subseteq [-R, R]$ for some R. If x is a Lebesgue point of f, then

$$\begin{aligned}
&\lim_{\lambda\to\infty} |f(x) - (f * k_\lambda)(x)| \\
&\qquad = \lim_{\lambda\to\infty} \left| f(x) \int k_\lambda(x-t)\, dt \;-\; \int f(t)\, k_\lambda(x-t)\, dt \right| \\
&\qquad \le \lim_{\lambda\to\infty} \lambda \int |f(x) - f(t)|\, |k(\lambda(x-t))|\, dt
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{\lambda\to\infty} \frac{2R\lambda}{2R} \int_{x-(R/\lambda)}^{x+(R/\lambda)} |f(x)-f(t)|\,|k(\lambda(x-t))|\,dt \\
&\le 2R\,\|k\|_\infty \lim_{h\to 0} \frac{1}{2h} \int_{x-h}^{x+h} |f(x)-f(t)|\,dt \\
&= 0,
\end{aligned}
$$

where the limit is zero by definition of Lebesgue point. Finally, since almost every x is a Lebesgue point of f, we conclude that $f * k_\lambda$ converges to f pointwise a.e. $\square$

The hypotheses on k in Theorem 1.58 can be relaxed. In particular, compact support of k is not required. For example, Stein and Weiss [SW71, p. 13] give a more intricate argument that shows that if $k \in L^1(\mathbb{R})$ satisfies $\int k = 1$ and there exists an even function $\phi \in L^1(\mathbb{R})$ that is decreasing and differentiable on $(0,\infty)$ that dominates k in the sense that $|k(x)| \le \phi(x)$ for all x, then $(f * k_\lambda)(x) \to f(x)$ for every x in the Lebesgue set of f.

1.5.5 Dense Sets of Nice Functions

Theorem B.60 in Appendix B gives a direct proof, based on Urysohn's Lemma, that the space $C_c(\mathbb{R})$ is dense in $L^1(\mathbb{R})$. It almost seems that we should be able to give a "simple" proof of this fact by using approximate identities and arguing as follows.

Choose any $f \in L^1(\mathbb{R})$. Then we can find a compactly supported $g \in L^1(\mathbb{R})$ that is close to f, e.g., if we take R large enough and set $g = f\,\chi_{[-R,R]}$ then we will have $\|f-g\|_1 < \varepsilon$. If we convolve g with an element k_λ of an approximate identity, then $g*k_\lambda$ will be close to g if λ is large enough, say $\|g - g*k_\lambda\|_1 < \varepsilon$. Further, if we choose our approximate identity so that k_λ is a "nice" function then $g*k_\lambda$ will inherit this "niceness" as well. For example, if $k_\lambda \in C_c(\mathbb{R})$ then $g * k_\lambda \in C_c(\mathbb{R})$, and so we have found an element of $C_c(\mathbb{R})$ that lies within 2ε of f in L^1-norm.

The flaw in this reasoning is that our proof in Theorem 1.53 that $g*k_\lambda \to g$ in L^1-norm relies on the fact that translation is a strongly continuous family of operators in $L^1(\mathbb{R})$. The proof of this strong continuity of translation is the content of Exercise 1.14. *However*, the proof of this exercise that we suggest in the solutions section requires us to *already* know that $C_c(\mathbb{R})$ is dense in $L^1(\mathbb{R})$. Hence the reasoning of the preceding paragraph is circular.

We could take a different approach, e.g., by first arguing that simple functions are dense in $L^1(\mathbb{R})$ and trying to proceed from there to show that $C_c(\mathbb{R})$ is dense. But it doesn't really matter, one way or the other we essentially have to "get our hands dirty" and show that some particular special subset is dense. The power of approximate identities really comes at the next step — once we know that one particular set is dense, we can use the spirit of the argument above (convolution with a "nice" approximate identity) to easily show that "much nicer" spaces are also dense. We even obtain results that almost seem too good to be true. For example, we will see that the space $C_c^\infty(\mathbb{R})$ consisting of *infinitely differentiable, compactly supported functions* is dense in $L^p(\mathbb{R})$ for every $1 \le p < \infty$! This fact is not just an abstract surprise, but will be of great utility to us throughout the remainder of this volume, especially when we turn to distribution theory in Chapter 3.

Exercise 1.59. (a) Show that $C_c^m(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ for each $m \ge 0$ and $1 \le p < \infty$, and is also dense in $C_0(\mathbb{R})$ in L^∞-norm.

(b) Show that $C_c^\infty(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ for each $1 \le p < \infty$, and is also dense in $C_0(\mathbb{R})$ in L^∞-norm.

After proving the Inversion Formula in Section 1.6, we will also be able to show that many spaces of functions with nice Fourier transforms are also dense. For example, in Problem 1.50 we will see that $\{f \in L^1(\mathbb{R}) : \widehat{f} \in C_c^\infty(\mathbb{R})\}$ is dense in $L^p(\mathbb{R})$ for $1 \le p < \infty$.

1.5.6 Urysohn's Lemma

Urysohn's Lemma is a general result that holds in a large class of topological spaces. One version of Urysohn's Lemma for the specific case of the real line is proved directly in Appendix A (see Theorem A.109). By using approximate identities, we now prove a much more refined C^∞-version of Urysohn's Lemma for subsets of the real line.

Theorem 1.60 (C^∞ Urysohn Lemma). *Let $K \subseteq \mathbb{R}$ be compact, and let $U \supseteq K$ be an open set. Then there exists $f \in C_c^\infty(\mathbb{R})$ such that $0 \le f \le 1$, $f = 1$ on K, and $\operatorname{supp}(f) \subseteq U$.*

Proof. Since K is compact and $\mathbb{R}\backslash U$ is closed, the distance between these sets is positive, i.e.,

$$d \;=\; \operatorname{dist}(K, \mathbb{R}\backslash U) \;=\; \inf\bigl\{|x-y| : x \in K, y \notin U\bigr\} \;>\; 0.$$

Let

$$V \;=\; \Bigl\{y \in \mathbb{R} : \operatorname{dist}(y, K) < \frac{d}{3}\Bigr\},$$

and let k be any function in $C_c^\infty(\mathbb{R})$ such that $k \ge 0$, $\int k = 1$, and $\operatorname{supp}(k) \subseteq [-\frac{d}{3}, \frac{d}{3}]$ (for example, dilate the function constructed in Exercise 1.37). Set $f = k * \chi_V$. Since k and χ_V are both compactly supported, their convolution is also compactly supported, and hence it follows from Corollary 1.36 that $f \in C_c^\infty(\mathbb{R})$. Since

$$f(x) \;=\; \int_V k(x-y)\, dy \;\le\; \int k \;=\; 1,$$

we have $0 \le f \le 1$ everywhere. Also, if $x \in K$ and $y \notin V$ then $|x-y| \ge \frac{d}{3}$ and so $k(x-y) = 0$. Therefore for $x \in K$ we have

$$f(x) = \int_V k(x-y)\,dy = \int k(x-y)\,dy = 1.$$

Similarly if $x \notin U$ then it follows that $f(x) = 0$. $\square$

1.5.7 Gibbs' Phenomenon

Gibbs' phenomenon refers to the behavior of the Fourier series of a periodic function near a jump discontinuity. Although this volume focuses on the Fourier transform on $\mathbb{R}$ rather than Fourier series on $\mathbb{T}$, there is an analogous phenomenon for the Fourier transform that we will discuss. For the formulation and proof of Gibbs' phenomenon on the torus, we refer to [DM72].

To illustrate this, let $H = \chi_{[0,\infty)}$ be the *Heaviside function*. Although H does not belong to $L^1(\mathbb{R})$, the important fact for this example is that H has a jump discontinuity at $x = 0$. Pointwise convergence of Fourier series corresponds to convolution with the Dirichlet kernel on the torus (see Section 2.2.2, and equation (2.7) in particular). The Dirichlet kernel on the real line is $\{d_\lambda\}_{\lambda>0}$, which is obtained by dilating the Dirichlet function

$$d(\xi) = \frac{\sin \xi}{\pi \xi}.$$

Since $d \notin L^1(\mathbb{R})$, the Dirichlet kernel does not form an approximate identity, but even so let us consider the pointwise behavior of $H * d_\lambda$ as $\lambda \to \infty$. Using the fact that, as an improper Riemann integral, $\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$ (see Problem B.18), we have

$$\begin{aligned}
(H * d_\lambda)(x) &= \int_0^\infty \frac{\sin \lambda(x-y)}{\pi(x-y)}\,dy \\
&= \int_{-\infty}^{\lambda x} \frac{\sin y}{\pi y}\,dy \\
&= \frac{1}{2} + \int_0^{\lambda x} \frac{\sin y}{\pi y}\,dy.
\end{aligned}$$

Since $\frac{\sin y}{\pi y} > 0$ for $0 < y < \pi$, $H * d_\lambda$ will increase on $(0, \pi/\lambda)$. Then it will decrease on $(\pi/\lambda, 2\pi/\lambda)$, then increase on $(2\pi/\lambda, 3\pi/\lambda)$ — but never back to the value it had at π/λ. Continuing in this way we see that $H * d_\lambda$ achieves its maximum at $x = \pi/\lambda$, and this maximum is

$$(H * d_\lambda)\Big(\frac{\pi}{\lambda}\Big) = \frac{1}{2} + \int_0^\pi \frac{\sin y}{\pi y}\,dy \approx 1.089\ldots.$$

Note that this maximum is *independent of* λ. Although $(H * d_\lambda)(x)$ converges pointwise to $H(x)$ as $\lambda \to \infty$ for all $x \neq 0$, this convergence is *not uniform*. In particular, the maximum distance between $(H * d_\lambda)(x)$ and $H(x)$ for $x > 0$ is a constant (approximately $0.089\ldots$) that is independent of λ, although its *location* at $x = \pi/\lambda$ decreases with λ. Figure 1.11 displays a plot of $H * d_\lambda$ for $\lambda = 16\pi$.

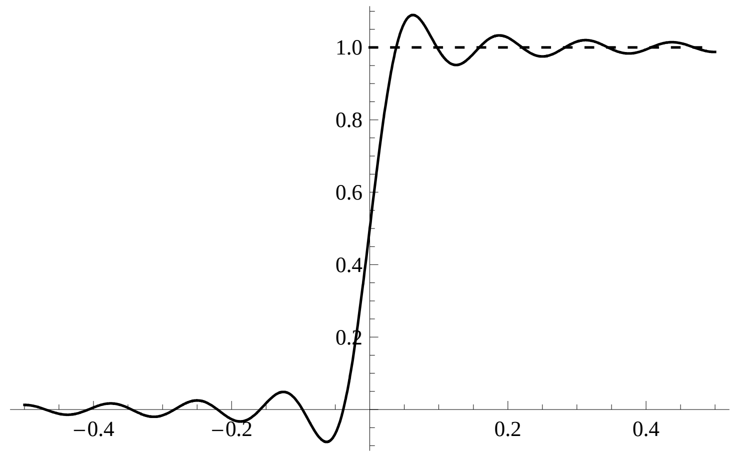

Fig. 1.11. Graph of $H * d_\lambda$ for $\lambda = 16\pi$.

1.5.8 Translation-Invariant Subspaces of $L^1(\mathbb{R})$

The following characterization of the closed, translation-invariant subspaces of $L^1(\mathbb{R})$ is due to Norbert Wiener (1894–1964); compare this to the characterization of the closed, translation-invariant subspaces of $L^2(\mathbb{R})$ given in Problem 3.10.

Definition 1.61 (Translation-Invariant Subspaces). We say that a subset J of $L^1(\mathbb{R})$ is *translation-invariant* if J is closed under all translations, i.e., if

$$f \in J,\ a \in \mathbb{R} \implies T_a f \in J.$$

Remark 1.62. A *shift-invariant* space V is one that is closed under *integer* translations, i.e., if $f \in V$ then $T_k f \in V$ for all integer k. These spaces play important roles in sampling theory and wavelet theory, see [Dau92].

Exercise 1.63. Let J be a closed subspace of $L^1(\mathbb{R})$, and prove that the following statements are equivalent.

(a) J is translation-invariant.

(b) J is an ideal in $L^1(\mathbb{R})$ under convolution.

Note that closedness is important — for example, $C_c(\mathbb{R})$ is a translation-invariant subspace of $L^1(\mathbb{R})$, but it is not an ideal with respect to convolution.

By using Exercise 1.63, we can give another characterization of the principal ideal $I(g) = \overline{g * L^1(\mathbb{R})}$ generated by a function $g \in L^1(\mathbb{R})$. By Exercise 1.39 and Problem 1.28, $I(g)$ is the smallest closed ideal in $L^1(\mathbb{R})$ that contains g.

Exercise 1.64. Given $g \in L^1(\mathbb{R})$, show that

$$I(g) \ = \ \overline{\operatorname{span}}\big(\{T_a g\}_{a\in\mathbb{R}}\big).$$

Thus, the smallest closed ideal that contains g equals the smallest closed subspace that contains all translates of g.

Recall that a set $S \subseteq L^1(\mathbb{R})$ is *complete* in $L^1(\mathbb{R})$ if $\operatorname{span}(S)$ is dense in $L^1(\mathbb{R})$ (see Definition A.76). It therefore follows from Exercise 1.64 that $\{T_a g\}_{a\in\mathbb{R}}$ is complete in $L^1(\mathbb{R})$ if and only if $I(g) = L^1(\mathbb{R})$. But when does this happen? The next exercise gives a necessary condition.

Exercise 1.65. Given $g \in L^1(\mathbb{R})$, show that

$$\{T_a g\}_{a\in\mathbb{R}} \text{ is complete in } L^1(\mathbb{R}) \quad \Longrightarrow \quad \widehat{g}(\xi) \neq 0 \text{ for all } \xi \in \mathbb{R}.$$

The converse of Exercise 1.65 is also true, but is a much deeper fact that is part of *Wiener's Tauberian Theorem*. In contrast, the analogous question in $L^2(\mathbb{R})$ is much simpler, see Problem 3.9.

Additional Problems

1.25. This problem constructs an example of an approximate identity $\{k_\lambda\}_{\lambda>0}$ where k_λ need not be a dilation of a single function k. For each $\lambda > 0$, let $k_\lambda \in L^1(\mathbb{R})$ be any function satisfying $k_\lambda \geq 0$, $\|k_\lambda\|_1 = \int k_\lambda = 1$, and $\operatorname{supp}(k_\lambda) \subseteq [-\frac{1}{\lambda}, \frac{1}{\lambda}]$. Show that $\{k_\lambda\}_{\lambda>0}$ is an approximate identity.

1.26. Show that if $\{k_\lambda\}_{\lambda>0}$ is any approximate identity then $\widehat{k_\lambda}(\xi) \to 1$ pointwise as $\lambda \to \infty$.

1.27. Assume $k \in L^1(\mathbb{R})$ is given and we set $r = \int k$ and $k_\lambda(x) = \lambda k(\lambda x)$. Prove that if $1 \leq p < \infty$, then for each $f \in L^p(\mathbb{R})$ we have that $f * k_\lambda \to rf$ in L^p-norm as $\lambda \to \infty$. Note that this includes the possibility that r may be complex or zero.

1.28. Fix $g \in L^1(\mathbb{R})$, and consider the ideals $g * L^1(\mathbb{R})$ and $I(g)$ introduced in Exercise 1.39. Show that we need not have $g \in g * L^1(\mathbb{R})$, but we always have $g \in I(g)$.

1.29. Let $k \in L^1(\mathbb{R})$ be any function such that $\int k = 1$ and $xk(x) \in L^1(\mathbb{R})$, and define an approximate identity by setting $k_\lambda(x) = \lambda k(\lambda x)$. Fix $1 \leq p < \infty$.

(a) Show that $\|Pk_\lambda\|_1 \to 0$ as $\lambda \to \infty$, where P is the position operator.

(b) Show that if $f \in L^p(\mathbb{R})$ then $f * Pk_\lambda \to 0$ in L^p-norm. Further, if we also have $xf(x) \in L^p(\mathbb{R})$, then $P(f * k_\lambda) \to Pf$ in L^p-norm.

1.30. This problem will illustrate one of the many different possible generalizations of the results of this section by considering particular *weighted L^p spaces*. Given $s \in \mathbb{R}$ let $v_s(x) = (1+|x|)^s$; we refer to v_s as a *polynomial weight* (since it has polynomial-like growth if $s \geq 0$, or decays like the reciprocal of a polynomial if $s \leq 0$). For this problem fix $1 \leq p < \infty$ and $s \in \mathbb{R}$. Then we define $L^p_s(\mathbb{R})$ to be the space of all measurable functions $f\colon \mathbb{R} \to \mathbb{C}$ such that

$$\|f\|_{p,s} \;=\; \|f\,v_s\|_p \;=\; \left(\int |f(x)|^p\,(1+|x|)^{ps}\,dx\right)^{1/p} \;<\; \infty.$$

$L^p_s(\mathbb{R})$ is a Banach space (see Problem C.11). Prove the following statements.

(a) If $s \ge 0$ then v_s is *submultiplicative*, i.e., $v_s(x+y) \le v_s(x)\,v_s(y)$ for $x,\, y \in \mathbb{R}$. If $s \le 0$ then v_s is v_{-s}*-moderate*, i.e., $v_s(x+y) \le v_s(x)\,v_{-s}(y)$ for $x,\, y \in \mathbb{R}$.

(b) For each $a \in \mathbb{R}$, the translation operator T_a is a continuous map of $L^p_s(\mathbb{R})$ into itself, with operator norm $\|T_a\|_{L^p_s \to L^p_s} \le v_{|s|}(a) = (1+|a|)^{|s|}$.

(c) Translation is strongly continuous on $L^p_s(\mathbb{R})$, i.e., for each $f \in L^p_s(\mathbb{R})$ we have $\lim_{a\to 0} \|f - T_a f\|_{p,s} = 0$.

(d) If $\{k_\lambda\}_{\lambda>0}$ is an approximate identity, then for each $f \in L^p_s(\mathbb{R})$ we have $\lim_{\lambda\to\infty} \|f - f * k_\lambda\|_{p,s} = 0$.

(e) $C_c^\infty(\mathbb{R})$ is dense in $L^p_s(\mathbb{R})$.

(f) Do parts (b)–(d) still hold if we replace v_s by an arbitrary weight $w\colon \mathbb{R} \to (0,\infty)$? What properties do we need w to possess?

1.6 The Inversion Formula

In this section we will prove Theorem 1.9, the Inversion Formula for the Fourier transform on $L^1(\mathbb{R})$, which states that if f and $\widehat{f}$ both belong to $L^1(\mathbb{R})$, then $f = \big(\widehat{f}\,\big)^{\vee} = \big(\check{f}\,\big)^{\wedge}$.

Unfortunately, the Inversion Formula does not apply to every function in $L^1(\mathbb{R})$. For example, by Exercise 1.7 we have that $\chi = \chi_{[-1/2,1/2]}$ belongs to $L^1(\mathbb{R})$ but $\widehat{\chi}(\xi) = d_\pi(\xi) = \frac{\sin \pi\xi}{\pi\xi}$ is not integrable. Another example is given in Problem 1.1: The one-sided exponential function $f(x) = e^{-x}\,\chi_{[0,\infty)}(x)$ is integrable, but its Fourier transform $\widehat{f}(\xi) = \frac{1}{2\pi i\xi+1}$ is not.

1.6.1 The Fejér Kernel

To prove the Inversion Formula, we will use the machinery of approximate identities that we developed in Section 1.5. Specifically, we will use a particular approximate identity, named after Lipót Fejér (1880–1959), that has some useful special properties.

Definition 1.66 (Fejér Kernel). The *Fejér function* is the square of the sinc function d_π:

$$w(x) \;=\; \left(\frac{\sin \pi x}{\pi x}\right)^2 \;=\; d_\pi(x)^2.$$

The *Fejér kernel* is $\{w_\lambda\}_{\lambda>0}$ where $w_\lambda(x) = \lambda w(\lambda x)$.