1.8. This problem provides an alternative proof to Theorem 1.17.

(a) Show that $\widehat{f} \in C_0(\mathbb{R})$ for every $f \in S = \operatorname{span}\{\chi_{[a,b]} : a < b \in \mathbb{R}\}$.

(b) Show that S is dense in $L^1(\mathbb{R})$ (see Exercise B.61), and use this to prove that $\widehat{f} \in C_0(\mathbb{R})$ for every $f \in L^1(\mathbb{R})$.

1.3 Convolution

Since $L^1(\mathbb{R})$ is a Banach space, we know that it has many useful properties. In particular the operations of addition and scalar multiplication are continuous. However, there are many other operations on $L^1(\mathbb{R})$ that we could consider. One natural operation is multiplication of functions, but unfortunately $L^1(\mathbb{R})$ is not closed under pointwise multiplication.

Exercise 1.18. Show that $f,\, g \in L^1(\mathbb{R})$ does not imply $fg \in L^1(\mathbb{R})$.

In this section we will define a different "multiplication-like" operation under which $L^1(\mathbb{R})$ is closed. This operation, convolution of functions, will be one of the most important tools in our further development of harmonic analysis. Therefore, in this section we set aside the Fourier transform for the moment, and concentrate on developing the machinery of convolution.

1.3.1 Some Notational Conventions

Before proceeding, there are some technical issues related to the definition of elements of $L^p(\mathbb{R})$ that we need to clarify (see Section B.6.2 for additional discussion of these issues).

The basic source of difficulty is that an element f of $L^p(\mathbb{R})$ is not a function but rather denotes an equivalence class of functions that are equal almost everywhere. Therefore we cannot speak of the "value of $f \in L^p(\mathbb{R})$ at a point $x \in \mathbb{R}$," and consequently concepts such as continuity or support do not apply in a literal sense to elements of $L^p(\mathbb{R})$. For example, the zero function 0 and the function $\chi_{\mathbb{Q}}$ both belong to the zero element of $L^p(\mathbb{R})$, which is the equivalence class of functions that are zero a.e., yet 0 is continuous and compactly supported while $\chi_{\mathbb{Q}}$ is discontinuous and its support is $\mathbb{R}$. Even so, it is often essential to consider smoothness or support properties of functions, and we therefore adopt the following conventions when discussing the smoothness or the support of elements of $L^p(\mathbb{R})$. More generally, these same issues and conventions apply to elements of

$$L^1_{\mathrm{loc}}(\mathbb{R}) \;=\; \bigl\{f \colon \mathbb{R} \to \mathbb{C} : f \cdot \chi_K \in L^1(\mathbb{R}) \text{ for every compact } K \subseteq \mathbb{R}\bigr\},$$

which is the *space of locally integrable functions* on $\mathbb{R}$. Note that $L^p(\mathbb{R}) \subseteq L^1_{\mathrm{loc}}(\mathbb{R})$ for every $1 \le p \le \infty$.

Notation 1.19 (Continuity for Elements of $L^1_{\text{loc}}(\mathbb{R})$). We will say that $f \in L^1_{\text{loc}}(\mathbb{R})$ is continuous if there is a representative of f that is continuous, i.e., there exists some continuous function f_0 such that f is the equivalence class of all functions that equal f_0 almost everywhere.

Conversely, if g is a continuous function such that $\int_K |g(x)|\,dx < \infty$ for every compact $K \subseteq \mathbb{R}$, then we write $g \in L^1_{\text{loc}}(\mathbb{R})$ with understanding that this means that the equivalence class of functions that equal g a.e. is an element of $L^1_{\text{loc}}(\mathbb{R})$. In this sense we write statements such as $C_c(\mathbb{R}) \subseteq L^p(\mathbb{R})$ even though $C_c(\mathbb{R})$ is a set of functions while $L^p(\mathbb{R})$ is a set of equivalence classes of functions.

Notation 1.20 (Support of Elements of $L^1_{\text{loc}}(\mathbb{R})$). We will say that $f \in L^1_{\text{loc}}(\mathbb{R})$ has compact support if there is a representative of f that has compact support. Thus f has compact support if there exists an $N > 0$ such that $f(x) = 0$ for a.e. $|x| > N$.

In many situations, this definition of compact support is all that we need, but in some circumstances it is important to discuss the support of $f \in L^1_{\text{loc}}(\mathbb{R})$ explicitly. We define the support of $f \in L^1_{\text{loc}}(\mathbb{R})$ to be

$$\operatorname{supp}(f) \;=\; \bigcap \bigl\{F \subseteq \mathbb{R} : F \text{ is closed and } f(x) = 0 \text{ for a.e. } x \notin F\bigr\}.$$

In particular, if F is a closed subset of $\mathbb{R}$, then

$$\operatorname{supp}(f) \subseteq F \iff f(x) = 0 \text{ for a.e. } x \notin F.$$

However, if T is a generic subset of $\mathbb{R}$, then the statements $\operatorname{supp}(f) \subseteq \mathrm{T}$ and $f(x) = 0$ for a.e. $x \notin \mathrm{T}$ need not be equivalent.

In the language of Chapter 3, we are taking the support of $f \in L^1_{\text{loc}}(\mathbb{R})$ to be the support of the distribution in $\mathcal{D}'(\mathbb{R})$ that is determined by f, see Section 4.5.

The reader should verify that if $f \in L^1_{\text{loc}}(\mathbb{R})$ is continuous (in the sense given in Notation 1.19), then the support of f in the sense of Notation 1.20 coincides with the usual definition of the support of f (see Definition A.19).

1.3.2 Definition and Basic Properties of Convolution

Now we can define convolution of functions.

Definition 1.21 (Convolution). Let $f\colon \mathbb{R} \to \mathbb{C}$ and $g\colon \mathbb{R} \to \mathbb{C}$ be Lebesgue measurable functions. Then the *convolution of f with g* is the function $f * g$ given by

$$(f * g)(x) \;=\; \int f(y)\, g(x-y)\, dy, \tag{1.8}$$

whenever this integral is well-defined.

For example, suppose that $1 \le p \le \infty$ and let p' be the dual index to p (see the review of notational conventions given in Section A.1). If $f \in L^p(\mathbb{R})$ and $g \in L^{p'}(\mathbb{R})$, then (as functions of y), $f(y)$ and $g(x-y)$ belong to dual spaces, and hence by Hölder's Inequality the integral defining $(f*g)(x)$ in equation (1.8) exists for every x, and furthermore is bounded as a function of x.

Exercise 1.22. Show that if $1 \le p \le \infty$, $f \in L^p(\mathbb{R})$, and $g \in L^{p'}(\mathbb{R})$, then $f*g \in L^\infty(\mathbb{R})$, and we have

$$\|f*g\|_\infty \le \|f\|_p \, \|g\|_{p'}. \tag{1.9}$$

We will improve on this exercise (in several ways) below. In particular, Exercise 1.22 does not give the only hypotheses on f and g which imply that $f*g$ exists — we will shortly see *Young's Inequality*, which is a powerful result that tells us that $f*g$ will belong to a particular Lebesgue space $L^r(\mathbb{R})$ whenever $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$, and we have the proper relationship among p, q, and r (specifically, $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$). Before turning to that general case, we prove the fundamental fact that $L^1(\mathbb{R})$ is closed under convolution, and that the Fourier transform interchanges convolution with multiplication.

Theorem 1.23. *If $f, g \in L^1(\mathbb{R})$ are given, then the following statements hold.*

(a) *$f(y)\,g(x-y)$ is Lebesgue measurable on $\mathbb{R}^2$.*

(b) *For almost every $x \in \mathbb{R}$, $f(y)\,g(x-y)$ is a measurable and integrable function of y, and hence $(f*g)(x)$ is defined for a.e. $x \in \mathbb{R}$.*

(c) *$f*g \in L^1(\mathbb{R})$, and*

$$\|f*g\|_1 \;\le\; \|f\|_1 \, \|g\|_1.$$

(d) *The Fourier transform of $f*g$ is the product of the Fourier transforms of f and g:*

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

Proof. (a) If we set $h(x,y) = f(x)$, then

$$h^{-1}(a,\infty) \;=\; \{(x,y) : h(x,y) > a\} \;=\; \{(x,y) : f(x) > a\} \;=\; f^{-1}(a,\infty) \times \mathbb{R},$$

which is a measurable subset of $\mathbb{R}^2$ since $f^{-1}(a,\infty)$ and $\mathbb{R}$ are measurable subsets of $\mathbb{R}$ (see Exercise B.15). Likewise $k(x,y) = g(y)$ is measurable. Since the product of measurable functions is measurable, we conclude that $F(x,y) = f(x)g(y)$ is measurable. Further, $T(x,y) = (y, x-y)$ is a linear transformation, so $H(x,y) = (F \circ T)(x,y) = F(y, x-y) = f(y)\,g(x-y)$ is measurable.

(b) Using the same notation as in part (a), we have

$$\begin{aligned} \iint |H(x,y)|\,dx\,dy \;&=\; \int\left(\int |g(x-y)|\,dx\right)|f(y)|\,dy \\ &=\; \int \|g\|_1\,|f(y)|\,dy \;=\; \|g\|_1\,\|f\|_1. \end{aligned}$$

Therefore $H(x,y) = f(y)\,g(x-y) \in L^1(\mathbb{R}^2)$, so Fubini's Theorem implies that the function $(f*g)(x) = \int f(y)\,g(x-y)\,dy$ exists for almost every x and is an integrable function of x.

(c) Using part (b), we have

$$\|f*g\|_1 \;=\; \int |(f*g)(x)|\,dx \;\le\; \iint |f(y)\,g(x-y)|\,dy\,dx \;=\; \|f\|_1\,\|g\|_1.$$

(d) Fubini's Theorem (exercise: justify its use) allows us to interchange integrals in the following calculation:

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

In the proof of Theorem 1.23, we carefully addressed the measurability of $f*g$. We will usually take issues of measurability for granted from now on, but it is a good idea for the reader to consider wherever appropriate why the measurability of the functions we encounter is ensured.

Exercise 1.24. Establish the following basic properties of convolution. Given $f,\, g,\, h \in L^1(\mathbb{R})$, prove the following.

(a) Commutativity: $f*g = g*f$.

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

(c) Distributive laws: $f*(g+h) = f*g + f*h$.

(d) Commutativity with translations: $f*(T_a g) = (T_a f)*g = T_a(f*g)$ for $a \in \mathbb{R}$.

(e) Behavior under involution: $(f*g)^\sim = \widetilde{f}*\widetilde{g}$.

1.3.3 Young's Inequality

As we have seen, $L^1(\mathbb{R})$ is closed under convolution, which we write in short as

$$L^1(\mathbb{R}) * L^1(\mathbb{R}) \;\subseteq\; L^1(\mathbb{R}).$$

It is not true that $L^p(\mathbb{R})$ is closed under convolution for $p > 1$. Instead we have the following fundamental result, known as *Young's Inequality* (although, since Young proved many inequalities, it may be advisable to refer to this as *Young's Convolution Inequality*).

Exercise 1.25 (Young's Inequality). Prove the following statements.

(a) If $1 \le p \le \infty$, then $L^p(\mathbb{R}) * L^1(\mathbb{R}) \subseteq L^p(\mathbb{R})$, and we have

$$\forall\, f \in L^p(\mathbb{R}), \quad \forall\, g \in L^1(\mathbb{R}), \quad \|f * g\|_p \;\le\; \|f\|_p\, \|g\|_1. \tag{1.10}$$

(b) If $1 \le p, q \le \infty$ and r satisfies $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, then $L^p(\mathbb{R}) * L^q(\mathbb{R}) \subseteq L^r(\mathbb{R})$, and we have

$$\forall\, f \in L^p(\mathbb{R}), \quad \forall\, g \in L^q(\mathbb{R}), \quad \|f * g\|_r \;\le\; \|f\|_p\, \|g\|_q. \tag{1.11}$$

Of course, statement (a) in Young's Inequality is a special case of statement (b), but it is so useful that it is worth stating separately. It is also instructive on first try to attempt to prove statement (a) rather than the more general statement (b) in order to see the appropriate technique needed. There are many ways to prove Young's Inequality, e.g., via Hölder's Inequality or Minkowski's Integral Inequality. See also Exercise 3.21, as well as Exercise C.23 in the Appendices.

Remark 1.26. If $q = p'$ (in which case $r = \infty$), then Young's Inequality tells us that the convolution of $f \in L^p(\mathbb{R})$ with $g \in L^{p'}(\mathbb{R})$ belongs to $L^\infty(\mathbb{R})$. In fact, it follows from our later Exercise 1.34 that $f * g$ is continuous in this case, and therefore $(f * g)(x)$ is defined for *every* x. However, for general values of p, q, r satisfying the hypotheses of Young's Inequality, we are only able to conclude that $f * g \in L^r(\mathbb{R})$, and hence we usually only have that $f * g$ is defined pointwise almost everywhere.

The inequalities given in Exercise 1.25 are in the form that we will most often need in practice, but it is very interesting to note that the implicit constant 1 on the right-hand side of equation (1.11) is not the optimal constant in general. Instead, if we define the *Babenko–Beckner constant* A_p by

$$A_p \;=\; \left(\frac{p^{1/p}}{p'^{1/p'}}\right)^{1/2}, \tag{1.12}$$

where we take $A_1 = A_\infty = 1$, then the optimal version of equation (1.11) is

$$\forall\, f \in L^p(\mathbb{R}), \quad \forall\, g \in L^q(\mathbb{R}), \quad \|f * g\|_r \;\le\; (A_p A_q A_{r'})\, \|f\|_p\, \|g\|_p, \tag{1.13}$$

and the constant $A_p A_q A_{r'}$ is typically not 1. The proof that $A_p A_q A_{r'}$ is the best constant in equation (1.13) is due to Beckner [Bec75] and Brascamp and Lieb [BrL76]. The Babenko–Beckner constant will make an appearance again when we consider the Hausdorff–Young Theorem in Chapter 2 (see Theorem 3.22).

1.3.4 Convolution as Filtering; Lack of an Identity

Theorem 1.23 gives us another way to view filtering (discussed in Section 1.1.2). Given $f \in L^1(\mathbb{R})$, we filter f by modifying its frequency content. That is, we create a new function h from f whose Fourier transform is

$$\widehat{h}(\xi) \;=\; \widehat{f}(\xi)\,\widehat{g}(\xi).$$

The Fourier transform of the function g tells us how to modify the frequency content of f. Assuming that the Inversion Theorem applies, we can recover h by the formula

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

which is a superposition of the "pure tones" $e^{2\pi i\xi x}$ with the modified amplitudes $\widehat{f}(\xi)\,\widehat{g}(\xi)$. Assuming that $g \in L^1(\mathbb{R})$, Theorem 1.23 tells us that we can also obtain h by convolution: we have $h = f * g$. *Filtering is convolution.*

Obviously, there are many details that we are glossing over by assuming all of the formulas are applicable. Some of these we will address later, e.g., is it true that a function $h \in L^1(\mathbb{R})$ is uniquely determined by its Fourier transform $\widehat{h}$? (Yes, we will show that $f \mapsto \widehat{f}$ is an injective map of $L^1(\mathbb{R})$ into $C_0(\mathbb{R})$, see Theorem 1.73.) Others we will leave for a course on digital signal processing, e.g., how do results for $L^1(\mathbb{R})$ relate to the processing of real-life digital signals whose domain is $\{1,\dots,n\}$ instead of $\mathbb{R}$? In any case, keeping our attention on the real line, let us ask one interesting question. If our goal is to filter f, then one of the possible filterings should be the identity operation, i.e., do nothing to the frequency content of f. Is there a $g \in L^1(\mathbb{R})$ such that $f \mapsto f * g$ is the identity operation on $L^1(\mathbb{R})$?

Exercise 1.27. Suppose that there existed a function $\delta \in L^1(\mathbb{R})$ such that

$$\forall\, f \in L^1(\mathbb{R}), \quad f * \delta \;=\; f.$$

Show that δ would satisfy $\widehat{\delta}(\xi) = 1$ for all ξ, which contradicts the Riemann–Lebesgue Lemma.

Consequently, *there is no identity element for convolution in* $L^1(\mathbb{R})$. This is problematic, and we will have several alternative ways of addressing this problem. In Section 1.5, we will construct functions which "approximate" an identity for convolution as closely as we like, according to a variety of meanings of approximation. In Chapter 3, we will create a distribution (or "generalized function") δ that is not itself a function but instead acts on functions and is an identity with respect to convolution. In Chapter 4, we will see that this distribution δ can also be regarded as a bounded measure on the real line. Thus, while there is no *function* δ that is an identity for convolution, we do have an identity δ in several generalized senses.

1.3.5 Convolution as Averaging; Introduction to Approximate Identities

Convolution can also be regarded as a kind of weighted averaging operator. For example, let us consider

$$\chi_T = \frac{1}{2T}\chi_{[-T,T]}.$$

Then given $f \in L^1(\mathbb{R})$, we have that

$$(f * \chi_T)(x) = \int f(y)\, \chi_T(x-y)\, dy = \frac{1}{2T}\int_{x-T}^{x+T} f(y)\, dy = \mathrm{Avg}_T f(x),$$

where $\mathrm{Avg}_T f(x)$ is the average of f on the interval $[x-T, x+T]$ (see Figure 1.6).

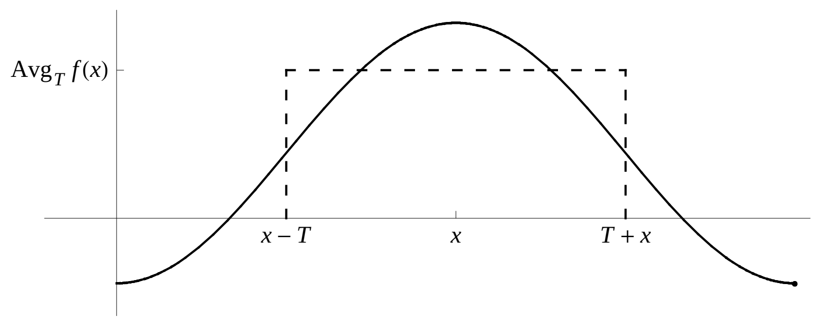

Fig. 1.6. The area of the dashed box equals $\int_{x-T}^{x+T} f(y)\, dy$, which is the area under the graph of f between $x-T$ and $x+T$.

For a general function g, the mapping $f \mapsto f * g$ can be regarded as a kind of weighted averaging of f, with g weighting some parts of the real line more than others. There is one technical point to observe in this viewpoint: while the function $\chi_T(x) = \frac{1}{2T}\chi_{[-T,T]}$ used in the preceding illustration is even, in general this will not be the case. Instead, in thinking of convolution as a weighted averaging, it is perhaps better to set $g^*(x) = g(-x)$ and write

$$(f * g)(x) = \int f(y)\, g^*(y-x)\, dy = \mathrm{Avg}_{g*} f(x),$$

the average of f around the point x corresponding to the weighting of the real line by $g^*(x) = g(-x)$. Alternatively, since convolution is commutative, we can equally view it as an averaging of g using the weighting corresponding to $f^*(x) = f(-x)$.

Looking ahead to Section 1.5, let us consider what happens to the convolution $f * \chi_T = \mathrm{Avg}_T f$ as $T \to 0$. The function $\chi_T = \frac{1}{2T}\chi_{[-T,T]}$ becomes a taller and taller "spike" centered at the origin, with the height of the spike being chosen so that the integral of χ_T is always 1. Intuitively, averaging over smaller and smaller intervals should give values $(f*\chi_T)(x)$ that are closer and closer to the original value $f(x)$. This intuition is made precise in Lebesgue's Differentiation Theorem (Theorem B.87), which implies that if $f \in L^1(\mathbb{R})$ then for almost every x (including all those in the Lebesgue set of f) we will have

$$f(x) \;=\; \lim_{T\to 0}\,(f*\chi_T)(x) \;=\; \lim_{T\to 0}\,\mathrm{Avg}_T f(x).$$

Thus $f \approx f*\chi_T$ when T is small. In this sense, while there is no identity element for convolution in $L^1(\mathbb{R})$, the function χ_T is *approximately* an identity for convolution, and the approximation becomes better and better the smaller T becomes.

Moreover, a similar phenomenon happens for the more general averaging operators $f \mapsto f*g = \mathrm{Avg}_{g*}f$. We can take any particular function $g \in L^1(\mathbb{R})$ and dilate it so that it becomes more and more compressed towards the origin, yet always keeping the total integral the same, by setting

$$g_\lambda(x) \;=\; \lambda g(\lambda x), \quad \lambda > 0,$$

as is done in Notation 1.5. Compressing g towards the origin corresponds to letting λ increase towards infinity (as opposed to $T \to 0$ in the discussion of χ_T above). Even if g is not compactly supported, it becomes more and more "spike-like" as λ increases (see the illustration in Figure 1.7).

If it is the case that $\int g = 1$ (so $\int g_\lambda = 1$ also), then we will see in Section 1.5 that, for any $f \in L^1(\mathbb{R})$, the convolution $f*g_\lambda$ converges to f in L^1-norm (and possibly in other senses as well, depending on properties of f and g). The family $\{g_\lambda\}_{\lambda>0}$ is an example of what we will call an *approximate identity* in Section 1.5.

From this discussion we can see at least an intuitive reason why there is no identity function for convolution in $L^1(\mathbb{R})$. Consider the functions g_λ, each an integrable function with integral 1 that become more and more spike-like as λ increases. Suppose that we could let $\lambda \to \infty$ and obtain in the limit an integrable function δ that, like each function g_λ, has integral 1, but is indeed a spike supported entirely at the origin. Then we would hopefully have that $f*\delta = \lim_{\lambda\to\infty} f*g_\lambda = f$, and so δ would be an identity function for convolution. And indeed, it is not uncommon to see informal wording similar to the following.

> "Let δ be the function on $\mathbb{R}$ that has the property that $\delta(x) = 0$ for all $x \neq 0$ and $\int \delta(x)\,dx = 1$. Then
>
> $$\int f(y)\,\delta(x-y)\,dy \;=\; f(x).\text{"} \tag{1.14}$$

However, there is *no such function* δ. Any function that is zero for all $x \neq 0$ is zero almost everywhere, and hence is the zero element of $L^1(\mathbb{R})$. If $\delta(x) = 0$

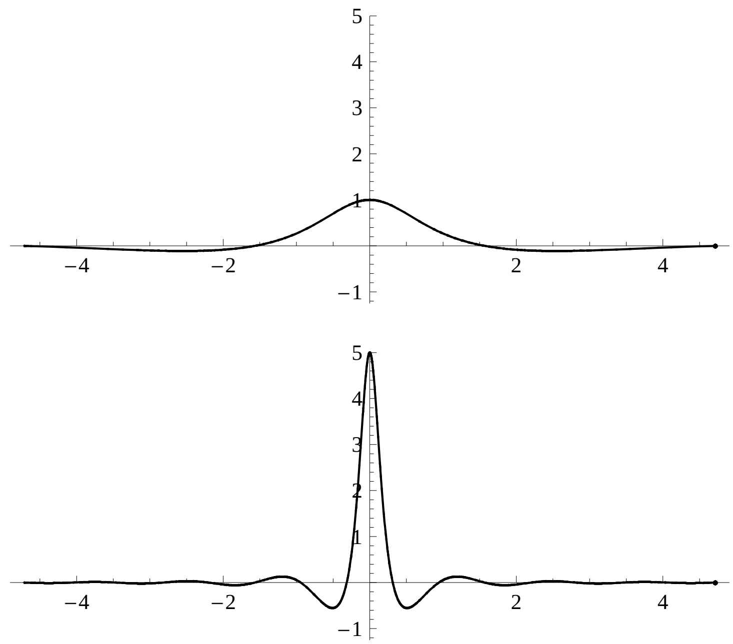

Fig. 1.7. Top: The function $g(x) = \cos(x)/(1+x^2)$. Bottom: The dilated function $g_5(x) = 5g(5x)$.

for $x \neq 0$, then the Lebesgue integral of δ is $\int \delta(x)\,dx = 0$, not 1, *even if* we define $\delta(0) = \infty$. Thus $f * \delta = 0$, not f.

We cannot construct a *function* δ that has the property that $f*\delta = f$ for all $f \in L^1(\mathbb{R})$. However, we can construct families $\{g_\lambda\}_{\lambda>0}$ that have the property that $f * g_\lambda$ converges to f in various senses, and these are the approximate identities of Section 1.5. We can also construct objects that are not functions but which are identities for convolution — we will see the δ-*distribution* in Chapter 3 and the δ-*measure* in Chapter 4. In effect, the integral appearing in equation (1.14) is not a Lebesgue integral but rather is simply a shorthand for something else, namely, the action of the distribution or measure δ on the function f.

1.3.6 Convolution as an Inner Product

It is often useful to write a convolution in one of the following forms (the involution $\widetilde{g}(x) = \overline{g(-x)}$ was introduced in Definition 1.12):

$$\begin{aligned}(f*g)(x) &= \int f(y)\,g(x-y)\,dy \\ &= \int f(y)\,\overline{\widetilde{g}(y-x)}\,dy \\ &= \int f(y)\,\overline{T_x\widetilde{g}(y)}\,dy \;=\; \langle f, T_x\widetilde{g}\,\rangle. \end{aligned} \tag{1.15}$$

Thus we can view the convolution of f with g at the point x as the inner product of f with the function $\widetilde{g}$ translated by x.

Notation 1.28. In equation (1.15), we have used the notation $\langle\cdot,\cdot\rangle$, which in the context of functions usually denotes the inner product on $L^2(\mathbb{R})$. However, neither f nor $T_x\widetilde{g}$ need belong to $L^2(\mathbb{R})$, so we are certainly taking some poetic license in speaking of $\langle f, T_x\widetilde{g}\,\rangle$ as an inner product of f and $T_x\widetilde{g}$. We do this because in this volume we so often encounter integrals of the form $\int f(x)\,\overline{g(x)}\,dx$ and direct generalizations of these integrals that it is extremely convenient for us to retain the notation $\langle f, g\rangle$ for such an integral whenever it makes sense. Specifically, if f and g are any measurable functions on $\mathbb{R}$, then we will write

$$\langle f, g\rangle \;=\; \int f(x)\,\overline{g(x)}\,dx$$

whenever this integral exists. In other language, $\langle\cdot,\cdot\rangle$ is a *sesquilinear form* (linear in the first variable, antilinear in the second) that extends the inner product on $L^2(\mathbb{R})$. Although an abuse of terminology, we will often refer to $\langle f, g\rangle$ as the *inner product of f with g* even when f and g are not in $L^2(\mathbb{R})$ or another Hilbert space. A more detailed discussion of this notation in the context of dual spaces is given in Appendix C (see especially Section C.5).

Note that in the calculation in equation (1.15), all that we know is that f and $T_x\widetilde{g}$ each belong to $L^1(\mathbb{R})$. Since the product of L^1 functions does not belong to L^1 in general, the integral appearing in equation (1.15) is not going to exist for every f, g, and x. Yet Theorem 1.23 implies the following perhaps unexpected fact.

Exercise 1.29. Show that

$$f, g \in L^1(\mathbb{R}) \quad\Longrightarrow\quad f\cdot\overline{T_x\widetilde{g}} \in L^1(\mathbb{R}) \text{ for a.e. } x.$$

Thus (thanks to Fubini and his theorem), even if we only assume that f and g are integrable, the "inner product" $(f*g)(x) = \langle f, T_x\widetilde{g}\,\rangle$ exists for almost every x.

1.3.7 Convolution and Smoothing

Since convolution is a type of averaging, it tends to be a smoothing operation. Generally speaking, a convolution $f*g$ inherits the "best" properties of both f and g. The following theorems and exercises will give several illustrations of this. We begin with an easy but very useful exercise.

Exercise 1.30. Show that

$$f \in C_c(\mathbb{R}),\ g \in C_c(\mathbb{R}) \implies f * g \in C_c(\mathbb{R}),$$

and that in this case we have

$$\operatorname{supp}(f*g) \subseteq \operatorname{supp}(f) + \operatorname{supp}(g) = \{x + y : x \in \operatorname{supp}(f),\ y \in \operatorname{supp}(g)\}.$$

Next we see an example of how a convolution $f*g$ can inherit smoothness from either f or g. This proof of this result uses a standard "extension by density" argument, which is a very useful technique for solving many of the exercises in this and other sections.

Theorem 1.31. *We have that*

$$f \in L^1(\mathbb{R}),\ g \in C_0(\mathbb{R}) \implies f * g \in C_0(\mathbb{R}).$$

Proof. Note that if $f \in L^1(\mathbb{R})$ and $g \in C_0(\mathbb{R})$, then Exercise 1.22 implies that $f*g$ exists and is bounded. Also, since $g \in C_0(\mathbb{R})$, we know that it is uniformly continuous. Therefore, for x, $h \in \mathbb{R}$ we have

$$\begin{aligned}
&\bigl|(f*g)(x) - (f*g)(x-h)\bigr| \\
&\qquad = \left|\int f(y)\, g(x-y)\, dy \; - \int f(y)\, g(x-h-y)\, dy\right| \\
&\qquad \le \int |f(y)|\, |g(x-y) - g(x-h-y)|\, dy \\
&\qquad \le \left(\sup_{u \in \mathbb{R}} |g(u) - g(u-h)|\right) \int |f(y)|\; dy \\
&\qquad = \|g - T_h g\|_\infty \, \|f\|_1 \;\to\; 0 \quad \text{as } h \to 0,
\end{aligned}$$

where the convergence follows from the fact that g is uniformly continuous. Thus $f*g \in C_b(\mathbb{R})$, and in fact $f*g$ is uniformly continuous. Actually, we can make this argument much more succinct by making use of the fact that convolution commutes with translation (Exercise 1.24). We need only write:

$$\begin{aligned}
\|f*g - T_h(f*g)\|_\infty &= \|f*g - f*(T_h g)\|_\infty \\
&= \|f*(g - T_h g)\|_\infty \\
&\le \|f\|_1 \, \|g - T_h g\|_\infty \;\to\; 0 \quad \text{as } h \to 0.
\end{aligned}$$

To show that $f*g \in C_0(\mathbb{R})$, consider first the case where $g \in C_c(\mathbb{R})$. Then $\operatorname{supp}(g) \subseteq [-N, N]$ for some $N > 0$. Hence

$$\begin{aligned}
|(f*g)(x)| &\le \int_{x-N}^{x+N} |f(y)|\, |g(x-y)|\, dy \\
&\le \|g\|_\infty \int_{x-N}^{x+N} |f(y)|\, dy \;\to\; 0 \quad \text{as } |x| \to \infty.
\end{aligned}$$

This shows that $f * g \in C_0(\mathbb{R})$ whenever $g \in C_c(\mathbb{R})$.

Now we extend by density to all of $C_0(\mathbb{R})$. Choose an arbitrary $g \in C_0(\mathbb{R})$. Since $C_c(\mathbb{R})$ is dense in $C_0(\mathbb{R})$, we can find functions $g_n \in C_c(\mathbb{R})$ such that $g_n \to g$ in L^∞-norm. Then $f * g_n \in C_0(\mathbb{R})$ for every n, and we have by equation (1.9) that

$$\|f * g - f * g_n\|_\infty \ \le\ \|f\|_1 \, \|g - g_n\|_\infty \ \to\ 0 \quad \text{ as } n \to \infty.$$

Thus $f * g_n \to f * g$ in L^∞-norm. Since $f * g_n \in C_0(\mathbb{R})$ for every n and since $C_0(\mathbb{R})$ is a closed subspace of $L^\infty(\mathbb{R})$, we conclude that $f * g \in C_0(\mathbb{R})$. □

Remark 1.32. It is important to observe that the function $f \in L^1(\mathbb{R})$ in the statement of Theorem 1.31 is really an equivalence class of functions that are equal almost everywhere. In this sense f is only defined a.e., yet the convolution $f * g$ is a continuous function that is defined everywhere. In particular, changing f on a set of measure zero has no effect on values $(f * g)(x) = \int f(y)\, g(x-y)\, dy$.

There are other ways to go about proving Theorem 1.31. For example, the next exercise suggests a slightly different way of giving an extension by density argument.

Exercise 1.33. Let $f_n,\, g_n \in C_c(\mathbb{R})$ be such that $f_n \to f$ in L^1-norm while $g_n \to g$ in L^∞-norm, and show that $f_n * g_n \to f * g$ in L^∞-norm. Since $f_n * g_n \in C_c(\mathbb{R})$ by Exercise 1.30, and since $C_0(\mathbb{R})$ is closed in L^∞-norm, it follows that $f * g \in C_0(\mathbb{R})$.

The following exercise extends the pairing (L^1, C_0) considered in Theorem 1.31 to pairings $(L^p, L^{p'})$ for each $1 < p < \infty$ (and in so doing improves on Exercise 1.22). A weaker conclusion also holds for the pairing (L^1, L^∞). An extension by density argument similar to that of either Theorem 1.31 or Exercise 1.33 is useful in proving this next exercise.

Exercise 1.34. (a) Show that if $1 < p < \infty$, then

$$f \in L^p(\mathbb{R}),\ g \in L^{p'}(\mathbb{R}) \quad\Longrightarrow\quad f * g \in C_0(\mathbb{R}).$$

(b) For the cases $p = 1$, $p' = \infty$ or $p = \infty$, $p' = 1$, show that

$$f \in L^1(\mathbb{R}),\ g \in L^\infty(\mathbb{R}) \quad\Longrightarrow\quad f * g \in C_b(\mathbb{R}),$$

and that $f * g$ is uniformly continuous. Show further that if g is compactly supported (see Notation 1.20), then $f * g \in C_0(\mathbb{R})$. However, give an example that shows that if g is not compactly supported then we need not have $f * g \in C_0(\mathbb{R})$, even if f is compactly supported.

1.3.8 Convolution and Differentiation

Not only is convolution well-behaved with respect to continuity, but we can extend to higher derivatives.

Exercise 1.35. Given $1 \le p < \infty$ and $m \ge 0$, show that

$$f \in L^p(\mathbb{R}),\ g \in C_c^m(\mathbb{R}) \implies f * g \in C_0^m(\mathbb{R}),$$

and

$$f \in L^\infty(\mathbb{R}),\ g \in C_c^m(\mathbb{R}) \implies f * g \in C_b^m(\mathbb{R}).$$

Further, writing $D^j g = g^{(j)}$ for the jth derivative, show that differentiation commutes with convolution, i.e.,

$$D^j(f * g) = f * D^j g, \quad j = 0, \dots, m.$$

In particular, for any $1 \le p \le \infty$, if $f \in L^p(\mathbb{R})$ is compactly supported and $g \in C_c^m(\mathbb{R})$, then $f * g \in C_c^m(\mathbb{R})$. This gives us an easy mechanism for generating new elements of $C_c^m(\mathbb{R})$ given any one particular element g. Moreover, if g is infinitely differentiable, then we can apply Exercise 1.35 for every m, and as a consequence obtain the following corollary.

Corollary 1.36. *If $1 \le p < \infty$, then*

$$f \in L^p(\mathbb{R}),\ g \in C_c^\infty(\mathbb{R}) \implies f * g \in C_0^\infty(\mathbb{R}),$$

and

$$f \in L^\infty(\mathbb{R}),\ g \in C_c^\infty(\mathbb{R}) \implies f * g \in C_b^\infty(\mathbb{R}).$$

Moreover, in either case, if f is also compactly supported then $f * g \in C_c^\infty(\mathbb{R})$.

So from one function in $C_c^\infty(\mathbb{R})$ we can generate many others. But this begs the question: *Are there* any functions that are both compactly supported and infinitely differentiable? It is not at all obvious at first glance whether there exist *any* functions with such extreme properties, but the following exercise constructs some examples (see also Problem 1.16).

Exercise 1.37. Define $f(x) = e^{-1/x^2}\,\chi_{(0,\infty)}(x)$.

(a) Show that for every $n \in \mathbb{N}$, there exists a polynomial p_n of degree $3n$ such that
$$f^{(n)}(x) = p_n(x^{-1})\,e^{-x^{-2}}\,\chi_{(0,\infty)}(x).$$
Conclude that $f \in C_b^\infty(\mathbb{R})$, and, in particular, $f^{(n)}(0) = 0$ for every $n \ge 0$. Note that $\operatorname{supp}(f) = [0, \infty)$.

(b) Show that if $a < b$, then $g(x) = f(x-a)\,f(b-x)$ belongs to $C_c^\infty(\mathbb{R})$ and satisfies $\operatorname{supp}(g) = [a, b]$ and $g(x) > 0$ on (a, b).

1.3.9 Convolution and Banach Algebras

The fact that $L^1(\mathbb{R})$ is a Banach space that is closed under convolution and satisfies $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$ says that $L^1(\mathbb{R})$ is an example of a *Banach algebra* (see Definition C.28). Since convolution is commutative, it is a commutative Banach algebra, but by Exercise 1.27 it is a Banach algebra without an identity: There is no function $\delta \in L^1(\mathbb{R})$ that satisfies $f * \delta = f$ for all $f \in L^1(\mathbb{R})$. We will explore some aspects of the Banach algebra structure of $L^1(\mathbb{R})$ and its relatives next.

First we create another Banach algebra that isometrically inherits its structure from $L^1(\mathbb{R})$ via the Fourier transform.

Exercise 1.38. Define

$$A(\mathbb{R}) \;=\; \{\widehat{f} : f \in L^1(\mathbb{R})\},$$

and set

$$\|\widehat{f}\|_A \;=\; \|f\|_1, \qquad f \in L^1(\mathbb{R}).$$

If we assume that the Fourier transform is injective on $L^1(\mathbb{R})$, which we prove later in Theorem 1.73, then $\|\cdot\|_A$ is well-defined. Given this, prove that $\|\cdot\|_A$ is a norm and that $A(\mathbb{R})$ is a Banach space with respect to this norm. Prove also that $A(\mathbb{R})$ is a commutative Banach algebra without identity with respect to the operation of pointwise products of functions.

The space $A(\mathbb{R})$ is known by a variety of names, including both the *Fourier algebra* and the *Wiener algebra*, although it should be noted that both of these terms are sometimes used to refer to other spaces.

Our definition of $A(\mathbb{R})$ is *implicit* — it is the set of all Fourier transforms of L^1 functions, and in fact is isometrically isomorphic to $L^1(\mathbb{R})$. However, we can ask whether there is an *explicit* description of $A(\mathbb{R})$ — is it possible to say that $F \in A(\mathbb{R})$ based directly on properties of F? For example, we know that $A(\mathbb{R}) \subseteq C_0(\mathbb{R})$ by the Riemann–Lebesgue Lemma (Theorem 1.17), so continuity and decay at infinity of F are necessary conditions for F to belong to $A(\mathbb{R})$, but are there further conditions that we can place on F that are both necessary and sufficient for membership in $A(\mathbb{R})$? No such explicit characterization is known. On the other hand, we will see later both implicit and explicit proofs that $A(\mathbb{R})$ is a *dense* but *proper* subset of $C_0(\mathbb{R})$, see Exercises 1.79, 1.80, and 1.81.

A subspace I of $L^1(\mathbb{R})$ is an *ideal* in $L^1(\mathbb{R})$ if

$$f \in L^1(\mathbb{R}),\ g \in I \implies f * g \in I.$$

Ideals play an important role in any Banach algebra (see Section C.4).

Exercise 1.39. Suppose that if $g \in L^1(\mathbb{R})$.

(a) Show that $g * L^1(\mathbb{R}) = \{g * f : f \in L^1(\mathbb{R})\}$ is an ideal in $L^1(\mathbb{R})$, called the *ideal generated by* g. Give an example that shows that g need not belong to $g * L^1(\mathbb{R})$.

(b) Show that

$$I(g) \;=\; \overline{g * L^1(\mathbb{R})}$$

is a closed ideal in $L^1(\mathbb{R})$, called the *closed ideal generated by* g. Problem 1.28 will later show that g always belongs to $I(g)$. Assuming this fact for now, prove that $I(g)$ is the *smallest* closed ideal that contains g.

An ideal of the form $I(g)$ is also called a *principal ideal* in $L^1(\mathbb{R})$. Is every closed ideal in $L^1(\mathbb{R})$ a principal ideal? Atzmon [Atz72] answered this longstanding question in 1972, showing that there exist closed ideals in $L^1(\mathbb{R})$ that are not of the form $I(g)$.

1.3.10 Convolution on General Domains

Convolution can be defined for functions with domains other than the real line. Let us give two specific examples.

First, consider the discrete analogue of functions, i.e., sequences.

Definition 1.40 (Convolution of Sequences). Let $a = (a_k)_{k\in\mathbb{Z}}$ and $b = (b_k)_{k\in\mathbb{Z}}$ be sequences of complex scalars. Then the *convolution of* a *with* b is the sequence $a * b = ((a * b)_k)_{k\in\mathbb{Z}}$ given by

$$(a * b)_k \;=\; \sum_{j\in\mathbb{Z}} a_j\, b_{k-j}, \tag{1.16}$$

whenever this series converges.

Exercise 1.41. Wherever it makes sense, derive analogues of the theorems of this section for convolution of sequences. For example, results dealing with the continuity or differentiability of convolution of functions will have no analogue for sequences, but results dealing with integrability will have an analogue in terms of summability. In particular, prove analogues of Young's Inequality for sequences. Show that, unlike $L^1(\mathbb{R})$, there does exist an element $\delta \in \ell^1(\mathbb{Z})$ that is an identity for convolution, and we have in fact that

$$\forall\, 1 \le p \le \infty, \quad \forall\, f \in \ell^p(Z), \quad f * \delta = f.$$

For a second example, consider periodic functions on the line. To be precise, we say that a function $f\colon \mathbb{R} \to \mathbb{C}$ is 1-*periodic* if

$$\forall\, x \in \mathbb{R}, \quad f(x+1) = f(x).$$

We define

$$L^p(\mathbb{T}) \;=\; \bigl\{f : f \text{ is 1-periodic and } f \in L^p[0,1)\bigr\}.$$

Definition 1.42 (Convolution of Periodic Functions). Let $f\colon \mathbb{R} \to \mathbb{C}$ and $g\colon \mathbb{R} \to \mathbb{C}$ be 1-periodic measurable functions. Then the *convolution of f with g* is the function $f * g$ given by

$$(f * g)(x) \;=\; \int_0^1 f(y)\, g(x-y)\, dy, \tag{1.17}$$

whenever this integral is well-defined.

Note that $f * g$, if it exists, will be a 1-periodic function.

Exercise 1.43. Wherever it makes sense, derive analogues of the theorems of this section for convolution of periodic functions. In particular, prove analogues of Young's Inequality for periodic functions. Show that there is no function $\delta \in L^1(\mathbb{T})$ that is an identity for convolution.

There is another way to view 1-periodic functions. If we define $\mathbb{T} = [0,1)$, then $\mathbb{T}$ is an abelian group under the operation $\oplus$ of *addition modulo* 1,

$$x \oplus y \;=\; x + y \bmod 1,$$

where $a \bmod 1$ is the fractional part of a. Topologically, $\mathbb{T}$ is homeomorphic to the unit circle in $\mathbb{R}^2$ or $\mathbb{C}$ (thus, "$\mathbb{T}$" for "torus", though only a 1-dimensional torus in our setting). Hence we can identify 1-periodic functions on $\mathbb{R}$ with functions on the group $\mathbb{T}$. Writing $+$ for the operation on $\mathbb{T}$ instead of $\oplus$, convolution of functions on the domain $\mathbb{T}$ is again defined by equation (1.17).

Clearly, there must be a broader context here. The three domains for convolution that we have considered, namely, $\mathbb{R}$, $\mathbb{Z}$, and $\mathbb{T}$, are all *locally compact abelian groups*, i.e., not only are they abelian groups, but they are also endowed with a topology that is locally compact (every point has a neighborhood that is compact). But convolution seems to involve more — we need the existence of Lebesgue measure on $\mathbb{R}$ and $\mathbb{T}$ in order to define the integral in the definition of convolution of functions, and we implicitly need counting measure on $\mathbb{Z}$ to define the series in the definition of the convolution of sequences. It is an amazing fact that every locally compact abelian group G has a positive, regular, translation-invariant Borel measure μ, and this measure is unique up to scaling by positive constants. This measure μ is called the *Haar measure* on G, see [Nac65] or [Fol99]. For $\mathbb{R}$ and $\mathbb{T}$ this is Lebesgue measure, and for $\mathbb{Z}$ it is counting measure. We are not going to define terms precisely here (positive, regular and translation-invariant measures on $\mathbb{R}$ are defined in Appendix D), but simply note that convolution is part of a much broader universe.

Indeed, this is true of this entire volume: Much of what we do for the Fourier transform on $\mathbb{R}$ has analogues that hold for the setting of locally compact abelian groups. There is a general abstract theory of the Fourier transform on locally compact abelian groups, but not every result for the Fourier transform on $\mathbb{R}$ has an analogue in that general setting. We explore this briefly in Sections 2.1 and 2.2. In some ways, our chosen setting of the real line is among the most "complex" of the locally compact abelian groups, as $\mathbb{R}$ is neither compact (like $\mathbb{T}$) nor discrete (like $\mathbb{Z}$).

And the story does not end with *abelian* groups. If G is a locally compact group, but not necessarily abelian, then we do have to distinguish between left and right translations, but still G will have a unique left Haar measure μ_L (unique up to scale, left-translation invariant), and a unique right Haar measure μ_R (unique up to scale, right-translation invariant). The reader can consider how to define convolution in that setting, and what properties that convolution will possess — there will be a difference between "left" and "right" convolution. Still, convolution carries over without too many complications. On the other hand, the analogue of the Fourier transform becomes a very difficult (and interesting) object to define and study on nonabelian groups, and this is the topic of the representation theory of locally compact groups. For these generalizations we must direct the reader to other texts, such as those by Folland [Fol95], Rudin [Rud62], or Hewitt and Ross [HR79].

Additional Problems

1.9. Let $f(x) = e^{-|x|}$, $g(x) = e^{-x^2}$, and $h(x) = xe^{-x^2}$. Show that

$$\begin{aligned}
(f * f)(x) &= (1 + |x|)\, e^{-|x|}, \\
(g * g)(x) &= \left(\frac{\pi}{2}\right)^{1/2} e^{-x^2/2}, \\
(h * h)(x) &= \frac{1}{4}\left(\frac{\pi}{2}\right)^{1/2} (x^2 - 1)\, e^{-x^2/2}.
\end{aligned}$$

1.10. Show that if $f,\, g \in L^1(\mathbb{R})$ and $f,\, g \geq 0$ a.e., then $\|f * g\|_1 = \|f\|_1\, \|g\|_1$. Find a function $h \in L^1(\mathbb{R})$ such that $\|h * h\|_1 < \|h\|_1^2$.

1.11. If $f,\, g \in L^1(\mathbb{R})$, then

$$\mathrm{supp}(f * g) \subseteq \overline{\mathrm{supp}(f) + \mathrm{supp}(g)}$$

(see Notation 1.20). Further, if f and g are both compactly supported then so is $f * g$ and, in this case, $\mathrm{supp}(f * g) \subseteq \mathrm{supp}(f) + \mathrm{supp}(g)$.

1.12. Let A_p be the Babenko–Beckner constant defined in equation (1.12). Prove the following.

(a) $A_p < 1$ for $1 < p < 2$.

(b) $A_p > 1$ for $2 < p < \infty$.

(c) $A_2 = \lim\limits_{p \to 1^+} A_p = \lim\limits_{p \to \infty} A_p = 1$.

1.13. Give an example of $f \in L^p(\mathbb{R})$ with $1 < p < \infty$ and $g \in C_0(\mathbb{R})$ such that $f * g$ is not defined. Compare this to Theorem 1.31 and Exercise 1.35, which show that $f * g \in C_0(\mathbb{R})$ if either $f \in L^1(\mathbb{R})$ and $g \in C_0(\mathbb{R})$, or if $f \in L^p(\mathbb{R})$ and $g \in C_c(\mathbb{R})$.

1.14. Prove the following variation on Exercise 1.35: If $f \in L^1(\mathbb{R})$ and $g \in C_b^m(\mathbb{R})$, then $f * g \in C_b^m(\mathbb{R})$.

1.15. Let $1 \le p \le \infty$ and $f \in L^p(\mathbb{R})$ be given. If there exists an $h \in L^p(\mathbb{R})$ such that

$$\lim_{a \to 0} \left\| h - \frac{f - T_a f}{a} \right\|_p \;=\; 0$$

then we call h a *strong* L^p *derivative of* f and denote it by $h = \partial_p f$. Show that, in this case, if $g \in L^{p'}(\mathbb{R})$ then $f * g$ is differentiable pointwise everywhere, and $(f * g)' = \partial_p f * g$.

1.16. Set $f(x) = e^{-1/x}\, \chi_{[0,\infty)}(x)$.

(a) Show that there exists a polynomial p_n of degree $2n$ such that $f^{(n)}(x) = p_n(x^{-1})\, e^{-1/x}\, \chi_{[0,\infty)}(x)$.

(b) Let $g(x) = f(1 - |x|^2)$, and show that $g \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}(g) = [-1, 1]$.

1.17. Create explicit examples of functions $f \in C_c^\infty(\mathbb{R})$ supported in the interval $[-1, 1]$ that have the following properties.

(a) $0 \le f(x) \le 1$ and $f(0) = 1$.

(b) $\int f(x)\, dx = 0$.

(c) $\int f(x)\, dx = 1$.

(d) $\int f(x)\, dx = 1$, $\int x f(x)\, dx = 0$.

1.18. Let $E \subseteq \mathbb{R}$ be measurable, and show that

$$I(E) \;=\; \{f \in L^1(\mathbb{R}) : \operatorname{supp}(\widehat{f}\,) \subseteq E\}$$

is a closed ideal in $L^1(\mathbb{R})$ (under the operation of convolution).

1.19. Define $A_c(\mathbb{R}) = \{F \in A(\mathbb{R}) : \operatorname{supp}(F) \text{ is compact}\}$. Show that $A_c(\mathbb{R})$ is an ideal in $A(\mathbb{R})$ (under the operation of pointwise multiplication). Compare Problem 1.41, which shows that $A_c(\mathbb{R})$ is dense in $A(\mathbb{R})$, and hence is not a closed ideal in $A(\mathbb{R})$.

1.20. (a) Let $E \subseteq \mathbb{R}$ be measurable with $0 < |E| < \infty$. By considering $\chi_E * \chi_{-E}$, show that $E - E = \{x - y : x, y \in E\}$ contains an open interval $(-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$.

(b) Consider the equivalence relation $x \sim y \iff x - y \in \mathbb{Q}$ on $\mathbb{R}$. By the Axiom of Choice (see Appendix G), there exists a set E that contains exactly one element of each of the equivalence classes of $\sim$. Use part (a) to show that E is not measurable.

(c) Show that every subset of $\mathbb{R}$ that has positive exterior Lebesgue measure contains a nonmeasurable subset.