Fourier Analysis, Distribution Theory, and Wavelets

Bengt Ove Turesson[1]

March 13, 2012

[1]Matematiska institutionen, Linköpings universitet, SE-581 83 Linköping, Sverige

Preface

Text

Bengt Ove Turesson

Linköping March 13, 2012

Contents

Part I

Introductory Material

Chapter 1
Convolutions

1.1. Definition of Convolutions

If f and g are two complex-valued, measurable functions on $\mathbf{R}^d$, their **convolution** $f * g$ is defined by

$$f * g(x) = \int_{\mathbf{R}^d} f(x-y)g(y)\,dy$$

for those values of $x \in \mathbf{R}^d$ for which the integral exists. We will in this chapter give a number of conditions on f and g under which the convolution $f * g$ exists at least a.e.

1.2. Basic Properties of Convolutions

We begin by showing that the convolution between two functions in $L^1(\mathbf{R}^d)$ is defined and belongs to $L^1(\mathbf{R}^d)$.

Proposition 1.2.1. *If $f, g \in L^1(\mathbf{R}^d)$, then the convolution $f * g$ is defined a.e. on $\mathbf{R}^d$. Moreover, $f * g \in L^1(\mathbf{R}^d)$ with $\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1$.*

Proof. We will use the fact that the function $(x, y) \mapsto f(x-y)g(y)$, $(x, y) \in \mathbf{R}^{2d}$, is measurable on $\mathbf{R}^{2d}$ without a proof. According to Tonelli's theorem (Theorem A.7.2),

$$\begin{aligned}
\iint_{\mathbf{R}^{2d}} |f(x-y)||g(y)|\,dx\,dy &= \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |f(x-y)||g(y)|\,dx \right) dy \\
&= \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |f(x-y)|\,dy \right) |g(y)|\,dy \qquad (1.1) \\
&= \int_{\mathbf{R}^d} |f(z)|\,dz \int_{\mathbf{R}^d} |g(y)|\,dy < \infty,
\end{aligned}$$

so it follows that $h \in L^1(\mathbf{R}^{2d})$. Fubini's theorem (Theorem A.7.1) then shows that it follows that $f * g \in L^1(\mathbf{R}^d)$ and, in particular, that the convolution $f * g(x)$ exists for a.e. $x \in \mathbf{R}^d$. The last assertion, finally, follows directly from (1.1). ■

The next proposition shows that convolution is both commutative and associative.

Proposition 1.2.2. *Suppose that $f, g, h \in L^1(\mathbf{R}^d)$. Then*

(i) $f * g = g * f$;

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

Proof.

(a) Making the substitution $z = x - y$, we obtain

$$f * g(x) = \int_{\mathbf{R}^d} f(x-y)g(y)\,dy = \int_{\mathbf{R}^d} f(z)g(x-z)\,dz = g * f(x).$$

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

$$\begin{aligned}(f * g) * h(x) &= \int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^d} f(z)g(x-y-z)\,dz\right)h(y)\,dy\\ &= \int_{\mathbf{R}^d} f(z)\left(\int_{\mathbf{R}^d} g(x-z-y)h(y)\,dy\right)dz\\ &= f*(g*h)(x).\end{aligned}$$ ■

Definition 1.2.3. The **support** of a function f, defined a.e. on $\mathbf{R}^d$, is the set

$$\operatorname{supp} f = \{x \in \mathbf{R}^d : f|_{B_\delta(x)} \neq 0 \text{ for every } \delta > 0\}.$$

Remark 1.2.4. A few remarks are in order.

(a) If x does not belong to $\operatorname{supp} f$, then there exists a ball $B_\delta(x)$ such that $f = 0$ a.e. on $B_\delta(x)$. This implies that the complement of $\operatorname{supp} f$ is open, so $\operatorname{supp} f$ is closed.

(b) Notice also that $f = 0$ a.e. on the complement of $\operatorname{supp} f$.

(c) It follows that if f is integrable on $\mathbf{R}^d$, then $\int_{\mathbf{R}^d} f(x)\,dx = \int_{\operatorname{supp} f} f(x)\,dx$.

(d) One can show that if f is continuous, then $\operatorname{supp} f = \overline{\{x \in \mathbf{R}^d : f(x) \neq 0\}}$. In general, however, this is not true. Take for instance $f = \chi_{\mathbf{Q}}$. Then $\operatorname{supp} f = \varnothing$, but $\overline{\{x \in \mathbf{R}^d : f(x) \neq 0\}} = \mathbf{R}$.

Proposition 1.2.5. *If $f, g \in L^1(\mathbf{R}^d)$, then* $\operatorname{supp}(f*g) \subset \overline{\operatorname{supp} f + \operatorname{supp} g}$.

Here, $\operatorname{supp} f + \operatorname{supp} g$ is the **algebraic sum** of $\operatorname{supp} f$ and $\operatorname{supp} g$, i.e.,

$$\operatorname{supp} f + \operatorname{supp} g = \{x + y : x \in \operatorname{supp} f \text{ and } y \in \operatorname{supp} g\}.$$

It follows from the theorem that if $\operatorname{supp} f$ and $\operatorname{supp} g$ are compact, then $\operatorname{supp} f * g$ is also compact.

Proof (Proposition 1.2.5). Let F be a closed superset to $\operatorname{supp} f + \operatorname{supp} g$. If x_0 does not belong to F, then there exists a number $\delta > 0$ such that $B_\delta(x_0)$ does not intersect F since F is closed. It follows that if $x \in B_\delta(x_0)$, then $x - y \notin \operatorname{supp} f$ for any point $y \in \operatorname{supp} g$, which implies that $f * g(x) = 0$. Hence, the restriction of $f*g$ to $B_\delta(x_0)$ is 0, so x_0 does not belong to $\operatorname{supp}(f*g)$. Thus, $\operatorname{supp}(f*g) \subset F$. This holds for every set F, thus proving the proposition. ■

1.3. Young's Inequality

Our next result about convolutions — often called **Young's inequality** — generalizes Theorem 1.2.1 considerably.

Theorem 1.3.1. *Suppose that $1 \leqslant p, q \leqslant \infty$ satisfy $\frac{1}{p} + \frac{1}{q} \geqslant 1$ and let $1 \leqslant r \leqslant \infty$ be defined by $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. If $f \in L^p(\mathbf{R}^d)$ and $g \in L^q(\mathbf{R}^d)$, then $f*g$ is defined a.e. on $\mathbf{R}^d$ and belongs to $L^r(\mathbf{R}^d)$ with $\|f*g\|_r \leqslant \|f\|_p \|g\|_q$.*

Remark 1.3.2. Before proving the theorem, let us mention a few special cases.

(a) If $p = q = 1$, then $r = 1$, and we retrieve the result in Theorem 1.2.1.

(b) More generally, if $1 \leqslant p \leqslant \infty$ and $q = 1$, then $r = p$, so $f * g \in L^p(\mathbf{R}^d)$ with $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$.

(c) Finally, if $q = p'$, then $r = \infty$, so $f * g \in L^\infty(\mathbf{R}^d)$ with $\|f * g\|_\infty \leqslant \|f\|_p \|g\|_{p'}$.

Proof (Theorem 1.3.1). We first consider the case $r = \infty$. Then $q = p'$, and Hölder's inequality shows that

$$\int_{\mathbf{R}^d} |f(x-y)||g(y)|\,dy \leqslant \|f\|_p \|g\|_{p'} \quad \text{for a.e. } x \in \mathbf{R}^d,$$

from which it follows that $f * g$ exists a.e. and $\|f * g\|_\infty \leqslant \|f\|_p \|g\|_{p'}$.

We next turn to the case $1 \leqslant r < \infty$. Notice that p and q are finite in this case and that $r \geqslant p, q$. Thus, if $\alpha = 1 - p/r$, then $0 \leqslant \alpha < 1$. Let also $\beta = r/q$, so that β satisfies $1 \leqslant \beta < \infty$. It now follows from Hölder's inequality that

$$\begin{aligned} h(x) &= \int_{\mathbf{R}^d} |f(x-y)||g(y)|\,dy = \int_{\mathbf{R}^d} |f(x-y)|^{1-\alpha} |g(y)||f(x-y)|^\alpha\,dy \\ &\leqslant \left(\int_{\mathbf{R}^d} |f(x-y)|^{(1-\alpha)q} |g(y)|^q\,dy \right)^{1/q} \||f|^\alpha\|_{q'} \end{aligned}$$

for a.e. $x \in \mathbf{R}^d$, which implies that

$$h(x)^q \leqslant \|f\|_p^{\alpha q} \int_{\mathbf{R}^d} |f(x-y)|^{(1-\alpha)q} |g(y)|^q\,dy.$$

Using this and Minkowski's integral inequality (Theorem A.6.4), it follows that

$$\begin{aligned} \|h\|_{\beta q}^q &= \|h^q\|_\beta \leqslant \|f\|_p^{\alpha q} \left(\int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |f(x-y)|^{(1-\alpha)q} |g(y)|^q\,dy \right)^\beta dx \right)^{1/\beta} \\ &\leqslant \|f\|_p^{\alpha q} \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |f(x-y)|^{(1-\alpha)\beta q}\,dx \right)^{1/\beta} |g(y)|^q\,dy \\ &= \|f\|_p^{\alpha q} \|f\|_{(1-\alpha)\beta q}^{(1-\alpha)q} \|g\|_q^q \end{aligned}$$

since $\alpha q' = p$. Finally, since $\beta q = r$ and $(1-\alpha)\beta q = p$, we obtain that

$$\|h\|_r \leqslant \|f\|_p^\alpha \|f\|_p^{1-\alpha} \|g\|_q = \|f\|_p \|g\|_q. \qquad \blacksquare$$

1.4. Regularity of Convolutions

We next study regularity properties, i.e., continuity and differentiability, of convolutions. We shall use the fact that translation is a continuous operation on $L^p(\mathbf{R}^d)$ for $1 \leqslant p < \infty$. Here, the **translate** $\tau_h f$ of a function f on $\mathbf{R}^d$ in the direction $h \in \mathbf{R}^d$ is defined by

$$\tau_h f(x) = f(x-h), \quad x \in \mathbf{R}^d.$$

Lemma 1.4.1. *If $f \in L^p(\mathbf{R}^d)$, where $1 \leqslant p < \infty$, then $\tau_h f \to f$ in $L^p(\mathbf{R}^d)$ as $h \to 0$.*

Proof. Let $\varepsilon > 0$ be an arbitrary number and choose a step function ϕ on $\mathbf{R}^d$ such that $\|f-\phi\|_p < \varepsilon$. Using direct calculations or the dominated convergence theorem, it is easy to see that $\tau_h\phi \to \phi$ in $L^p(\mathbf{R}^d)$. It follows that

$$\begin{aligned}\|f - \tau_h f\|_p &\leqslant \|f-\phi\|_p + \|\phi - \tau_h\phi\|_p + \|\tau_h\phi - \tau_h f\|_p \\ &= 2\|f-\phi\|_p + \|\phi-\tau_h\phi\|_p < 3\varepsilon\end{aligned}$$

if $|h|$ is small enough. ■

As noticed in Remark 1.3.2 (c), $f * g \in L^\infty(\mathbf{R}^d)$ if $f \in L^p(\mathbf{R}^d)$ and $g \in L^{p'}(\mathbf{R}^d)$, where $1 \leqslant p \leqslant \infty$. We next show that $f*g$ is actually uniformly continuous in this case and also that $f*g(x)$ tends to 0 as $|x| \to \infty$ if $1 < p < \infty$.

Theorem 1.4.2. *Suppose that $f \in L^p(\mathbf{R}^d)$ and $g \in L^{p'}(\mathbf{R}^d)$, where $1 \leqslant p \leqslant \infty$. Then $f * g$ is uniformly continuous on $\mathbf{R}^d$. For $1 < p < \infty$, there also holds that $\lim_{|x|\to\infty} f * g(x) = 0$.*

Proof. To prove that $f*g$ is uniformly continuous, we may assume that $1 \leqslant p < \infty$ (if $p = \infty$, we let f och g change roles). An application of Hölder's inequality then shows that

$$\begin{aligned}|f*g(x+h) - f*g(x)| &\leqslant \int_{\mathbf{R}^d} |f(x+h-y) - f(x-y)||g(y)|\,dy \\ &\leqslant \|\tau_{-h}f - f\|_p\|g\|_{p'}.\end{aligned}$$

According to Lemma 1.4.1, $\|\tau_{-h}f - f\|_p \to 0$ as $|h| \to 0$, so it follows that the convolution $f*g$ is uniformly continuous. For the proof of the second assertion, we let $f_n = \chi_{B_n(0)}f$ and $g_n = \chi_{B_n(0)}g$ for $n = 1, 2, \ldots$. Then $f_n \to f$ in $L^p(\mathbf{R}^d)$ and $g_n \to g$ in $L^{p'}(\mathbf{R}^d)$. The first part of the proof together with Theorem 1.2.5 also shows that $f_n * g_n \in C_c(\mathbf{R}^d)$. Moreover,

$$\|f_n * g_n - f*g\|_\infty \leqslant \|f\|_p\|g_n - g\|_{p'} + \|f_n - f\|_p\|g\|_{p'}.$$

This shows that $f_n * g_n \to f*g$ uniformly, from which it follows that $f*g(x) \to 0$ as $|x| \to \infty$. ■

We now consider differentiability of convolutions. In general, one expects $f*g$ to be at least as smooth as either f or g. Formally, this follows by differentiating $f*g$ under the integral sign:

$$\partial^\alpha(f*g)(x) = \partial^\alpha \int_{\mathbf{R}^d} f(x-y)g(y)\,dy = \int_{\mathbf{R}^d} \partial_x^\alpha f(x-y)g(y)\,dy = (\partial^\alpha f) * g(x)$$

if $\partial^\alpha f$ exists, so that $\partial^\alpha(f*g) = (\partial^\alpha f)*g$. Similarly, $\partial^\alpha(f*g) = f*\partial^\alpha g$ if $\partial^\alpha g$ exists. We will now show that these formal calculations can be justified if certain conditions are imposed on f and g.

Theorem 1.4.3. *Suppose that $f \in L^p(\mathbf{R}^d)$, where $1 \leqslant p \leqslant \infty$, and $g \in C^m(\mathbf{R}^d)$ with $\partial^\alpha g \in L^{p'}(\mathbf{R}^d)$ for $|\alpha| \leqslant m$. Then $f*g \in C^m(\mathbf{R}^d)$ and*

$$\partial^\alpha(f*g) = f*\partial^\alpha g \quad \text{for } |\alpha| \leqslant m. \tag{1.2}$$

Proof. It suffices to prove the theorem for $m = 1$; the general case follows by induction. The fact that $f * g$ is continuous is a consequence of Theorem 1.4.2. To prove (1.2), we first consider the case $p = 1$. Let e_j be one of the elements in the standard basis for $\mathbf{R}^d$. We will use the following notation:

$$D_j(f * g)(x, h) = \frac{f * g(x + he_j) - f * g(x)}{h},$$

where $x \in \mathbf{R}^d$ and $h \in \mathbf{R} \setminus \{0\}$. Using the assumptions on f and g together with the dominated convergence theorem, we see that

$$\begin{aligned} D(f * g)(x, h) &= \int_{\mathbf{R}^d} f(x - y) \frac{g(y + he_j) - g(y)}{h} \, dy \\ &\longrightarrow \int_{\mathbf{R}^d} f(x - y) \partial_j g(y) \, dy \quad \text{as } h \to 0, \end{aligned}$$

which shows that $\partial_j(f * g) = f * \partial_j g$ for $j = 1, \ldots, d$. Now suppose that $1 < p \leqslant \infty$. Given $\varepsilon > 0$, choose $R \geqslant 2$ so large that

$$\int_{|y| \geqslant R/2} |g(y)|^{p'} \, dy < \varepsilon^{p'}.$$

According to the mean-value theorem,

$$\begin{aligned} |f * \partial_j g(x) - D_j(f * g)(x, h)| \leqslant & \int_{|y| < R} |f(x - y)| |\partial_j g(y) - \partial_j g(y + \theta h e_j)| \, dy \\ & + \int_{|y| \geqslant R} |f(x - y)| |\partial_j g(y) - \partial_j g(y + \theta h e_j)| \, dy \end{aligned}$$

for some $\theta \in [0, 1]$. In this inequality, the first integral in the right-hand side tends to 0 as $h \to 0$ since the integrand tends to 0, f is locally integrable, and $\partial_j g$ is locally bounded. If $|h| \leqslant 1$, we also have

$$\begin{aligned} \int_{|y| \geqslant R} |f(x - y)| |\partial_j g(y) - \partial_j g(y + \theta h e_j)| \, dy &\leqslant 2\|f\|_p \left(\int_{|y| \geqslant R/2} |\partial_j g(y)|^{p'} \, dy \right)^{1/p'} \\ &< 2\|f\|_p \varepsilon. \end{aligned}$$

This establishes (1.2) in the case $1 < p \leqslant \infty$. Finally, $f * \partial_j g \in C(\mathbf{R}^d)$ according to Theorem 1.4.2, so $f * g \in C^1(\mathbf{R}^d)$. ■

Remark 1.4.4. Using exactly the same technique, one can show that the assertion in Theorem 1.4.3 also holds true if we assume that $f \in L^1_{\text{loc}}(\mathbf{R}^d)$ and $g \in C_c^m(\mathbf{R}^d)$.

1.5. Approximate Identities

According to Theorem 1.2.1 and Theorem 1.2.2 (a), $L^1(\mathbf{R}^d)$ is a commutative **Banach algebra**[1] with convolution as the product. A natural question to ask

[1] A Banach algebra is a Banach space B equipped with a product $*$ such that $\|f * g\| \leqslant \|f\| \|g\|$ for all elements $f, g \in B$.

is whether this algebra has an multiplicative identity, i.e., if there exists a function $K \in L^1(\mathbf{R}^d)$ such that

$$K * f = f * K = f \text{ for every } f \in L^1(\mathbf{R}^d).$$

The answer to this question is in fact "no". Indeed, suppose that K were such a function. Then $K * f = f$ for every function $f \in L^1(\mathbf{R}^d) \cap L^\infty(\mathbf{R}^d)$. This is a contradiction since $K * f$ is continuous in this case according to Theorem 1.4.2.

There are, however, sequences $(K_n)_{n=1}^\infty \subset L^1(\mathbf{R}^d)$ that approximate a multiplicative identity in the sense that $K_n * f \to f$ in $L^1(\mathbf{R}^d)$ as $n \to \infty$ for every $f \in L^1(\mathbf{R}^d)$. We will now see how such sequences can be constructed.

Definition 1.5.1. A sequence $(K_n)_{n=1}^\infty$ of integrable functions on $\mathbf{R}^d$ is called an **approximate identity** if

(i) $\int_{\mathbf{R}^d} K_n(x)\,dx = 1$ for every n;
(ii) there exists a constant $C \geqslant 0$ such that $\int_{\mathbf{R}^d} |K_n(x)|\,dx \leqslant C$ for every n;
(iii) $\lim_{n\to\infty} \int_{|x|\geqslant\delta} |K_n(x)|\,dx = 0$ for every $\delta > 0$.

Notice that if $K \geqslant 0$, then (ii) follows from (i). A simple recipe for constructing an approximate identity is given by the following proposition:

Proposition 1.5.2. *Suppose that $K \in L^1(\mathbf{R}^d)$ satisfies the conditions $K \geqslant 0$ and $\int_{\mathbf{R}^d} K(x)\,dx = 1$. Put*

$$K_n(x) = n^d K(nx),\ x \in \mathbf{R}^d, \quad \textit{for } n = 1, 2, \ldots. \tag{1.3}$$

Then $(K_n)_{n=1}^\infty$ is an approximate identity.

Proof. Changing variables $y = nx$, we see that $\int_{\mathbf{R}^d} K_n(x)\,dx = 1$ for every n. The same change of variables shows that

$$\int_{|x|\geqslant\delta} |K_n(x)|\,dx = \int_{|y|\geqslant n\delta} |K(y)|\,dy \longrightarrow 0 \quad \text{as } n \to \infty. \qquad \blacksquare$$

Theorem 1.5.3. *Suppose that $(K_n)_{n=1}^\infty$ is an approximate identity and moreover that $f \in L^p(\mathbf{R}^d)$, where $1 \leqslant p < \infty$. Then $K_n * f \in L^p(\mathbf{R})$ and $K_n * f \to f$ in $L^p(\mathbf{R}^d)$ as $n \to \infty$.*

Proof. The fact that $K_n * f \in L^p(\mathbf{R})$ follows from Young's inequality (see Remark 1.3.2 (b)). Minkowski's integral inequality (Theorem A.6.4) now shows that

$$\begin{aligned}\left(\int_{\mathbf{R}^d} |f(x) - K_n f(x)|^p\,dx\right)^{1/p} &= \left(\int_{\mathbf{R}^d}\left|\int_{\mathbf{R}^d} (f(x) - f(x-y))K_n(y)\,dy\right|^p dx\right)^{1/p} \\ &\leqslant \int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^d} |f(x) - f(x-y)|^p\,dx\right)^{1/p} |K_n(y)|\,dy \\ &= \int_{\mathbf{R}^d} \|f - \tau_y f\|_p |K_n(y)|\,dy.\end{aligned}$$

We then split the last integral into two parts:

$$\int_{|y|<\delta} \|f-\tau_y f\|_p |K_n(y)|\,dy + \int_{|y|\geqslant\delta} \|f-\tau_y f\|_p |K_n(y)|\,dy.$$

Since $\|f-\tau_y f\|_p \to 0$ as $y \to 0$ according to Lemma 1.4.1 and (ii) in Definition 1.5.1 holds, the first integral can be made arbitrarily small by choosing δ sufficiently small. Moreover, using the fact that $\|f-\tau_y f\|_p \leqslant 2\|f\|_p$ and (iii) in Definition 1.5.1, we see that the second integral tends to 0 as $n \to \infty$. ■

The next result concerns pointwise and uniform convergence of convolutions with approximate identities.

Theorem 1.5.4. *Suppose that $(K_n)_{n=1}^{\infty}$ is an approximate identity and moreover that $f \in L^p(\mathbf{R}^d)$, where $1 \leqslant p \leqslant \infty$. Suppose also that*

$$\|K_n\|_{L^{p'}(\{x\in\mathbf{R}^d : |x|\geqslant\delta\})} \longrightarrow 0 \quad \text{as } n \to \infty \tag{1.4}$$

for every $\delta > 0$. Then $K_n * f \to f$ uniformly as $n \to \infty$ on a every compact set $K \subset \mathbf{R}^d$ where f is continuous.

Remark 1.5.5.

(i) Notice that pointwise convergence corresponds to the case when K consists of just one point.

(ii) Notice also that if $p = \infty$, then (1.4) coincides with (iii) in Definition 1.5.1.

(iii) Suppose that K is of the form (1.3), where $K(x) = o(|x|^{-d})$ as $|x| \to \infty$, i.e., $K(x) = |x|^{-d} r(x)$, where $r(x) \to 0$ as $|x| \to \infty$. Then, for $p = 1$,

$$\|K_n\|_{L^{p^\infty}(\{x\in\mathbf{R}^d : |x|\geqslant\delta\})} = \sup_{|x|\geqslant\delta} \frac{|r(nx)|}{|x|^\delta} \leqslant \delta^{-d} \sup_{|y|\geqslant n\delta} |r(y)| \longrightarrow 0 \quad \text{as } n \to \infty,$$

and, for $1 < p \leqslant \infty$,

$$\begin{aligned} \|K_n\|_{L^{p'}(\{x\in\mathbf{R}^d : |x|\geqslant\delta\})} &= \left(\int_{|x|\geqslant\delta} \frac{|r(nx)|^{p'}}{|x|^{dp'}}\,dx\right)^{1/p'} \\ &\leqslant C_{d,p} \sup_{|y|\geqslant n\delta} |r(y)| \longrightarrow 0 \quad \text{as } n \to \infty. \end{aligned}$$

Therefore, (1.4) holds.

Proof (Theorem 1.5.4). Given $\varepsilon > 0$, choose $\delta > 0$ such that $|f(x)-f(x-y)| < \varepsilon$ for every $x \in K$ and every $y \in \mathbf{R}^d$ that satisfies $|y| < \delta$. Suppose that $|f(x)| \leqslant M$ for every $x \in K$. Then

$$\begin{aligned} |f(x) - K_n * f(x)| &\leqslant \int_{|y|<\delta} |f(x) - f(x-y)||K_n(y)|\,dy \\ &\quad + |f(x)| \int_{|y|\geqslant\delta} |K_n(y)|\,dy + \int_{|y|\geqslant\delta} |f(x-y)||K_n(y)|\,dy \\ &\leqslant C\varepsilon + M \int_{|y|\geqslant\delta} |K_n(y)|\,dy + \int_{|y|\geqslant\delta} |f(x-y)||K_n(y)|\,dy. \end{aligned}$$

The first integral in the last member tends to 0 as $n \to \infty$ because of (iii). Also,

$$\int_{|y|\geqslant\delta} |f(x-y)||K_n(y)|\,dy \leqslant \|f\|_p \|K_n\|_{L^{p'}(\{y\in\mathbf{R}^d:|y|\geqslant\delta\})},$$

which shows that also the second integral tends to 0 as $n \to \infty$. It easy to see that the convergence is uniform. ■

Remark 1.5.6.

(a) In the definition of an approximate identity, the indices are the positive integers and the statements in the theorems just proved hold when $n \to \infty$. In many cases, however, the indices naturally belong to some other subset of the reals. One can, for instance, consider sequences (K_ε), where the index ε belongs to $(0,\infty)$ and the limiting value for ε is 0. We will also call such sequences approximate identities if they satisfy the properties in Definition 1.5.1 (with appropriate modifications). Let us also mention that Theorem 1.5.3 and Theorem 1.5.4 hold true in such cases with identical proofs.

(b) In one dimension and under the assumption that every K_n is even, it is possible to modify the proof of Theorem 1.5.4 to handle jump discontinuities. Suppose that $f \in L^\infty(\mathbf{R})$ and that the one-sided limits $f(x^+) = \lim_{y\to 0^+} f(x+y)$ and $f(x^-) = \lim_{y\to 0^+} f(x-y)$ exist. Using the fact that K_n is even, we see that

$$\int_0^\infty K_n(y)\,dy = \frac{1}{2} \quad \text{for } n = 1,2,\dots,$$

from which it follows that

$$\begin{aligned}\frac{1}{2}(f(x^+)+f(x^-)) - K_n * f(x) &= \int_0^\infty (f(x^+)-f(x+y))K_n(y)\,dy \\ &\quad + \int_0^\infty (f(x^-)-f(x-y))K_n(y)\,dy.\end{aligned}$$

As in the proof of Theorem 1.5.4, one then shows that both these integrals tend to 0 as $n \to \infty$, so $K_n * f(x) \to \frac{1}{2}(f(x^+)+f(x^-))$.

Example 1.5.7. Put

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

and

$$P_\varepsilon(x) = \varepsilon^{-1}P(\varepsilon^{-1}x) = \frac{1}{\pi}\frac{\varepsilon}{\varepsilon^2+x^2}, \quad x \in \mathbf{R},\ \varepsilon > 0.$$

Notice that $P \in L^p(\mathbf{R})$ for $1 \leqslant p \leqslant \infty$. Since $\int_{-\infty}^\infty P(x)\,dx = 1$ and $P \geqslant 0$, $(P_\varepsilon)_{\varepsilon>0}$ is an approximate identity. We call P_ε the **Poisson kernel**. The integral

$$P_\varepsilon * f(x) = \int_{-\infty}^\infty f(t)P_\varepsilon(x-t)\,dt, \quad x \in \mathbf{R},$$

where $f \in L^p(\mathbf{R})$, is called the **Poisson integral** of f. Let $u(x,y) = P_y * f(x)$, where (x,y) belongs to the upper half plane

$$H = \{(x,y) \in \mathbf{R}^2 : -\infty < x < \infty \text{ and } y > 0\}.$$

Notice that $P_y(x)$ is the imaginary part of $-z^{-1}$, where $z = x + iy \in \mathbf{C}$, so $P_y(x)$ is harmonic in H. This fact together with Theorem 1.4.3 imply that

$$\Delta u(x, y) = \int_{-\infty}^{\infty} f(t) \Delta P_y(x - t)\, dt = 0 \quad \text{for } (x, y) \in H,$$

which shows that u is harmonic in H. If moreover $f \in C(\mathbf{R})$, then Theorem 1.5.4 shows that $u(x, y) \to f(x)$ as $y \to 0$ for every $x \in \mathbf{R}$. Thus, u is a solution to the **Dirichlet problem** in H, i.e., a solution to **Laplace's equation** $\Delta u = 0$ in H with boundary values f. □

Example 1.5.8. Put

$$W(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbf{R},$$

and

$$W_t(x) = t^{-1/2} W(t^{-1/2} x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}, \quad x \in \mathbf{R},\ t > 0.$$

Then $(W_t)_{t>0}$ is an approximate identity since $\int_{-\infty}^{\infty} W(x)\, dx = 1$ and $W > 0$; the kernel W_t is known as the **Gauss kernel**. If $f \in L^p(\mathbf{R})$, where $1 \leqslant p \leqslant \infty$, then the function $u(x, t) = W_t * f(x)$, $(x, t) \in H$, solves the **heat equation**:

$$u'_t - u''_{xx} = 0 \quad \text{for } (x, t) \in H.$$

If also $f \in C(\mathbf{R})$, then $u(x, t) \to f(x)$ as $t \to 0$ for every $x \in \mathbf{R}$, so u has boundary values f. □

Example 1.5.9. Put

$$K(x) = \frac{1}{\pi} \left(\frac{\sin x}{x} \right)^2, \quad x \in \mathbf{R},$$

and

$$K_n(x) = nK(nx) = \frac{1}{\pi} \frac{\sin^2 x}{nx^2}, \quad x \in \mathbf{R},\ n = 1, 2, \ldots.$$

One can show that $\int_{-\infty}^{\infty} K(x)\, dx = 1$, so $(K_n)_{n=1}^{\infty}$ is an approximate identity. The kernel K_n is called the **Fejér kernel** for the real line. If $f \in L^p(\mathbf{R}) \cap C(\mathbf{R})$, where $1 \leqslant p \leqslant \infty$, then $K_n * f(x) \to f(x)$ as $n \to \infty$ for every $x \in \mathbf{R}$. □

1.6. Regularization

In many situations, it is important to be able to approximate an L^p-function with smooth functions. The standard procedure for this is to use mollifiers.

Definition 1.6.1. A **mollifier** is a function $\phi \in C_c^\infty(\mathbf{R}^d)$ that satisfies the conditions $\phi \geqslant 0$, $\operatorname{supp} \phi \subset \overline{B_1(0)}$, and $\int_{\mathbf{R}^d} \phi\, dx = 1$.

The following example contains the standard example of a mollifier.

Example 1.6.2. It is not so hard to show that the function ψ on $\mathbf{R}$, defined by

$$\psi(t) = \begin{cases} e^{-1/t} & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t \leqslant 0 \end{cases}$$

belongs to $C^\infty(\mathbf{R})$; this comes down to showing that all right-hand derivatives of ψ are 0 at $t = 0$. Now put $\phi(x) = C\psi(1 - |x|^2)$ for $x \in \mathbf{R}^d$, where the constant C is chosen so that $\int_{\mathbf{R}^d} \phi \, dx = 1$. Then $\phi \in C_c^\infty(\mathbf{R}^d)$ with support in the closed unit ball $\{x \in \mathbf{R}^d : |x| \leqslant 1\}$. □

If ϕ is a mollifier and $\varepsilon > 0$, put

$$\phi_\varepsilon(x) = \varepsilon^{-d}\phi(\varepsilon^{-1}x), \ x \in \mathbf{R}^d.$$

According to Proposition 1.5.7 and Remark 1.5.6, $(\phi_\varepsilon)_{\varepsilon>0}$ is then an approximate identity. Notice also that $\operatorname{supp} \phi_\varepsilon \subset \overline{B_\varepsilon(0)}$.

The following theorem, which is a consequence of Theorem 1.2.5, Remark 1.4.4, Theorem 1.5.3, and Theorem 1.5.4, summarizes a number of useful properties of convolutions with mollifiers.

Theorem 1.6.3. *Suppose that ϕ is a mollifier and moreover that $f \in L^p(\mathbf{R}^d)$, where $1 \leqslant p \leqslant \infty$. Then the following properties hold:*

(i) the convolution $\phi_\varepsilon * f$ exists a.e. on $\mathbf{R}^d$ and belongs to $L^p(\mathbf{R}^d)$;

(ii) $\phi_\varepsilon * f \in C^\infty(\mathbf{R}^d)$;

(iii) the support of $\phi_\varepsilon * f$ is a subset of the closed ε-neighbourhood of $\operatorname{supp} f$;

(iv) if $1 \leqslant p < \infty$, then $\phi_\varepsilon * f \to f$ in $L^p(\mathbf{R}^d)$ as $\varepsilon \to 0$;

(v) $\phi_\varepsilon * f \to f$ uniformly as $\varepsilon \to 0$ on every compact subset to $\mathbf{R}^d$ where f is continuous.

Notice that if $\operatorname{supp} f$ is compact, then $\phi_\varepsilon * f \in C_c^\infty(\mathbf{R}^d)$. By an ε-**neighbourhood** of a subset E to $\mathbf{R}^d$, we mean the set

$$\{x \in \mathbf{R}^d : \operatorname{dist}(x, E) < \varepsilon\}.$$

Its closure, i.e., the set obtained by replacing strict inequality with inequality, is called a **closed** ε-**neighbourhood** of E.

1.7. Partitions of Unity

The next proposition shows how the characteristic function of a compact set can be regularized.

Proposition 1.7.1. *Suppose that $X \subset \mathbf{R}^d$ is open and that K is a compact subset to X. Then there exists a function $\psi \in C_c^\infty(X)$ such that $0 \leqslant \psi \leqslant 1$ and $\psi = 1$ on K.*

Proof. Let 3δ be the distance from K to X^c and let χ_δ be the characteristic function of a δ-neighbourhood of K. If ϕ is a mollifier and if ε satisfies $0 < \varepsilon < \delta$, then the function $\psi = \phi_\varepsilon * \chi_\delta$ belongs to $C^\infty(X)$ with support in the closed 2δ-neighbourhood of K. Moreover, it is easily checked that $0 \leqslant \psi \leqslant 1$ and $\psi = 1$ on K. ■

Corollary 1.7.2. *Suppose that* $X_1 \dots, X_m \subset \mathbf{R}^d$ *are open and that* $\phi \in C_c^\infty(X)$, *where* $X = \bigcup_{j=1}^m X_j$. *Then there exist functions* $\phi_j \in C_c^\infty(X_j)$, $j = 1, \dots, m$, *such that*

$$\phi = \sum_{j=1}^m \phi_j. \tag{1.5}$$

If $\phi \geqslant 0$, *then* $\phi_j \geqslant 0$ *for* $j = 1, \dots, m$.

Proof. It is easy to see that there exist compact sets $K_1, \dots, K_m \subset X$ such that $K_j \subset X_j$ for every j and $\operatorname{supp}\phi \subset \bigcup_{j=1}^m K_j$. Now, using Proposition 1.7.1, choose functions $\psi_j \in C_c^\infty(X_j)$ that satisfy $0 \leqslant \psi_j$ and $\psi_j = 1$ on K_j, and put

$$\phi_1 = \phi\psi_1,\ \phi_2 = \phi\psi_2(1-\psi_1), \dots,\ \phi_m = \phi\psi_m(1-\psi_1)\cdot \dots \cdot (1-\psi_{m-1}).$$

Then these functions satisfy (1.5) since

$$\sum_{j=1}^m \phi_j - \phi = -\phi \prod_{j=1}^m (1-\psi_j) = 0. \qquad \blacksquare$$

By combining Proposition 1.7.1 with Corollary 1.7.2, we obtain following result. The functions ϕ_j in the Proposition are called a **partition of unity** subordinate to the covering $\bigcup_{j=1}^m X_j$ of K.

Corollary 1.7.3. *Suppose that* $X_1 \dots, X_m \subset \mathbf{R}^d$ *are open and that* K *is a compact subset to* $\bigcup_{j=1}^m X_j$. *Then there exist functions* $\phi_j \in C_c^\infty(X_j)$, $j = 1, \dots, m$, *such that* $0 \leqslant \phi_j \leqslant 1$ *for every* j *and* $\sum_{j=1}^m \phi_j \leqslant 1$ *with equality on* K.

1.8. A Density Theorem

The following density theorem is a consequence of Theorem 1.6.3.

Theorem 1.8.1. *If* $1 \leqslant p < \infty$ *and* $X \subset \mathbf{R}^d$ *is open, then* $C_c^\infty(X)$ *is dense in* $L^p(X)$.

In the proof of the theorem, we use the following lemma.

Lemma 1.8.2. *Suppose that* X *is an open subset to* $\mathbf{R}^d$ *and that* $f \in L^p(X)$, *where* $1 \leqslant p < \infty$. *Then there exists a sequence* $(f_n)_{n=1}^\infty$ *such that every function* f_n *has compact support and* $f_n \to f$ *in* $L^p(X)$.

Proof. Put $K_n = \{x \in X : |x| \leqslant n \text{ and } \operatorname{dist}(x, X^c) \geqslant 1/n\}$ and $f_n = \chi_{K_n} f$ for $n = 1, 2, \dots$. Then every set K_n is compact and

$$\|f - f_n\|_p^p = \int_X |f - f_n|^p \, dx \longrightarrow 0 \quad \text{as } n \to \infty$$

due to dominated convergence. $\blacksquare$

Proof (Theorem 1.8.1). Let $f \in L^p(\mathbf{R}^d)$. Given $\varepsilon > 0$, choose a function g with compact support in X such that $\|f - g\|_p < \varepsilon$. Then extend g to $\mathbf{R}^d$ by letting $g = 0$ outside X. If ϕ is a mollifier on $\mathbf{R}^d$, then $\phi_\eta * g \in C_c^\infty(X)$ if η is chosen so small that $\operatorname{supp}(\phi_\eta * g) \subset X$. Moreover, $\|g - \phi_\eta * g\|_p < \varepsilon$ for a possibly even smaller value of η. Thus, for a sufficiently small η,

$$\|f - \phi_\eta * g\|_p \leqslant \|f - g\|_p + \|g - \phi_\eta * g\|_p < 2\varepsilon. \qquad \blacksquare$$

1.9. Periodic Convolutions

There is a corresponding convolution for functions f and g on $\mathbf{R}$ with period 2π, namely

$$f * g(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t-s) g(s)\, ds.$$

In this chapter, we will concentrate on the non-periodic case; let us just mention that all results remain true in the periodic case.

Part II

Fourier Series

Chapter 2

L^1-theory for Fourier Series

2.1. Function Spaces

For $1 \leqslant p \leqslant \infty$, we let $L^p(\mathbb{T})$ denote the class of measurable functions f, defined a.e. on $\mathbf{R}$, such that f has period 2π, i.e.,

$$f(t+2\pi) = f(t) \quad \text{for a.e. } t \in \mathbf{R},$$

and $f \in L^p(-\pi,\pi)$. In the case $1 \leqslant p < \infty$, we equip $L^p(\mathbb{T})$ with the norm

$$\|f\|_p = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^p \, dt\right)^{1/p},$$

and for $p = \infty$, we use the norm of $L^\infty(-\pi,\pi)$:

$$\|f\|_\infty = \inf\{C : |f(t)| \leqslant C \text{ a.e.}\}.$$

With these norms, $L^p(\mathbb{T})$ are Banach spaces. Notice that if f belongs to $L^p(\mathbb{T})$, where $1 \leqslant p < \infty$, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)| \, dt \leqslant \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^p \, dt\right)^{1/p},$$

according to Hölder's inequality, which shows that $L^p(\mathbb{T}) \subset L^1(\mathbb{T})$. We similarly have $L^\infty(\mathbb{T}) \subset L^1(\mathbb{T})$.

For $m = 0, 1, \dots$, we denote by $C^m(\mathbb{T})$ the class of m times continuously differentiable functions on $\mathbf{R}$ with period 2π, equipped with the norm

$$\|f\|_{C^m(\mathbb{T})} = \sum_{j=0}^{m} \|f^{(j)}\|_\infty.$$

It is known that $C^m(\mathbb{T})$ is a Banach space which is dense in $L^p(\mathbb{T})$ for $1 \leqslant p < \infty$.

2.2. Fourier Series and Fourier Coefficients

Definition 2.2.1. The **Fourier series** of a function $f \in L^1(\mathbb{T})$ is the formal series

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{int}, \tag{2.1}$$

where the **Fourier coefficients** $\hat{f}(n)$ are defined by

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) e^{-int} \, dt, \quad n = 0, \pm 1, \dots. \tag{2.2}$$

The series (2.1) is **convergent** at $t \in \mathbf{R}$ with value S if

$$\lim_{N\to\infty} \sum_{n=-N}^{N} \hat{f}(n) e^{int} = S. \tag{2.3}$$

Since we do not assume that the Fourier series of a function f is absolutely convergent, it is necessary to define in what sense (2.1) should be interpreted. Interpreting the Fourier series as the limit of symmetric partial sums as in (2.3) gives a satisfactory theory with nice formulae and also allows for cancellation. Let us stress that we — at this stage — consider the Fourier series of a function as a purely formal object and that we do not assume that it converges in any sense.

Example 2.2.2. Let $f \in L^1(\mathbb{T})$ be defined by $f(t) = t$ for $-\pi \leqslant t < \pi$. Then

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} te^{-int}\,dt = \begin{cases} 0 & \text{if} \quad n = 0 \\ i\dfrac{(-1)^n}{n} & \text{if} \quad n \neq 0 \end{cases}.$$

The Fourier series of f is thus

$$i\sum_{n\neq 0}\frac{(-1)^n}{n}e^{int} = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin nt.$$

The last identity holds if either side converges because of the way we have defined convergence for a Fourier series. We will return to this function and its Fourier series in Example 2.6.3. □

2.3. Trigonometric Series

Definition 2.3.1. A **trigonometric series** is a formal series of the form

$$\sum_{n=-\infty}^{\infty} c_n e^{int},$$

where $(c_n)_{n=-\infty}^{\infty}$ is some sequence of complex numbers.

Every Fourier series is of course a trigonometric series. There are, however, trigonometric series that are not Fourier series. We now give an example of a trigonometric series that later will be shown not to be a Fourier series (see Example 2.10.4). To prove convergence, we will use a little discrete analysis.

If $(a_n)_{n=0}^{\infty}$ is a sequence of complex numbers, we define the **forward difference** Δa_n by

$$\Delta a_n = a_{n+1} - a_n, \quad n = 0, 1, \dots .$$

Then the following product rule holds:

$$\Delta(a_n b_n) = (\Delta a_n)b_n + a_n \Delta b_n, \quad n = 0, 1, \dots$$

for all sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$. Summing both sides in this identity from M to N, where $N > M \geqslant 1$, we obtain the formula for **summation by parts**:

$$\sum_{n=M}^{N} a_n \Delta b_n = a_{N+1}b_{N+1} - a_M b_M - \sum_{n=M}^{N}(\Delta a_n)b_n. \tag{2.4}$$

The reader should compare this formula with the formula for integration by parts. Notice also that if we put

$$A_n = \begin{cases} \sum_{k=0}^{n-1} a_k & \text{for} \quad n = 1, 2, \ldots \\ 0 & \text{for} \quad n = 0 \end{cases},$$

then $(A_n)_{n=1}^{\infty}$ is a **primitive** to $(a_n)_{n=1}^{\infty}$ in the sense that $\Delta A_n = a_n$ for $n = 0, 1, \ldots$.

Proposition 2.3.2. *Suppose that $(a_n)_{n=0}^{\infty}$ is a decreasing sequence of real numbers such that $a_n \to 0$ as $n \to \infty$. Then the trigonometric series $\sum_{n=0}^{\infty} a_n e^{int}$ is convergent for $t \notin 2\pi\mathbf{Z}$. The series also converges uniformly on every compact subset K to $\mathbf{R}$ such that $K \subset (2k\pi, 2(k+1)\pi)$ for some number $k \in \mathbf{Z}$.*

Proof. For $t \in \mathbf{R}$, put $B_n(t) = \sum_{k=0}^{n-1} e^{kt}$, $n = 1, 2, \ldots$, and $B_0(t) = 0$. Using the fact that $|e^{it} - 1| = 2|\sin \frac{t}{2}|$, we see that

$$|B_n(t)| = \left| \frac{e^{int} - 1}{e^{it} - 1} \right| \leqslant \frac{1}{|\sin \frac{t}{2}|}$$

for every $t \notin 2\pi\mathbf{Z}$ and $n = 0, 1, \ldots$. It also follows from (2.4) that

$$\sum_{n=M}^{N} a_n e^{int} = a_{N+1} B_{N+1}(t) - a_M B_M(t) + \sum_{n=M}^{N} (\Delta a_n) B_n(t).$$

The first two terms in the right-hand side of this equation tend to 0 as $M, N \to \infty$. This also applies to the third term since

$$\left| \sum_{n=M}^{N} (\Delta a_n) B_n(t) \right| \leqslant \frac{1}{|\sin \frac{t}{2}|} \sum_{n=M}^{N} \Delta a_n = \frac{1}{|\sin \frac{t}{2}|} (a_{N+1} - a_M).$$

The statement about uniform convergence on a compact subset K to $\mathbf{R}$ such that $K \subset (2n\pi, 2(n+1)\pi)$ holds since $|\sin \frac{t}{2}|^{-1}$ is bounded on K. ■

Example 2.3.3. Proposition 2.3.2 shows that the series

$$\sum_{n=2}^{\infty} \frac{e^{int}}{\ln n}$$

is convergent for $t \notin 2\pi\mathbf{Z}$. It follows that the imaginary part of this series, namely the series

$$\sum_{n=2}^{\infty} \frac{\sin nt}{\ln n},$$

converges for every $t \in \mathbf{R}$. We will show in Example 2.10.4 that this series is in fact not a Fourier series. □

2.4. Properties of Fourier Coefficients

We next collect some useful properties of the Fourier coefficients of a function. The mapping, which maps a function $f \in L^1(\mathbb{T})$ to the sequence $(\hat{f}(n))_{n=-\infty}^{\infty}$ is called the **finite Fourier transform**. According to the following result, which follows directly from the definition, this map is linear.

Proposition 2.4.1. *Suppose that $f, g \in L^1(\mathbb{T})$ and $\alpha, \beta \in \mathbf{C}$. Then*

$$\widehat{\alpha f + \beta g}(n) = \alpha \hat{f}(n) + \beta \hat{g}(n) \quad \textit{for every } n \in \mathbf{Z}.$$

The next proposition shows that the discrete Fourier transform of a convolution is the product of the transforms of the functions involved. Recall that the convolution between two functions f and g with period 2π is defined by

$$f * g(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t-s) g(s)\, ds,$$

and that $f * g$ exists a.e. and belongs to $L^1(-\pi, \pi)$; see Section 1.9.

Proposition 2.4.2. *Suppose that $f, g \in L^1(\mathbb{T})$. Then $f * g \in L^1(\mathbb{T})$ and*

$$\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n) \quad \textit{for every } n \in \mathbf{Z}. \tag{2.5}$$

Proof. We prove (2.5) by changing the order of integration and using the fact that f has period 2π:

$$\begin{aligned}\widehat{f * g}(n) &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \left(\int_{-\pi}^{\pi} f(t-s) g(s)\, ds \right) e^{-int}\, dt \\ &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \left(\int_{-\pi}^{\pi} f(t-s) e^{-in(t-s)}\, dt \right) g(s) e^{-ins}\, ds \\ &= \hat{f}(n)\hat{g}(n).\end{aligned}$$

■

The first part of the next proposition shows that the finite Fourier transform maps $L^1(\mathbb{T})$ into ℓ^∞ (the space of bounded sequences of complex numbers), while the second shows that that the image of $L^1(\mathbb{T})$ is a subset of $\mathbf{c}_0$ (the space of sequences of complex numbers that tend to 0 at $\pm\infty$). In Example 2.10.4, we will show that the last inclusion in fact is proper. We will refer to second property in the proposition as the **Riemann–Lebesgue lemma**.

Proposition 2.4.3. *Suppose that $f \in L^1(\mathbb{T})$. Then the following properties hold*:

(i) $|\hat{f}(n)| \leqslant \|f\|_1$ *for every* $n \in \mathbf{Z}$;

(ii) $\hat{f}(n) \to 0$ *as* $n \to \pm\infty$.

Proof. The first property follows directly from the definition of $\hat{f}(n)$. To prove the second property, notice that

$$\begin{aligned}2\pi \hat{f}(n) &= \int_{-\pi}^{\pi} f(t) e^{-int}\, dt = -\int_{-\pi}^{\pi} f(t) e^{-in(t+\pi/n)}\, dt \\ &= -\int_{-\pi}^{\pi} f(t-\pi/n) e^{-int}\, dt,\end{aligned}$$

so that

$$\widehat{f}(n) = \frac{1}{4\pi}\int_{-\pi}^{\pi} (f(t) - f(t-\pi/n))e^{-int}\,dt. \tag{2.6}$$

It now follows from (2.6) and Lemma 1.4.1 that

$$|\widehat{f}(n)| \leqslant \tfrac{1}{2}\|f - \tau_{\pi/n}f\|_1 \longrightarrow 0 \quad \text{as } n \to \pm\infty. \qquad \blacksquare$$

Remark 2.4.4. Notice that if $f \in L^1(\mathbb{T})$, then

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt - \frac{i}{2\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt.$$

This identity together with the Riemann–Lebesgue lemma show that if f is real-valued, then both integrals in the right-hand side tend to 0 as $n \to \pm\infty$. By splitting a complex-valued function into its real and imaginary parts, we see that this is also true in general.

According to the Riemann–Lebesgue lemma, $\widehat{f}(n) = o(1)$ as $n \to \pm\infty$ for every function $f \in L^1(\mathbb{T})$. We now show that if f has additional regularity, then $\widehat{f}(n)$ will decay faster. The main tool used is integration by parts. The largest class of functions, for which it is possible to integrate by parts, is the class of absolutely continuous functions.

Definition 2.4.5. Denote by $AC(\mathbb{T})$ the class of absolutely continuous functions on $\mathbf{R}$ with period 2π.

Proposition 2.4.6. *Suppose that $f \in C^{k-1}(\mathbb{T})$ and $f^{(k-1)} \in AC(\mathbb{T})$, where $k \geqslant 1$. Then*

$$\widehat{f^{(k)}}(n) = (in)^k \widehat{f}(n), \quad n \in \mathbf{Z}. \tag{2.7}$$

Moreover, $\widehat{f}(n) = o(n^{-k})$ as $n \to \pm\infty$, i.e., $\lim_{n\to\pm\infty} n^k \widehat{f}(n) = 0$.

Proof. The identity (2.7) follows by integrating the left-hand side k times by parts using the fact that f is periodic:

$$\begin{aligned}\widehat{f^{(k)}}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f^{(k)}(t)e^{-int}\,dt \\ &= \frac{1}{2\pi}\big(f^{(k-1)}(\pi)e^{-in\pi} - f^{(k-1)}(-\pi)e^{in\pi}\big) + \frac{in}{2\pi}\int_{-\pi}^{\pi} f^{(k-1)}(t)e^{-int}\,dt \\ &= \ldots = \frac{(in)^k}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt = (in)^k\widehat{f}(n).\end{aligned}$$

Since $f^{(k)} \in L^1(\mathbb{T})$, this formula together with the Riemann–Lebesgue lemma shows that $\widehat{f}(n) = o(n^{-k})$ as $n \to \pm\infty$. ■

Remark 2.4.7. The assertions in the proposition of course hold true if $f \in C^k(\mathbb{T})$.

Definition 2.4.8. Suppose that the function f is defined on an interval $I \subset \mathbf{R}$. We say that f satisfies a **Hölder condition** at a point $t \in I$ if there exist constants $C \geqslant 0$, $\alpha > 0$, and $\delta > 0$ such that

$$|f(s) - f(t)| \leqslant C|s-t|^{\alpha} \quad \text{for every } s \in I \text{ satisfying } |s-t| \leqslant \delta.$$

If f satisfies a Hölder condition at every $t \in I$ with the same constants C and α, and if δ can be taken as the length of I, then we say that f is **Hölder continuous**.

When $\alpha = 1$, one usually uses the terms **Lipschitz condition** and **Lipschitz continuous**. Notice that if f satisfies a Hölder condition at t, then f is continuous at t, and if f is Hölder continuous, then f is also uniformly continuous.

Example 2.4.9.

(a) The function $f(t) = \sqrt{|t|}$, $t \in \mathbf{R}$, is Hölder continuous on $\mathbf{R}$ with exponent $\frac{1}{2}$:

$$\left|\sqrt{|s|} - \sqrt{|t|}\right| \leqslant \sqrt{|s-t|} \quad \text{for } s, t \in \mathbf{R}.$$

(b) If f is differentiable on an interval I and $|f'(t)| \leqslant C$ for every $t \in I$, then f is Lipschitz continuous on I; this follows directly from the mean value theorem:

$$|f(s) - f(t)| = |f'(\xi)||s-t| \leqslant C|s-t| \quad \text{for } s, t \in I,$$

where ξ is some point between s and t. □

Definition 2.4.10. Denote by $\Lambda_\alpha(\mathbb{T})$ the class of Hölder continuous functions on $\mathbf{R}$ with period 2π. The norm of $f \in \Lambda_\alpha(\mathbb{T})$ is given by

$$\|f\|_{\Lambda_\alpha(\mathbb{T})} = \sup_{s,t \in \mathbf{R},\ s \neq 0} \frac{|f(s) - f(t)|}{|s-t|^{\alpha}}.$$

Notice that if $\alpha > 1$, then $\Lambda_\alpha(\mathbb{T})$ contains only constants. The next result is a direct consequence of (2.6).

Corollary 2.4.11. *Suppose that $f \in \Lambda_\alpha(\mathbb{T})$. Then there exists a constant $C \geqslant 0$ such that*

$$|\widehat{f}(n)| \leqslant C\|f\|_{\Lambda_\alpha(\mathbb{T})}|n|^{-\alpha} \quad \textit{for every } n \neq 0.$$

Definition 2.4.12. Denote by $BV(\mathbb{T})$ the class of 2π-periodic functions on $\mathbf{R}$ that are of bounded variation.

Proposition 2.4.13. *Suppose that $f \in BV(\mathbb{T})$. Then*

$$|\widehat{f}(n)| \leqslant \frac{\mathrm{V}(f)}{2\pi|n|} \quad \textit{for every } n \neq 0. \tag{2.8}$$

Proof. The inequality (2.8) follows by integration by parts:

$$|\widehat{f}(n)| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt\right| = \left|\frac{1}{2\pi n}\int_{-\pi}^{\pi} e^{-int}\,df(t)\right| \leqslant \frac{\mathrm{V}(f)}{2\pi|n|}.$$ ■

Remark 2.4.14. In Proposition 2.4.3, we saw that the Fourier coefficients of a function $f \in L^1(\mathbb{T})$ belong to $\mathbf{c}_0$, i.e., $\widehat{f}(n) \to 0$ as $n \to \pm\infty$. It is natural to ask if anything more can be said about the rate of convergence of $\widehat{f}(n)$. This is, in fact, not possible; one can show that the Fourier coefficients of a L^1-function can tend to 0 arbitrarily slowly. To be more precise, if $(c_n)_{n=0}^{\infty}$ is a sequence of nonnegative numbers, such that $\lim_{n\to\infty} c_n = 0$, that satisfies the following convexity condition:

$$c_{n+1} + c_{n-1} \geqslant 2c_n \quad \text{for } n \geqslant 1,$$

then there exists a function $f \in L^1(\mathbb{T})$ such that $\widehat{f}(n) = c_{|n|}$ for every $n \in \mathbf{Z}$.

2.5. Partial Sums of Fourier Series

In the next section, we will prove a number of criteria for pointwise convergence of Fourier series. As a preparation, we now study the partial sums of the Fourier series for a function $f \in L^1(\mathbb{T})$. Denote by $S_N f$ the N-th symmetric partial sum of the series (2.1), that is

$$S_N f(t) = \sum_{n=-N}^{N} \widehat{f}(n) e^{int}, \quad N = 0, 1, \dots .$$

We rewrite $S_N f(t)$ as a convolution in the following way:

$$\begin{aligned} S_N f(t) &= \sum_{n=-N}^{N} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) e^{-ins}\, ds \right) e^{int} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \left(\sum_{n=-N}^{N} e^{in(t-s)} \right) ds \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) D_N(t-s)\, ds = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t-s) D_N(s)\, ds \qquad (2.9) \\ &= D_N * f(t), \end{aligned}$$

where D_N is the **Dirichlet kernel**:

$$D_N(t) = \sum_{n=-N}^{N} e^{int}, \quad t \in \mathbf{R},\ N = 0, 1, \dots . \qquad (2.10)$$

The next proposition summarizes some of the most important properties of the Dirichlet kernel.

Proposition 2.5.1. *The following properties hold for the Dirichlet kernel D_N:*

(i) $D_N(t) = \begin{cases} \dfrac{\sin(N + \frac{1}{2})t}{\sin \frac{t}{2}} & \text{for } t \in \mathbf{R} \setminus 2\pi\mathbf{Z} \\ 2N+1 & \text{for } t \in 2\pi\mathbf{Z} \end{cases}$;

(ii) D_N *is even*;

(iii) $\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} D_N(t)\, dt = 1.$

Proof. The first property follows by summing the geometric series in (2.10), the second is obvious, while the third is obtained by integrating both sides of (2.10) over $(-\pi, \pi)$. ∎

Let us return to our investigation of the partial sums of a function $f \in L^1(\mathbb{T})$. Since D_N is even, we have that

$$S_N f(t) = \frac{1}{2\pi}\int_0^\pi (f(t+s)+f(t-s))\frac{\sin(N+\frac{1}{2})s}{\sin\frac{s}{2}}\,ds. \tag{2.11}$$

Suppose now that $0 < \delta < \pi$. Then

$$\begin{aligned} S_N f(t) &= \frac{1}{\pi}\int_0^\delta \frac{f(t+s)+f(t-s)}{s}\sin(N+\tfrac{1}{2})s\,ds \\ &+\frac{1}{\pi}\int_\delta^\pi \frac{f(t+s)+f(t-s)}{s}\sin(N+\tfrac{1}{2})s\,ds \\ &+\frac{1}{2\pi}\int_0^\pi (f(t+s)+f(t-s))\left(\frac{1}{\sin\frac{s}{2}}-\frac{2}{s}\right)\sin(N+\tfrac{1}{2})s\,ds. \end{aligned} \tag{2.12}$$

The second integral in (2.12) tends to 0 as $N \to \infty$ according to the Riemann–Lebesgue lemma (see Remark 2.4.4) since the integrand belongs to $L^1(\delta, \pi)$. This also applies to the last integral since the function

$$g(s) = \frac{1}{\sin\frac{s}{2}} - \frac{2}{s}, \quad 0 < s \leqslant \pi,$$

is bounded (because g is continuous and $g(s) \to 0$ as $s \to 0$).[1] We thus have the following asymptotic representation for $S_N f(t)$.

Proposition 2.5.2. *Suppose that $f \in L^1(\mathbb{T})$ and $0 < \delta < \pi$. Then*

$$S_N f(t) = \frac{1}{\pi}\int_0^\delta \frac{f(t+s)+f(t-s)}{s}\sin(N+\tfrac{1}{2})s\,ds + \varepsilon_N(t) \tag{2.13}$$

for every $t \in \mathbf{R}$, where $\varepsilon_N(t) \to 0$ as $N \to \infty$.

Taking $f = 1$ in (2.13), we obtain that

$$1 = \frac{2}{\pi}\int_0^\delta \frac{\sin(N+\frac{1}{2})s}{s}\,ds + \varepsilon_N, \tag{2.14}$$

where $\varepsilon_N \to 0$ as $N \to \infty$. By combining (2.13) with (2.14), we obtain a necessary and sufficient condition for the convergence of the Fourier series of f at a point t.

Proposition 2.5.3. *Suppose that $f \in L^1(\mathbb{T})$. Then $\lim_{N\to\infty} S_N f(t) = S$ if and only there exists a number $\delta > 0$ such that*

$$\lim_{N\to\infty}\int_0^\delta \frac{f(t+s)+f(t-s)-2S}{s}\sin(N+\tfrac{1}{2})s\,ds = 0. \tag{2.15}$$

Proof. Multiply (2.14) with S and subtract from (2.13):

$$S_N f(t) - S = \frac{1}{\pi}\int_0^\delta \frac{f(t+s)+f(t-s)-2S}{s}\sin(N+\tfrac{1}{2})s\,ds + (\varepsilon_N(t) - S\varepsilon_N).$$

Then use the fact that $\varepsilon_N(t) - S\varepsilon_N \to 0$ as $N \to \infty$. ∎

[1] One can in fact show that $|g(s)| \leqslant \pi^2/24$ for $0 \leqslant s \leqslant \pi$.

2.6. Criteria for Pointwise Convergence

We now establish a number of corollaries to Theorem 2.5.3. The first is Dini's classical criterion.

Corollary 2.6.1 (Dini's Criterion). *Suppose that* $f \in L^1(\mathbb{T})$ *satisfies a* **Dini condition** *at* $t \in \mathbf{R}$, *meaning that there exist numbers* $\delta > 0$ *and* $S \in \mathbf{C}$ *such that*

$$\int_0^\delta \frac{|f(t+s) + f(t-s) - 2S|}{s}\, ds < \infty.$$

Then $\lim_{N\to\infty} S_N f(t) = S$.

In particular, if

$$\int_0^\delta \frac{|f(t+s) + f(t-s) - 2f(t)|}{s}\, ds < \infty \tag{2.16}$$

for some $\delta > 0$, then $\lim_{N\to\infty} S_N f(t) = f(t)$.

Proof (Corollary 2.6.1). The quotient in (2.15) belongs by the assumption to the space $L^1(0, \delta)$. The assertion therefore follows from the Riemann–Lebesgue lemma. ∎

The next corollary is the convergence criterion one usually meets in introductory courses in Fourier analysis.

Corollary 2.6.2. *Suppose that* $f \in L^1(\mathbb{T})$. *If the one-sided limits*

$$f(t^+) = \lim_{s\to 0^+} f(t+s) \quad \text{and} \quad f(t^-) = \lim_{s\to 0^+} f(t-s)$$

and the one-sided derivatives

$$f'(t^+) = \lim_{s\to 0^+} \frac{f(t+s) - f(t^+)}{s} \quad \text{and} \quad f'(t^-) = \lim_{s\to 0^+} \frac{f(t-s) - f(t^-)}{-s}$$

exist, then

$$\lim_{N\to\infty} S_N f(t) = \frac{f(t^+) + f(t^-)}{2}. \tag{2.17}$$

Proof. Let S denote the right-hand side of (2.17). Then the quotient in (2.15) is bounded for every $\delta > 0$. ∎

Example 2.6.3. If we apply the result in Corollary 2.6.2 to the function in Example 2.2.2, we see that

$$t = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nt \quad \text{for } -\pi < t < \pi.$$

For $t = \pm\pi$, the series equals 0, which is in accordance with the corollary. □

Corollary 2.6.4. *Suppose that* $f \in L^1(\mathbb{T})$. *If* f *satisfies a Hölder condition at a point* $t \in \mathbf{R}$, *then* $\lim_{N\to\infty} S_N f(t) = f(t)$.

Proof. The assumption means that there exist numbers $C \geqslant 0$, $\alpha > 0$, and $\delta > 0$ such that

$$|f(t+s) - f(t)| \leqslant C|s|^\alpha \quad \text{for } |s| < \delta.$$

This implies that the integrand in (2.16) is bounded by $g(s) = 2Cs^{\alpha-1}$, $0 < s < \delta$, which is a integrable function on $(0, \delta)$. ■

Example 2.6.5. Let $f \in C(\mathbb{T})$ be defined by $f(t) = \sqrt{|t|}$ for $-\pi \leqslant t \leqslant \pi$. Notice that we cannot apply Corollary 2.6.2 to show that the Fourier series of f is convergent at $t = 0$ since both one-sided derivatives are infinite. However, f satisfies a Hölder condition at 0:

$$|f(s) - f(0)| = \sqrt{|s|} = |s - 0|^{1/2} \quad \text{for } -\pi \leqslant s \leqslant \pi,$$

so the Fourier series of f converges to 0 at $t = 0$. □

In the proof of our next result, we will use the Si function:

$$\operatorname{Si}(t) = \int_0^t \frac{\sin\tau}{\tau}\,d\tau, \quad 0 \leqslant t < \infty.$$

The following lemma is often proved using calculus of residues. We will, however, give a proof that uses techniques from this chapter.

Lemma 2.6.6. *There holds* $\lim_{t\to\infty} \operatorname{Si}(t) = \frac{\pi}{2}$.

Proof. Using integration by parts, we see that if $t \geqslant 1$, then

$$\operatorname{Si}(t) = \int_0^1 \frac{\sin\tau}{\tau}\,d\tau + \cos 1 - \frac{\cos t}{t} - \int_1^t \frac{\cos\tau}{\tau^2}\,d\tau.$$

Moreover, since the integral $\int_1^\infty \tau^{-2}\cos\tau\,d\tau$ is absolutely convergent, it follows that the limit $\lim_{t\to\infty}\operatorname{Si}(t)$ exists. From (2.14), we also have that

$$\int_0^\delta \frac{\sin(N+\frac{1}{2})s}{s}\,ds = \frac{\pi}{2} + \varepsilon_N,$$

where $\varepsilon_N \to 0$ as $N \to \infty$. The claim now follows if we change variables in the last integral and let $N \to \infty$:

$$\frac{\pi}{2} = \lim_{N\to\infty}\int_0^\delta \frac{\sin(N+\frac{1}{2})s}{s}\,ds = \lim_{N\to\infty}\int_0^{(N+\frac{1}{2})\delta} \frac{\sin\tau}{\tau}\,d\tau = \int_0^\infty \frac{\sin\tau}{\tau}\,d\tau.$$ ■

The following convergence criterion for functions of bounded variation was proved by C. Jordan in 1881.

Theorem 2.6.7. *Suppose that $f \in L^1(\mathbb{T})$. If f is of bounded variation on an interval $[t-\delta, t+\delta]$ for some $\delta > 0$, then*

$$\lim_{N\to\infty} S_N f(t) = \frac{f(t^+) + f(t^-)}{2}.$$

Proof. Put $F(s) = \frac{1}{2}(f(t+s) + f(t-s))$ for $|s| \leqslant \delta$, $S = F(0^+)$, and $m = N + \frac{1}{2}$. Then

$$\begin{aligned}\int_0^\delta \frac{f(t+s)+f(t-s)-2S}{s}\sin(N+\tfrac{1}{2})s\,ds &= \int_0^\delta (F(s)-S)\,d\mathrm{Si}(ms)\\ &= (F(\delta^-)-S)\mathrm{Si}(m\delta)\\ &\quad - \int_0^\delta \mathrm{Si}(ms)\,dF(s).\end{aligned}$$

If we now use the fact that $\mathrm{Si}(ms) \to \frac{\pi}{2}$ as $m \to \infty$ and the dominated convergence theorem, we obtain that

$$\begin{aligned}\lim_{N\to\infty}\int_0^\delta \frac{f(t+s)+f(t-s)-2S}{s}\sin(N+\tfrac{1}{2})s\,ds &= \frac{\pi}{2}(F(\delta^-)-S) - \frac{\pi}{2}\int_0^\delta dF(s)\\ &= 0.\end{aligned}$$ ■

Since every absolutely continuous function is of bounded variation, we have the following corollary.

Corollary 2.6.8. *Suppose that $f \in AC(\mathbb{T})$. Then*

$$\lim_{N\to\infty} S_N f(t) = f(t) \quad \textit{for every } t \in \mathbf{R}.$$

2.7. The Riemann Localization Principle

Although the Fourier coefficients of a function $f \in L^1(\mathbb{T})$ depend on the global behaviour of f, the convergence of the Fourier series of f at a point in fact only depends on the behaviour of f in an arbitrarily small neighbourhood of the point. This is the content of the following theorem, known as the **Riemann localization principle**.

Theorem 2.7.1. *Suppose that $f, g \in L^1(\mathbb{T})$. If $f = g$ in a neighbourhood of a point $t_0 \in \mathbf{R}$, then the Fourier series of f and g either both converge to the same value or both diverge.*

Proof. Suppose that $f(t) = g(t)$ for $|t - t_0| < \delta$. Then, according to (2.13),

$$\begin{aligned}S_N f(t_0) &= \frac{1}{\pi}\int_0^\delta \frac{f(t_0+s)+f(t_0-s)}{s}\sin(N+\tfrac{1}{2})s\,ds + o(1)\\ &= S_N g(t_0) + o(1).\end{aligned}$$ ■

2.8. A Uniqueness Theorem for Fourier Series

The following theorem shows that the Fourier coefficients determine a function completely. Notice that we do not assume that the Fourier series are convergent; if the series converge to the involved functions, the result is of course obvious.

Theorem 2.8.1. *Suppose that $f, g \in L^1(\mathbb{T})$ and $\widehat{f}(n) = \widehat{g}(n)$ for every $n \in \mathbf{Z}$. Then $f = g$ a.e.*

Proof. By the linearity of the Fourier coefficients, we may assume that $g = 0$. First put $F(t) = \int_{-\pi}^{t} f(\tau)\,d\tau + C,\ t \in \mathbf{R}$, where C is chosen so that $\widehat{F}(0) = 0$, i.e., C has to satisfy the equation

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F(t)\,dt + C = 0.$$

The function F has period 2π since

$$F(t+2\pi) - F(t) = \int_{t}^{t+2\pi} f(\tau)\,d\tau = \int_{-\pi}^{\pi} f(\tau)\,d\tau = 2\pi\widehat{f}(0) = 0$$

for every $t \in \mathbf{R}$. Then put $G(t) = \int_{-\pi}^{t} F(s)\,ds,\ t \in \mathbf{R}$. Since $\widehat{F}(0) = 0$, G also has period 2π. It now follows from Proposition 2.4.6 that

$$(in)^2\widehat{G}(n) = \widehat{G''}(n) = \widehat{f}(n) = 0,$$

so $\widehat{G}(n) = 0$ for every $n \neq 0$. Corollary 2.6.2 now shows that $G(t) = \widehat{G}(0)$ for every $t \in \mathbf{R}$. Differentiating this identity twice, we obtain that $f = 0$ a.e. ■

2.9. Uniform Convergence of Fourier Series

We next consider uniform convergence of Fourier series. Suppose first that the Fourier series $\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{int}$ of a function $f \in L^1(\mathbb{T})$ is absolutely convergent:

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty.$$

It then follows from the Weierstrass M-test that the Fourier series converges uniformly, and hence that its sum, which we denote $g(t)$, is a continuous function. Integrating the series termwise, which is allowed because it converges uniformly, we see that $\widehat{g}(n) = \widehat{f}(n)$ for every $n \in \mathbf{Z}$. The uniqueness theorem (Theorem 2.8.1) therefore shows that $g = f$ a.e. In particular, the Fourier series of f converges to f a.e. The following theorem summarizes these observations.

Theorem 2.9.1. *Suppose that $f \in L^1(\mathbb{T})$. If the Fourier series of f is absolutely convergent, then the series converges uniformly to a function belonging to $C(\mathbb{T})$, which coincides with f a.e. In particular, the Fourier series of f converges to f a.e. and everywhere if f is continuous.*

For instance, if $f \in C^2(\mathbb{T})$, then $\widehat{f}(n) = o(n^{-2})$ as $n \to \pm\infty$ according to Theorem 2.4.6, so $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$, and we can apply Theorem 2.9.1 to conclude that the Fourier series of f converges uniformly to f. It is possible to obtain precise information about the rate of convergence of the series. Indeed, for every $\varepsilon > 0$, there exists a number $M \geqslant 0$ such that

$$|\widehat{f}(n)| \leqslant \frac{\varepsilon}{n^2} \quad \text{if } |n| \geqslant M.$$

Now, if $N \geqslant M$, then

$$|f(t) - S_N f(t)| = \left|\sum_{|n|\geqslant N+1} \widehat{f}(n)e^{int}\right| \leqslant 2\varepsilon \sum_{n=N+1}^{\infty} \frac{1}{n^2} \leqslant \int_{N}^{\infty} \frac{d\tau}{\tau^2} = \frac{2\varepsilon}{N}$$

for every $t \in \mathbf{R}$. It follows that $\|f - S_N f\|_\infty = o(N^{-1})$ as $N \to \infty$.

Theorem 2.9.2. *If $f \in C^2(\mathbb{T})$, then $\|f - S_N f\|_\infty = o(N^{-1})$ as $N \to \infty$. In particular, the Fourier series of f converges uniformly to f.*

The result in the next theorem is much stronger than the previous one. As expected, the proof is harder.

Theorem 2.9.3. *Suppose that $f \in L^1(\mathbb{T})$ is Hölder continuous on (a, b). Then the Fourier series of f converges uniformly to f on every interval $(c, d) \subset (a, b)$ such that $\overline{(c, d)} \subset (a, b)$.*

In particular, if f is Hölder continuous on $\mathbf{R}$, then the Fourier series of f converges uniformly to f on $\mathbf{R}$.

Notce that if f is Hölder continuous, then $\widehat{f}(n) = O(n^{-\alpha})$ as $n \to \pm\infty$ for some number $\alpha > 0$ according to Corollary 2.4.11, so just looking at the Fourier coefficients, it is not at all obvious that the Fourier series should converge uniformly (or even pointwise). We will use the following definition and lemma.

Definition 2.9.4. A sequence $(g_n)_{n=1}^\infty$ of functions on a set $E \subset \mathbf{R}$ is said to be **equicontinuous** if the following condition is satisfied: For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $s, t \in E$ and $|s - t| < \delta$, then

$$|g_n(s) - g_n(t)| < \varepsilon \quad \text{for every } n \geqslant 1.$$

To put it differently, a sequence is equicontinuous if it is uniformly continuous, where the continuity is uniform both with respect to the variable and the index.

Lemma 2.9.5. *Suppose that $(g_n)_{n=1}^\infty$ is a sequence of functions on a compact set $K \subset \mathbf{R}$. If $g_n(t) \to 0$ as $n \to \infty$ for every $t \in K$ and $(g_n)_{n=1}^\infty$ is equicontinuous, then $(g_n)_{n=1}^\infty$ converges uniformly to 0 on K.*

Proof. The proof proceeds by contradiction. Suppose that there exists a number $\varepsilon > 0$, indices $n_1 < n_2 < \dots$, and points $t_1, t_2, \dots \in K$ such that

$$|g_{n_k}(t_k)| \geqslant \varepsilon \quad \text{for } k = 1, 2, \dots .$$

By compactness, there exists a subsequence to $(t_k)_{k=1}^\infty$, which we may assume is the whole sequence, that converges to some point $t_0 \in K$. We then have

$$\varepsilon \leqslant |g_{n_k}(t_k)| \leqslant |g_{n_k}(t_k) - g_{n_k}(t_0)| + |g_{n_k}(t_0)|.$$

This is a contradiction since the right-hand side can be made arbitrarily small by choosing k sufficiently large. ■

Proof (Theorem 2.9.3). The assumption means that

$$|f(t) - f(u)| \leqslant C|t - u|^\alpha \quad \text{for all } t, u \in (a, b).$$

For $N = 1, 2, \dots$, put $g_N(t) = S_N f(t) - f(t)$, $c < t < d$. Since we know (Corollary 2.6.4) that $g_N(t)$ converges to 0 as $N \to \infty$ for $c < t < d$, it suffices to show that the sequence $(g_N)_{N=1}^\infty$ is equicontinuous. Let $\varepsilon > 0$ be arbitrary. From (2.11) and (iii) in Proposition 2.5.1, we have

$$g_N(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(t + s) - f(t)) \frac{\sin{(N + \frac{1}{2})s}}{\sin \frac{s}{2}}\, ds \quad \text{for } c < t < d.$$

It follows that if $c < t, u < d$, then

$$|g_N(t) - g_N(u)| \leqslant \frac{1}{2\pi}\int_{|s|<\eta} \frac{|f(t+s)-f(t)| + |f(u+s)-f(u)|}{|\sin\frac{s}{2}|}\,ds + \frac{1}{2\pi}\int_{\eta\leqslant|s|<\pi} \frac{|f(t+s)-f(u+s)| + |f(t)-f(u)|}{|\sin\frac{s}{2}|}\,ds,$$

where $0 < \eta < \pi$ satisfies $\eta \leqslant \min(c-a, b-d)$. Using this ineqality and the fact that $|\sin s| \geqslant \frac{2}{\pi}|s|$ for $|s| \leqslant \frac{\pi}{2}$, we obtain

$$|g_N(t) - g_N(u)| \leqslant C\eta^\alpha + C\eta^{-1}|t-u|^\alpha.$$

Finally choose η so small that the first term in the right-hand side is less than $\varepsilon/2$ and then δ so small that the second term is less than $\varepsilon/2$ whenever $|t-u| < \delta$. ■

Corollary 2.9.6. *Suppose that $f \in C(\mathbb{T})$ with piecewise continuous derivative. Then the Fourier series of f converges uniformly to f on $\mathbf{R}$.*

In Theorem 4.5.1, we will show that the convergence is also absolute.

Proof (Corollary 2.9.6). It follows from the assumption, that there exist points

$$-\pi = t_1 < t_2 < ... < t_n = \pi$$

such that f is continuously differentiable on each interval $[t_i, t_{i+1}]$, $1 = 1, 2, \dots, n-1$. But then f is Lipschitz continuous on every interval $[t_i, t_{i+1}]$. This implies that f is Lipschitz continuous on $[-\pi, \pi]$ and therefore on $\mathbf{R}$. ■

2.10. Termwise integration of Fourier Series

A quite surprising result is the fact that the Fourier series of a L^1-function may be integrated termwise and the resulting series is pointwise convergent everywhere (even uniformly convergent), irrespective if the original series is convergent or not. Suppose that $f \in L^1(\mathbb{T})$. Then the function

$$F(t) = \int_{-\pi}^{t} (f(\tau) - \hat{f}(0))\,d\tau, \quad t \in \mathbf{R},$$

is absolutey continuous and satisfies $F' = f$ a.e. Moreover, F has period 2π:

$$F(t+2\pi) - F(t) = \int_{t}^{t+2\pi} (f(\tau) - \hat{f}(0))\,d\tau = \int_{-\pi}^{\pi} f(\tau)\,d\tau - 2\pi\hat{f}(0) = 0$$

for every $t \in \mathbf{R}$. According to Proposition 2.4.6, $\hat{f}(n) = in\hat{F}(n)$ for every $n \neq 0$. It now follows from Corollary 2.6.8 that

$$F(t) = \hat{F}(0) + \sum_{n\neq 0} \frac{\hat{f}(n)}{in} e^{int} \quad \text{for every } t \in \mathbf{R}. \tag{2.18}$$

In Theorem 4.5.1, we will show that the series in (2.18) actually converges uniformly on **R**. Now, if $-\infty < s < t < \infty$, then

$$F(t) - F(s) = \sum_{n \neq 0} \frac{\widehat{f}(n)}{in}(e^{int} - e^{ins}) = \sum_{n \neq 0} \widehat{f}(n) \int_s^t e^{in\tau}\, d\tau,$$

so that

$$\int_s^t f(\tau)\, d\tau = \widehat{f}(0)(t-s) + \sum_{n \neq 0} \widehat{f}(n) \int_s^t e^{in\tau}\, d\tau = \sum_{n=-\infty}^{\infty} \widehat{f}(n) \int_s^t e^{in\tau}\, d\tau.$$

Theorem 2.10.1. *Suppose that $f \in L^1(\mathbb{T})$. Then*

$$\int_s^t f(\tau)\, d\tau = \sum_{n=-\infty}^{\infty} \widehat{f}(n) \int_s^t e^{in\tau}\, d\tau \quad \textit{for } -\infty < s < t < \infty. \tag{2.19}$$

Formally, the equation (2.19) may be written

$$\int_s^t \left(\sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{in\tau} \right) d\tau = \sum_{n=-\infty}^{\infty} \widehat{f}(n) \int_s^t e^{in\tau}\, d\tau.$$

Notice also that it follows from (2.19) that

$$\int_s^t f(\tau)\, d\tau = \lim_{N \to \infty} \int_s^t S_N f(\tau)\, d\tau.$$

This fact, however, does not imply that $S_N f \to f$ in $L^1(\mathbb{T})$.

Example 2.10.2. From Example 2.6.3, we know that

$$t = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nt \quad \text{for } -\pi < t < \pi.$$

Integrating this identity from 0 to t, we obtain that

$$\frac{t^2}{2} = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} + 2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nt \quad \text{for } -\pi \leqslant t \leqslant \pi.$$

To evaluate the first series in the right-hand side, we integrate both sides once more, this time from $-\pi$ to π:

$$\frac{\pi^3}{3} = 4\pi \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}, \quad \text{which shows that} \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}.$$

We have thus shown that

$$t^2 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nt \quad \text{for } -\pi \leqslant t \leqslant \pi.$$

Putting $t = \pi$ in this identity, we obtain that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}. \qquad \square$$

The following corollary is a consequence to (2.18).

Corollary 2.10.3. *Suppose that $f \in L^1(\mathbb{T})$. Then the series*

$$\sum_{n \neq 0} \frac{\hat{f}(n)}{n} e^{int}$$

is convergent for every $t \in \mathbf{R}$.

Example 2.10.4. In Example 2.3.3, we saw that the trigonometric series

$$\sum_{n=2}^{\infty} \frac{\sin nt}{\ln n} \tag{2.20}$$

is convergent for every $t \in \mathbf{R}$. However, since

$$\sum_{n=2}^{\infty} \frac{1}{n \ln n} = \infty,$$

this is not the Fourier series of any function belonging to $L^1(\mathbb{T})$. It is not so hard to show that the function, defined by (2.20), in fact does not belong to $L^1(\mathbb{T})$. Notice that this also shows that the image of $L^1(\mathbb{T})$ under the finite Fourier transform is not the whole of $\mathbf{c}_0$. □

2.11. Divergence of Fourier Series

Let us end this chapter with a few comments and results about divergence of Fourier series. The first convergence criterion for Fourier series was proved by L. Dirichlet in 1829. Dirichlet and many others in this period seem to have believed that the Fourier series of a continuous function converges to the function everywhere. In 1873, P. du Bois-Reymond however proved that there exists a continuous function whose Fourier series diverges on a dense subset to $\mathbf{R}$. Dirichlet's construction was later simplified by L. Fejér in 1909. In 1923, A. Kolmogorov proved that there even exists a L^1-function (although not continuous) whose Fourier series diverges *everywhere*. It was therefore not unreasonable to expect that there could exist a continuous function with an everywhere divergent Fourier series.

On the other hand, N. Lusin conjectured in 1915 that the Fourier series of a L^2-function and, in particular, of a continuous function, converges a.e. Lusin's conjecture was proved by L. Carleson as late as 1966. According to Carleson's theorem, the Fourier series of a continuous thus converges a.e. Carlesons result was generalized in 1968 by R. A. Hunt to L^p for $1 < p < \infty$, and a new proof of Carleson's theorem was given by C. Fefferman in 1973. In this connection, we should mention that J.-P. Kahane and Y. Katznelson in 1966 showed that, for any set $E \subset \mathbf{R}$ with measure 0, there exists a continuous function whose Fourier series diverges at every point of E.

We will here prove that there exists a continuous function with the the property that the Fourier series of the function diverges at one point. Although the existence of such a function can be proved constructively, we prefer to use a "soft" argument, which is due to Kolmogorov, based on the Banach–Steinhaus theorem which we state without a proof.

Theorem 2.11.1 (Banach–Steinhaus). *Suppose that X is a Banach space, Y is a normed linear space, and $(T_n)_{n=1}^{\infty}$ is a sequence of bounded, linear operators from X to Y. Then either there exists a constant C such that*

$$\|T_n\| \leqslant C \quad \text{for every } n \geqslant 1$$

or

$$\sup_{n\geqslant 1} \|T_n x\| = \infty$$

for every x that belongs to a dense G_δ set in X.

Theorem 2.11.2. *There exists a function in $C(\mathbb{T})$ whose Fourier series diverges at a point.*

In the proof, we will use the following notation: For $f \in C(\mathbb{T})$, put

$$S^* f(t) = \sup_{N\geqslant 1} |S_N f(t)|, \quad t \in \mathbf{R}.$$

Since every convergent sequence is bounded, it is obvious that the Fourier series of f diverges at t if $S^* f(t) = \infty$.[2] We also need to know something about the norms of the Dirichlet kernels. Put $L_N = \|D_N\|_1$ for $N = 0, 1, \dots$. The numbers L_N are known as the **Lebesgue constants**.

Lemma 2.11.3. *The Lebesgue constants have the following asymptotics*:

$$L_N = \frac{4}{\pi^2} \ln N + O(1) \quad \text{as } N \to \infty. \tag{2.21}$$

The asymptotic behavior of the Lebesgue constants was first investigated by L. Fejér in 1910. In the proof, presented below, we use a technique due to L. Lorch (1954).

Proof. Using the fact that $(\sin \frac{s}{2})^{-1} - 2s^{-1}$ is bounded for $0 < s \leqslant \pi$, we see that

$$\begin{aligned} L_N &= \frac{1}{\pi} \int_0^\pi \frac{|\sin (N + \frac{1}{2})s|}{\sin \frac{s}{2}}\, ds \\ &= \frac{2}{\pi} \int_0^\pi \frac{|\sin (N + \frac{1}{2})s|}{s}\, ds + \frac{1}{\pi} \int_0^\pi |\sin (N + \tfrac{1}{2})s| \Big(\frac{1}{\sin \frac{s}{2}} - \frac{2}{s}\Big)\, ds \\ &= \frac{2}{\pi} \int_0^\pi \frac{|\sin (N + \frac{1}{2})s|}{s}\, ds + O(1). \end{aligned}$$

We then change variables in the last integral and split it into two parts:

$$\begin{aligned} L_N &= \frac{2}{\pi} \int_0^{(N+\frac{1}{2})\pi} \frac{|\sin s|}{s}\, ds + O(1) \\ &= \frac{2}{\pi} \int_1^{(N+\frac{1}{2})\pi} \frac{|\sin s|}{s}\, ds + \frac{2}{\pi} \int_0^1 \frac{|\sin s|}{s}\, ds + O(1) \\ &= \frac{2}{\pi} \int_1^{(N+\frac{1}{2})\pi} \frac{|\sin s|}{s}\, ds + O(1). \end{aligned}$$

[2] The converse is in fact also true: If the Fourier series of f diverges at t, then $S^* f(t) = \infty$.

Let $m = \frac{1}{\pi}\int_0^\pi |\sin t|\,dt = \frac{2}{\pi}$ be the mean-value of $t \mapsto |\sin t|$, $t \in \mathbf{R}$, over a period. Then, according to the second mean-value theorem of integral calculus,

$$\int_1^{(N+\frac{1}{2})\pi} \frac{|\sin s| - m}{s}\,ds = \frac{1}{\pi}\int_1^\eta (|\sin s| - m)\,ds$$

for some number η such that $1 \leqslant \eta \leqslant (N+\frac{1}{2})\pi$. The integral in the right member is easily seen to be bounded with respect to N. Hence,

$$\begin{aligned} L_N &= \frac{2}{\pi}\int_1^{(N+\frac{1}{2})\pi} \frac{m}{s}\,ds + \frac{2}{\pi}\int_1^{(N+\frac{1}{2})\pi} \frac{|\sin s| - m}{s}\,ds + O(1) \\ &= \frac{4}{\pi^2}\ln(N+\tfrac{1}{2})\pi + O(1) = \frac{4}{\pi^2}\ln N + O(1). \end{aligned}$$ ∎

Proof (Theorem 2.11.2). For $N = 1, 2, \ldots$, define the functional $T_N : C(\mathbb{T}) \to \mathbf{C}$ by $T_N f = S_N f(0)$ for $f \in C(\mathbb{T})$. It is not so hard to show that $\|T_N\| = \|D_N\|_1$. So, according to Lemma 2.11.3, $\sup_{N\geqslant 1}\|T_N\| = \infty$. It thus follows from the Banach–Steinhaus theorem that $S^*f(0) = \infty$ for every f that belongs to a dense G_δ set in $C(\mathbb{T})$. For any of these functions f, the Fourier series diverges at 0. ∎

The result in Theorem 2.11.2 can be strengthened considerably. There is of course nothing special with the point $t = 0$ in the proof, so for every $t \in \mathbf{R}$, there exists a dense G_δ set $E_t \subset C(\mathbb{T})$ such that $S^*f(t) = \infty$ for every $f \in E_t$. Let $(t_i)_{i=1}^\infty$ be a dense subset to $\mathbf{R}$ and put $E = \bigcap_{i=1}^\infty E_{t_i}$. Then, according to Baire's theorem, E is also a dense G_δ set and has the property that for every $f \in E_t$,

$$S^*f(t_i) = \infty \quad \text{for all points } t_i.$$

Notice that the set $\{t \in \mathbf{R} : S^*f(t) = \infty\}$ is G_δ in $\mathbf{R}$ for every continuous function f since S^*f is lower semicontinuous (being the supremum of a sequence of continuous functions). Let us summarize:

Theorem 2.11.4. *There exists a a dense G_δ set $E \subset C(\mathbb{T})$ such that, for every function $f \in E$, the set $\{t \in \mathbf{R} : S^*f(t) = \infty\}$ is a dense G_δ set in $\mathbf{R}$.*

We can rephrase the theorem in the following way: There exists a dense subset E to $C(\mathbb{T})$, which is G_δ and has the property that for any function in E, the Fourier series diverges on a dense G_δ set. Let us mention that it follows from Baire's theorem that E is even uncountable.

We end this section by briefly returning to Theorem 2.8.1. This theorem may also be formulated by saying that the finite Fourier transform, which maps $L^1(\mathbb{T})$ into $\mathbf{c}_0$, is injective. As we saw Example 2.10.4, there are sequences in $\mathbf{c}_0$ that are not Fourier coefficients of any function in $L^1(\mathbb{T})$, i.e., the Fourier transform is not surjective. We shall now prove this by an abstract argument. Suppose that the Fourier transform were surjective and hence bijective. Then, according to the inverse mapping theorem, the inverse of the Fourier transform would be bounded, so there would exist a constant $C \geqslant 0$ such that

$$\|f\|_1 \leqslant C\|\widehat{f}\|_\infty \quad \text{for every } f \in L^1(\mathbb{T}).$$

But if $f = D_N$, then the right-hand side is 1 since the Fourier coefficients of D_N are either 1 or 0, while the left-hand side tends to ∞ as $N \to \infty$, which then gives a contradiction.

Chapter 3

Hilbert Spaces

Let X denote a complex vector space.

3.1. Inner Product Spaces, Hilbert Spaces

Inner Products

Definition 3.1.1. A function $(\,\cdot\,,\,\cdot\,) : X \times X \to \mathbf{C}$ is called an **inner product** if

(i) the function $(\,\cdot\,, z) : X \to \mathbf{C}$ is linear for every $z \in X$, that is,

$$(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z) \quad \text{for all } x, y \in X,\ \alpha, \beta \in \mathbf{C};$$

(ii) $(x, y) = \overline{(y, x)}$ for all $x, y \in X$;

(iii) $(x, x) \geqslant 0$ for every $x \in X$;

(iv) $(x, x) = 0$ if and only if $x = 0$.

Equipped with an inner product, X is called an **inner product space**.

It follows from (i) and (ii) that

$$(x, y + z) = (x, y) + (x, z) \quad \text{and} \quad (x, \alpha y) = \overline{\alpha}(x, y)$$

for $x, y, z \in X$ and $\alpha \in \mathbf{C}$. This means that $(\,\cdot\,,\,\cdot\,)$ is **sesquilinear** (linear in the first argument, but only additive in the second).

For the rest of this chapter, X will always denote an inner product space.

Example 3.1.2. Let us give a few examples of inner product spaces:

(a) The space $\mathbf{C}^d$ with

$$(x, y) = \sum_{j=1}^{d} x_j \overline{y_j}, \quad x, y \in \mathbf{C}^d;$$

(b) The space ℓ^2 with

$$(c, d) = \sum_{n=-\infty}^{\infty} c_n \overline{d_n}, \quad c, d \in \ell^2;$$

the series is absolutely convergent since $2|c_n \overline{d_n}| \leqslant |c_n|^2 + |d_n|^2$ for all n;

(c) The space $L^2(\mathbb{T})$ with

$$(f, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g(t)}\, dt, \quad f, g \in L^2(\mathbb{T});$$

this definition makes sense since $f\overline{g}$ is measurable and belongs to $L^1(\mathbb{T})$ because $2|f\overline{g}| \leqslant |f|^2 + |g|^2$, where $|f|^2 + |g|^2 \in L^1(\mathbb{T})$.

(d) The space $L^2(\mathbf{R}^d)$ with

$$(f, g) = \int_{\mathbf{R}^d} f(x)\overline{g(x)}\, dx, \quad f, g \in L^2(\mathbf{R}^d).$$

The Cauchy–Schwarz Inequality

Theorem 3.1.3 (The Cauchy–Schwarz Inequality). *For $x, y \in X$,*

$$|(x, y)| \leqslant \sqrt{(x, x)}\sqrt{(y, y)}.$$

Equality holds if and only if x and y are linearly dependent.

Proof. The inequality obviously holds true if $y = 0$. If $y \neq 0$, put $e = ty$, where $t^{-1} = \sqrt{(y, y)}$. Then $(e, e) = 1$, and

$$0 \leqslant (x - (x, e)e, x - (x, e)e) = (x, x) - |(x, e)|^2 = (x, x) - \frac{|(x, y)|^2}{(y, y)},$$

from which the Cauchy–Schwarz inequality follows directly. Equality holds if and only if $x - (x, e)e = x - t^2(x, y)y = 0$, which means that x and y are linearly dependent. ∎

Example 3.1.4. The Cauchy–Schwarz inequality for $L^2(\mathbb{T})$ is

$$\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt\right| \leqslant \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt\right)^{1/2}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |g(t)|^2\,dt\right)^{1/2}$$

for $f, g \in L^2(\mathbb{T})$. Notice that this inequality coincides with Hölder's inequality. □

The Norm on an Inner Product Space

Definition 3.1.5. For $x \in X$, we define $\|x\| = \sqrt{(x, x)}$.

With this notation, the Cauchy–Schwarz inequality may be written

$$|(x, y)| \leqslant \|x\|\|y\|, \quad x, y \in X.$$

Proposition 3.1.6. *The function $\|\cdot\|$ is a norm on X, that is,*

(i) $\|x\| \geqslant 0$ *for every* $x \in X$ *and* $\|x\| = 0$ *if and only if* $x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$ *for every* $\alpha \in \mathbf{C}$ *and every* $x \in X$;

(iii) $\|x + y\| \leqslant \|x\| + \|y\|$ *for all* $x, y \in X$.

The third property is called the **triangle inequality**.

Proof. It is only the triangle inequality that really requires a proof. We deduce this from the Cauchy–Schwarz inequality in the following way:

$$\begin{aligned}\|x + y\|^2 &= \|x\|^2 + 2\operatorname{Re}(x, y) + \|y\|^2 \leqslant \|x\|^2 + 2|(x, y)| + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$ ∎

Example 3.1.7. The norm of a function $f \in L^2(\mathbb{T})$ is

$$\|f\|_2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt\right)^{1/2}$$ □

The next simple, but useful corollary follows directly from the Cauchy–Schwarz inequality.

Corollary 3.1.8. *The function* $(\,\cdot\,, z) : X \to \mathbf{C}$ *is Lipschitz continuous for every fixed* $z \in X$:

$$|(x, z) - (y, z)| \leqslant \|x - y\|\|z\| \quad \text{for all} \quad x, y \in X.$$

In vector algebra, the following identity is known as the **parallelogram law**.

Proposition 3.1.9. *For* $x, y \in X$, $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

Proof. Expand the left-hand side as in the proof of Proposition 3.1.6. ■

Hilbert Spaces

With the norm, there comes a notion of convergence.

Definition 3.1.10.

(a) A sequence $(x_n)_{n=1}^{\infty}$ in X is said to be **convergent** if there exists a vector $x \in X$ such that $\|x - x_n\| \to 0$ as $n \to \infty$.

(b) A sequence $(x_n)_{n=1}^{\infty}$ is said to be a **Cauchy sequence** if $\|x_m - x_n\| \to 0$ as $m, n \to \infty$.

(c) The space X is said to be **complete** if every Cauchy sequence is convergent.

(d) A **Hilbert space** is a complete inner product space.

Example 3.1.11. One can show that the spaces in Example 3.1.2 are all Hilbert spaces. □

3.2. Orthogonality

Orthogonality, Orthonormal Sets

Definition 3.2.1. Two vectors $x, y \in X$ are said to be **orthogonal** if $(x, y) = 0$. This relation is denoted $x \perp y$.

The next proposition generalizes Pythagoras' Theorem in classical geometry.

Proposition 3.2.2 (Pythagoras' Theorem). *If* $x_1, \dots, x_N \in X$ *are pairwise orthogonal, that is,* $(x_m, x_n) = 0$ *if* $m \neq n$, *then*

$$\left\| \sum_{n=1}^{N} x_n \right\|^2 = \sum_{n=1}^{N} \|x_n\|^2.$$

Proof. Just expand the left-hand side in the identity using the properties of the inner product and the fact that the vectors are pairwise orthogonal:

$$\left\| \sum_{n=1}^{N} x_n \right\|^2 = \left(\sum_{m=1}^{N} x_m, \sum_{n=1}^{N} x_n \right) = \sum_{m,n=1}^{N} (x_m, x_n) = \sum_{n=1}^{N} (x_n, x_n) = \sum_{n=1}^{N} \|x_n\|^2. \quad ■$$

Orthonormal Sets

Definition 3.2.3. A subset E to X is called **orthonormal** if the elements in E are pairwise orthogonal and have all norm 1. A sequence $(e_n)_{n=1}^{\infty} \subset X$ is orthonormal if the corresponding set $E = \{e_1, e_2, ...\}$ is orthonormal.

Example 3.2.4. The sequence $(e^{int})_{n=-\infty}^{\infty} \subset L^2(\mathbb{T})$ is orthonormal:

$$(e^{imt}, e^{int}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)t}\,dt = \begin{cases} 1 \text{ if } m = n \\ 0 \text{ if } m \neq n \end{cases}. \qquad \square$$

Lemma 3.2.5. *Suppose that H is a Hilbert space and that $(e_n)_{n=1}^{\infty}$ is a orthonormal sequence in H. Let $(c_n)_{n=1}^{\infty}$ be a sequence of complex numbers. Then the series $\sum_{n=1}^{\infty} c_n e_n$ is convergent in H if and only if $\sum_{n=1}^{\infty} |c_n|^2 < \infty$.*

We remark that the convergence of the series $\sum_{n=1}^{\infty} c_n e_n$ means that there exists an element $x \in H$ such that $\|x - \sum_{n=1}^{N} c_n e_n\| \to 0$ as $N \to \infty$.

Proof. According to Pythagoras' theorem (Theorem 3.2.2),

$$\left\| \sum_{n=N}^{M} c_n e_n \right\|^2 = \sum_{n=N}^{M} |e_n|^2$$

for $M > N$. It follows that the series $\sum_{n=1}^{\infty} c_n e_n$ is convergent in H if and only if $\sum_{n=1}^{\infty} |c_n|^2$ is convergent. ■

Example 3.2.6. If the sequence $(c_n)_{n=-\infty}^{\infty} \subset \mathbf{C}$ satisfies $\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty$, then the function $f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}$, $t \in \mathbf{R}$, belongs to $L^2(\mathbb{T})$. Compare this with the following result: If we assume that $\sum_{n=-\infty}^{\infty} |c_n| < \infty$ (a stronger assumption), then it follows from Weierstrass' theorem that f is continuous on $\mathbf{R}$. □

3.3. Least Distance, Orthogonal Projections

Distance to a Subspace

In this and the following subsection, H will denote a Hilbert space. A subspace Y to H is said to be **closed** if Y contains all its limit points, i.e., if $(y_n)_{n=1}^{\infty}$ is a sequence in H and $y_n \to y \in H$, then, in fact, $y \in Y$.

Theorem 3.3.1. *Let Y be a closed subspace to H. Then, for every $x \in H$, there exists a unique vector $y \in Y$ such that*

$$\|x - y\| = \inf_{z \in Y} \|x - z\|.$$

Proof. First choose $(y_n)_{n=1}^{\infty} \subset Y$ such that $\|x - y_n\| \to d = \inf_{z \in Y} \|x - z\|$. By the parallelogram law (Theorem 3.1.9),

$$4\left\|x - \frac{y_m + y_n}{2}\right\|^2 + \|y_m - y_n\|^2 = 2(\|x - y_m\|^2 + \|x - y_n\|^2).$$

Notice that the first term in the left-hand side is at least $4d^2$. On the other hand, the right-hand side tends to $4d^2$, so it follows that $\|y_m - y_n\| \to 0$. If y denotes the limit of the sequence $(y_n)_{n=1}^{\infty}$, then $y \in Y$ since Y is closed. Moreover, since

$$d \leqslant \|x - y\| \leqslant \|x - y_n\| + \|y_n - y\| \longrightarrow d \quad \text{as } n \to \infty,$$

it follows that $\|x - y\| = d$. To prove that y is unique, suppose that $\|x - y'\| = d$ for some $y' \in Y$. Then, as above,

$$\left\|x - \frac{y + y'}{2}\right\|^2 + \|y - y'\|^2 = 2(\|x - y\|^2 + \|x - y'\|^2).$$

Since the first term in the left member is at least $4d^2$ and the right member is exactly $4d^2$, it follows that $\|y - y'\| = 0$, so $y = y'$. ■

Theorem 3.3.2. *Suppose that Y is a closed subspace to H. Then*

$$\|x - y\| = \inf_{z \in Y} \|x - z\| \quad \textit{if and only if} \quad (x - y, z) = 0 \quad \textit{for every } z \in Y.$$

Proof. Suppose first that $\|x - y\| = d = \inf_{z \in Y} \|x - z\|$. Given $z \in Y$, choose a scalar $\lambda \in \mathbf{C}$ such that $(x - y, \lambda z) = -|(x - y, z)|$. Then

$$\begin{aligned} d^2 &\leqslant \|(x - y) + t\lambda z\|^2 = \|x - y\|^2 + 2t \operatorname{Re}(x - y, \lambda z) + t^2|\lambda|^2\|z\|^2 \\ &= d^2 - 2t|(x - y, z)| + t^2|\lambda|^2\|z\|^2 \end{aligned}$$

for every $t \in \mathbf{R}$. This implies that $2|(x - y, z)| \leqslant t|\lambda|^2\|z\|^2$ for every $t \geqslant 0$, from which it follows that $(x - y, z) = 0$.

The converse is easier; in fact, by Pythagoras' theorem (Theorem 3.2.2),

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \geqslant \|x - y\|^2$$

for every $z \in Y$ since $x - y$ and $y - z$ are orthogonal. ■

Orthogonal Projections

Definition 3.3.3. Let Y be a closed subspace to H and let $x \in H$. The unique vector $y \in Y$, that satisfies $(x - y, z) = 0$ for every $z \in Y$, is called the **orthogonal projection** of x on Y. We will denote this vector by $P_Y x$.

Example 3.3.4. Suppose that $\{e_1, \ldots, e_N\} \subset H$ is orthonormal and let Y be the linear span of $\{e_1, \ldots, e_N\}$. Then the orthogonal projection of a vector $x \in H$ on Y is $P_Y x = \sum_{n=1}^{N} (x, e_n)e_n$ since $x - P_Y x \perp e_m$ for $m = 1, 2, \ldots, N$:

$$(x - P_Y x, e_m) = (x, e_m) - \sum_{n=1}^{N} (x, e_n)(e_n, e_m) = (x, e_m) - (x, e_m) = 0. \qquad \square$$

Example 3.3.5. The Fourier coefficients of $f \in L^2(\mathbb{T})$ are defined by

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)e^{-int}\, dt, \quad n = 0, \pm 1, \pm 2, \ldots .$$

Notice that $\hat{f}(n) = (f, e_n)$, where $e_n(t) = e^{int}$, $t \in \mathbf{R}$. It follows that the partial sum $\sum_{n=-N}^{N} \hat{f}(n)e^{int}$ to the Fourier series of f is nothing but the orthogonal projection of f on the linear span of the functions $e^{iNt}, \ldots, e^{-iNt}$. □

3.4. Orthonormal Bases

The Finite-Dimensional Case

Suppose that $\dim(X) = d < \infty$ and that $\{e_1, \dots, e_d\}$ is an orthonormal basis for X. Then every vector $x \in X$ can be written

$$x = \sum_{n=1}^{d} x_n e_n.$$

Taking the inner product of both sides in this identity with e_n, $n = 1, \dots, d$, we find that $x_n = (x, e_n)$, so that

$$x = \sum_{n=1}^{d} (x, e_n) e_n.$$

It now follows from Pythagoras' theorem that

$$\|x\|^2 = \sum_{n=1}^{d} |(x, e_n)|^2.$$

We shall next investigate to what extent these observations can be generalized to infinite-dimensional spaces.

Bessel's Inequality

Theorem 3.4.1 (Bessel's Inequality). *If $(e_n)_{n=1}^{\infty} \subset X$ is orthonormal, then, for every $x \in X$,*

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leqslant \|x\|^2.$$

Proof. According to Example 3.3.4, the orthogonal projection of x on the subspace $\operatorname{span}\{e_1, \dots, e_N\}$ to X is the vector $\sum_{n=1}^{N} (x, e_n) e_n$. Two applications of Pythagoras' theorem now shows that

$$\begin{aligned} \|x\|^2 &= \left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 + \left\| \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 \\ &= \left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 + \sum_{n=1}^{N} |(x, e_n)|^2 \geqslant \sum_{n=1}^{N} |(x, e_n)|^2. \end{aligned}$$

Since this inequality holds for any N, Bessel's inequality follows. ■

Example 3.4.2. For $L^2(\mathbb{T})$, Bessel's inequality takes the form

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 \, dt, \quad f \in L^2(\mathbb{T}). \qquad \square$$

Combining Bessels inequality with Lemma 3.2.5, we obtain the following result.

Corollary 3.4.3. *If $(e_n)_{n=1}^{\infty} \subset X$ is orthonormal, then the series $\sum_{n=1}^{\infty} (x, e_n) e_n$ is convergent for every $x \in X$.*

Orthonormal Bases, Parseval's Identity

Let H be a Hilbert space.

Definition 3.4.4. An orthonormal sequence $(e_n)_{n=1}^{\infty} \subset H$ is said to be an **orthonormal basis** for H if every $x \in H$ can be written

$$x = \sum_{n=1}^{\infty} (x, e_n) e_n.$$

Theorem 3.4.5. *For an orthonormal sequence $(e_n)_{n=1}^{\infty} \subset H$, the following conditions are equivalent.*

(i) The sequence $(e_n)_{n=1}^{\infty} \subset H$ is an orthonormal basis for H.

(ii) For every $x \in H$, $\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2$.

(iii) If $(x, e_n) = 0$ for every n, then $x = 0$.

The identity in (ii) is known as **Parseval's identity**.

Proof. We first assume that (i) holds true and deduce (ii). As in the proof of Bessel's inequality,

$$\|x\|^2 - \sum_{n=1}^{N} |(x, e_n)|^2 = \left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 .$$

The right-hand side tends to 0 as $N \to \infty$, so Parseval's identity holds.

The fact that (ii) implies (iii) is self-evident.

Finally, suppose that (iii) holds. Then, according to Corollary 3.4.3, the series $\sum_{n=1}^{\infty} (x, e_n) e_n$ is convergent; denote the sum by y. Since

$$(x - y, e_m) = (x, e_m) - (x, e_m) = 0$$

for every m, we have that $y = x$, and hence that $x = \sum_{n=1}^{\infty} (x, e_n) e_n$. ∎

Chapter 4

L^2-theory for Fourier Series

In the present chapter, we first establish Parseval's identity for $L^2(\mathbb{T})$. A consequence is the fact that $(e^{int})_{n=-\infty}^{\infty}$ is an orthonormal basis for $L^2(\mathbb{T})$, another is a for $L^2(\mathbb{T})$. We also prove the so called Riesz–Fischer theorem and a result about uniform convergence of Fourier series.

4.1. The Space $L^2(\mathbb{T})$

Let us summarize the definitions and results in Chapter 3 that concerned Fourier series.

(a) In Example 3.1.2, we defined an inner product for $L^2(\mathbb{T})$:

$$(f,g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt, \quad f,g \in L^2(\mathbb{T}).$$

(b) With this inner product, $L^2(\mathbb{T})$ becomes a Hilbert space.

(c) We also saw that $(e^{int})_{n=-\infty}^{\infty}$ is an orthonormal sequence in $L^2(\mathbb{T})$ in Example 3.2.4.

(d) Then, using the fact that $\hat{f}(n) = (f(t), e^{int})$ for $n \in \mathbf{Z}$ and $f \in L^2(\mathbb{T})$, we showed in Example 3.4.2 that Bessel's inequality for $L^2(\mathbb{T})$ has the form

$$\sum_{-\infty}^{\infty} |\hat{f}(n)|^2 \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt \quad \text{for } f \in L^2(\mathbb{T}). \tag{4.1}$$

In particular, the sequence $(\hat{f}(n))_{n=-\infty}^{\infty}$ of Fourier coefficients of $f \in L^2(\mathbb{T})$ belongs to ℓ^2.

Notice that Bessel's inequality implies that $\hat{f}(n) \to 0$ as $n \to \pm\infty$ for every function $f \in L^2(\mathbb{T})$; this is a weaker form of the Riemann–Lebesgue lemma (Proposition 2.4.3).

4.2. Parseval's Identity

Theorem 4.2.1 (Parseval's Identity). *Suppose that $f, g \in L^2(\mathbb{T})$. Then*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt = \sum_{n=-\infty}^{\infty} \hat{f}(n)\overline{\hat{g}(n)}.$$

Taking $g = f$, where $f \in L^2(\mathbb{T})$, in Parseval's identity, we see that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2.$$

Proof. We first assume that f belongs to $C^2(\mathbb{T})$. Then $\hat{f}(n) = o(n^{-2})$ as $n \to \pm\infty$ according to Theorem 2.4.6, which implies that the Fourier series of f is uniformly convergent. Using this fact together with Corollary 2.6.2, we obtain that

$$\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt = \int_{-\pi}^{\pi}\left(\sum_{n=-\infty}^{\infty}\widehat{f}(n)e^{int}\right)\overline{g(t)}\,dt = \sum_{n=-\infty}^{\infty}\widehat{f}(n)\int_{-\pi}^{\pi}\overline{g(t)}e^{int}\,dt$$
$$= 2\pi\sum_{n=-\infty}^{\infty}\widehat{f}(n)\overline{\widehat{g}(n)}.$$

In the general case, we choose as a sequence $(f_k)_{k=-\infty}^{\infty}$ of functions in $C^2(\mathbb{T})$ such that $\|f-f_k\|_2 \to 0$ as $k\to\infty$. Bessel's inequality (4.1) then shows that

$$\|\widehat{f}-\widehat{f}_k\|_{\ell^2} \leqslant \|f-f_k\|_2,$$

so $\widehat{f}_k \to \widehat{f}$ in ℓ^2. We finally obtain from the first case that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt = \lim_{k\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} f_k(t)\overline{g(t)}\,dt = \lim_{k\to\infty}\sum_{n=-\infty}^{\infty}\widehat{f}_k(n)\overline{\widehat{g}(n)}$$
$$= \sum_{n=-\infty}^{\infty}\widehat{f}(n)\overline{\widehat{g}(n)}. \qquad \blacksquare$$

Example 4.2.2. In Example 2.2.2, we showed that the Fourier series of the function $f \in L^2(\mathbb{T})$, defined by $f(t)=t,\ -\pi \leqslant t < \pi$, is

$$i\sum_{n\neq 0}\frac{(-1)^n}{n}e^{int}.$$

Parseval's inequality now shows that

$$\sum_{n\neq 0}\frac{1}{n^2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} t^2\,dt, \quad \text{which implies that} \quad \sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}. \qquad \square$$

The next two results are consequences of Theorem 3.4.5. Notice that the second corollary is a special case of the more general Theorem 2.8.1.

Corollary 4.2.3. *The sequence $(e^{int})_{n=-\infty}^{\infty}$ is an orthonormal basis for $L^2(\mathbb{T})$.*

The statement means that if $f \in L^2(\mathbb{T})$, then $f(t) = \sum_{n=-\infty}^{\infty}\widehat{f}(n)e^{int}$ in the sense of $L^2(\mathbb{T})$, that is,

$$\|f - S_N f\|_2 \to 0 \quad \text{as } N\to\infty.$$

Corollary 4.2.4. *Suppose that $f, g \in L^2(\mathbb{T})$ and $\widehat{f}(n) = \widehat{g}(n)$ for every $n \in \mathbf{Z}$. Then $f = g$ a.e.*

4.3. The Riesz–Fischer Theorem

As noticed in Section 4.1, the finite Fourier transform $\mathscr{F}$, defined by

$$\mathscr{F}f(n) = \widehat{f}(n),\ n\in\mathbf{Z}, \quad \text{for } f\in L^2(\mathbb{T}),$$

maps $L^2(\mathbb{T})$ into ℓ^2. This mapping is obviously linear. According to Parseval's identity, it is also an isometry:

$$\|\mathscr{F} f\|_{\ell^2} = \|\widehat{f}\|_{\ell^2} = \|f\|_2 \quad \text{for } f \in L^2(\mathbb{T}),$$

and according to the uniqueness theorem, it is injective (this, of course, also follows from the fact that every linear isometry is injective). To show that $\mathscr{F}$ is surjective, we assume that $(c_n)_{n=-\infty}^{\infty}$ is an arbitrary sequence in ℓ^2. Lemma 3.2.5 then shows that the function $f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}$ belongs to $L^2(\mathbb{T})$. Moreover, since the inner product is continuous according to Corollary 3.1.8,

$$\widehat{f}(m) = (f(t), e^{imt}) = \sum_{n=-\infty}^{\infty} c_n (e^{int}, e^{imt}) = c_m \quad \text{for every } m \in \mathbf{Z},$$

which shows that $\mathscr{F} f(n) = c_n$ for every $n \in \mathbf{Z}$. These observations are summarized in the following theorem.

Theorem 4.3.1 (The Riesz–Fischer Theorem). *The space $L^2(\mathbb{T})$ is isometrically isomorphic to ℓ^2.*

The isomorphism in the theorem is thus the finite Fourier transform.

4.4. Characterization of Function Spaces

In some cases, function spaces can be characterized in terms of Fourier coefficients. For instance, a function $f \in L^1(\mathbb{T})$ belongs to $L^2(\mathbb{T})$ if and only if

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 < \infty.$$

The necessity of this condition follows from Bessel's inequality and the sufficiency from Riesz–Fischer's Theorem in conjunction with the uniqueness theorem.

Now suppose that $f \in AC(\mathbb{T})$ with $f' \in L^2(\mathbb{T})$. According to Proposition 2.4.6, we have $\widehat{f'}(n) = in\widehat{f}(n)$ for every $n \in \mathbf{Z}$, so Parseval's identity shows that

$$\sum_{n=-\infty}^{\infty} n^2 |\widehat{f}(n)|^2 = \|f'\|_2 < \infty.$$

We shall now adress the converse.

Theorem 4.4.1. *Suppose that $f \in L^1(\mathbb{T})$ satisfies*

$$\sum_{n=-\infty}^{\infty} n^2 |\widehat{f}(n)|^2 < \infty. \tag{4.2}$$

Then there exists a function $g \in AC(\mathbb{T})$ with $g' \in L^2(\mathbb{T})$ such that $f = g$ a.e.

Thus, if $f \in L^1(\mathbb{T})$, then $f \in AC(\mathbb{T})$ with $f' \in L^2(\mathbb{T})$ if and only if (4.2) holds (in the sufficiency part, we assume that f is redefined on a set of measure 0).

Proof. Using Riesz–Fischer's theorem, it follows from (4.2) that there exists a function $h \in L^2(\mathbb{T})$ such that $\widehat{h}(n) = in\widehat{f}(n)$ for every $n \in \mathbf{Z}$. If

$$H(t) = \int_{-\pi}^{t} g(\tau)\,d\tau, \quad t \in \mathbf{R},$$

then H has period 2π since $\widehat{h}(0) = 0$. Moreover, H is absolutely continuous with $H' = h$ a.e. We also have

$$in\widehat{f}(n) = \widehat{h}(n) = \widehat{H'}(n) = in\widehat{H}(n) \quad \text{for every } n \in \mathbf{Z},$$

so $\widehat{H}(n) = \widehat{f}(n)$ for $n \neq 0$. The uniqueness theorem (Theorem 2.8.1) now shows that $H - f = \widehat{H}(0) - \widehat{f}(0)$ a.e. Finally, put $g = H - (\widehat{H}(0) - \widehat{f}(0))$. ■

4.5. More About Uniform Convergence

In Theorem 2.9.2, we proved that if $f \in C^2(\mathbb{T})$, then the Fourier series for f is uniformly and absolutely convergent. We shall now show that this also holds under the weaker assumption that $f \in AC(\mathbb{T})$ with $f' \in L^2(\mathbb{T})$. This of course implies that the same conclusion holds if $f \in C^1(\mathbb{T})$.

Theorem 4.5.1. *Suppose that $f \in AC(\mathbb{T})$ with $f' \in L^2(\mathbb{T})$. Then the Fourier series of f is absolutely convergent. Moreover,*

$$\|f - S_N f\|_\infty \leqslant \sqrt{\frac{2}{N}}\|f'\|_2. \tag{4.3}$$

In particular, the Fourier series of f converges uniformly to f.

Proof. Using the identity $in\widehat{f}(n) = \widehat{f'}(n)$ together with the Cauchy–Schwarz inequality for ℓ^2 and Corollary 2.6.8, we obtain that

$$\begin{aligned}\|f - S_N f\|_\infty &\leqslant \sum_{|n| \geqslant N+1} |\widehat{f}(n)| = \sum_{|n|\geqslant N+1} \frac{1}{|n|}|in\widehat{f}(n)| \\ &\leqslant \Big(\sum_{|n|\geqslant N+1} \frac{1}{n^2}\Big)^{1/2}\Big(\sum_{|n|\geqslant N+1} |\widehat{f'}(n)|^2\Big)^{1/2}.\end{aligned}$$

The inequality (4.3) now follows from Bessel's inequality (4.1) and the fact that

$$\sum_{n=N+1}^{\infty} \frac{1}{n^2} \leqslant \int_N^\infty \frac{dt}{t^2} = \frac{1}{N}.$$ ■

Chapter 5

Summation of Fourier Series

5.1. Cesàro Convergence

Given a sequence $(a_n)_{n=0}^{\infty}$ of complex numbers, we denote by σ_N the arithmetic mean of the first $N+1$ terms in the sequence, i.e.,

$$\sigma_N = \frac{a_0 + a_1 + \dots + a_N}{N+1}, \quad N = 0, 1, \dots .$$

Definition 5.1.1. Let $(a_n)_{n=0}^{\infty}$ be a sequence of complex numbers.

(a) The sequence $(a_n)_{n=0}^{\infty}$ is said to be **Cesàro convergent** with (**Cesàro**) **limit** a if $\sigma_N \to a$ as $N \to \infty$.

(b) The series $\sum_{k=0}^{\infty} a_k$ is said to be **Cesàro summable** with (**Cesàro**) **sum** S if the sequence of partial sums $S_n = \sum_{k=0}^{n} a_k$, $n = 0, 1, \dots$, is Cesàro convergent with limit S.

Sequences and series, that are divergent in the usual sense, may in fact be convergent in this new sense as the following examples show.

Example 5.1.2. The sequence $1, 0, 1, 0, \dots$ is Cesàro convergent with limit $\frac{1}{2}$. Indeed,

$$\sigma_{2k+1} = \frac{1}{2} \quad \text{and} \quad \sigma_{2k} = \frac{k+1}{2k+1} \quad \text{for } k = 0, 1, \dots . \qquad \square$$

Example 5.1.3. The series $1 - 1 + 1 - 1 + \dots$ is Cesàro summable with sum $\frac{1}{2}$. In fact, the sequence of partial sum is $S_0 = 1,\ S_1 = 0,\ S_2 = 1,\ S_3 = 0, \dots$, which has Cesàro limit $\frac{1}{2}$ according to the previous example. $\square$

The following proposition shows that if a sequence is convergent, then it is also Cesàro convergent with the same limit. The converse is false according to Example 5.1.2.

Proposition 5.1.4. *Suppose that $(a_n)_{n=0}^{\infty}$ is a convergent sequence of complex numbers with limit a. Then $\lim_{N\to\infty} \sigma_N = a$.*

Proof. Let $\varepsilon > 0$ be arbitrary and choose M so large that $|a - a_n| < \varepsilon$ if $n > M$. For $N > M$, we then have

$$|a - \sigma_N| = \frac{1}{N+1}\left|\sum_{n=0}^{N}(a - a_n)\right| \leqslant \frac{1}{N+1}\left|\sum_{n=0}^{M}(a - a_n)\right| + \frac{N-(M+1)}{N+1}\varepsilon.$$

The second term in the right-hand side of this inequality is less than ε. The claim thus follows if we choose N so large that the first term is also less than ε. ■

5.2. The Fejér Kernel

We next consider Cesàro summability of Fourier series. The Cesàro means or **Fejér means** $\sigma_N f$ for the Fourier series of a function $f \in L^1(\mathbb{T})$ are defined by

$$\sigma_N f(t) = \frac{1}{N+1} \sum_{n=0}^{N} S_n f(t), \quad t \in \mathbf{R},\ N = 0, 1, \ldots .$$

Using Equation (2.9), we see that

$$\sigma_N f(t) = \frac{1}{N+1} \sum_{n=0}^{N} D_n * f(t) = \left(\frac{1}{N+1} \sum_{n=0}^{N} D_n \right) * f(t).$$

The expression within brackets in the right-hand side of this equation is known as the **Fejér kernel** and denoted $K_N,\ N = 0, 1, \ldots$. To obtain an explicit expression for K_N, we use (2.10):

$$\begin{aligned} (N+1)\sin^2 \tfrac{t}{2} K_N(t) &= \sum_{n=0}^{N} \sin \tfrac{t}{2} \sin{(N+\tfrac{1}{2})t} = \frac{1}{2} \sum_{n=0}^{N} (\cos nt - \cos{(n+1)t}) \\ &= \frac{1}{2}(1 - \cos{(N+1)t}) = \sin^2 \tfrac{N+1}{2} t \end{aligned}$$

for every $t \in \mathbf{R}$. We thus have

$$K_N(t) = \begin{cases} \dfrac{1}{N+1} \left(\dfrac{\sin \frac{N+1}{2} t}{\sin \frac{t}{2}} \right)^2 & \text{for} \quad t \notin 2\pi\mathbf{Z} \\ N+1 & \text{for} \quad t \in 2\pi\mathbf{Z} \end{cases} .$$

Proposition 5.2.1. *The Fejér kernel K_N has the following properties*:

(i) $K_N \geqslant 0$;

(ii) $\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(t)\, dt = 1$;

(iii) for every $\delta > 0$, $\int_{\delta \leqslant |t| < \pi} K_N(t)\, dt \to 0$ *as* $N \to \infty$;

(iv) K_N *is even*;

(v) $K_N(t) \leqslant N+1$ *for every* $t \in \mathbf{R}$.

Properties (i)–(iii) show that $(K_N)_{N=1}^{\infty}$ is an approximate identity (see Definition 1.5.1).

Proof. Out of these five properties, the first and the fourth are obvious. The second holds because the same is true for the Dirichlet kernel. If we use the fact that $|\sin t/2| \geqslant |t|/\pi$ for $|t| \leqslant \pi$, we obtain that

$$K_N(t) \leqslant \frac{\pi^2}{(N+1)t^2} \quad \text{for } 0 < |t| \leqslant \pi, \tag{5.1}$$

form which the third property follows. Finally, to prove the fifth property, notice that

$$|D_n(t)| = \left| 1 + 2 \sum_{k=1}^{n} \cos kt \right| \leqslant 1 + 2n$$

for $t \in \mathbf{R}$ and $n = 0, 1, \ldots$, so that

$$K_N(t) = \left| \frac{1}{N+1} \sum_{n=0}^{N} D_n(t) \right| \leqslant \frac{1}{N+1} \sum_{n=0}^{N} (1 + 2n) = N + 1 \quad \text{for } t \in \mathbf{R}. \qquad \blacksquare$$

Proposition 5.2.2. *Suppose that $f \in L^1(\mathbb{T})$. Then*

$$\sigma_N f(t) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N+1}\right) \hat{f}(n) e^{int} \quad \textit{for } t \in \mathbf{R} \textit{ and } N = 0, 1, \ldots . \tag{5.2}$$

Proof. The identity (5.2) is proved by changing the order of summation:

$$\begin{aligned} \sigma_N f(t) &= \frac{1}{N+1} \sum_{n=0}^{N} S_n f(t) = \frac{1}{N+1} \sum_{n=0}^{N} \sum_{k=-n}^{n} \hat{f}(k) e^{ikt} \\ &= \frac{1}{N+1} \sum_{k=-N}^{N} \sum_{n=|k|}^{N} \hat{f}(k) e^{ikt} = \sum_{k=-N}^{N} \left(1 - \frac{|k|}{N+1}\right) \hat{f}(k) e^{ikt}. \qquad \blacksquare \end{aligned}$$

5.3. Fejér's Theorem

The next theorem, which was proved by L. Fejér in 1904, is a consequence of the fact that $(K_N)_{N=1}^{\infty}$ is an approximate identity. The theorem shows that the Fourier series of a L^1-function f is Cesàro summable at every point, where f has one-sided limits (and, in particular, at every point where f is continuous) and uniformly Cesàro summable on every compact set, where f is continuous.

Theorem 5.3.1. *Suppose that $f \in L^1(\mathbb{T})$.*

(a) If the one-sided limits $f(t^+)$ and $f(t^-)$ exist at some point $t \in \mathbf{R}$, then $\sigma_N f(t)$ converges to $(f(t^+) + f(t^-))/2$ as $N \to \infty$.

(b) If f is continuous on a closed set $F \subset \mathbf{R}$, then $\sigma_N f$ converges uniformly to f on F as $N \to \infty$.

According to du Bois-Reymond's example (see Theorem 2.11.2), the corresponding theorem with $\sigma_N f$ replaced by $S_N f$ is false. We have, however, the corollary below, which follows directly from Proposition 5.1.4 and Theorem 5.3.1.

Corollary 5.3.2. *Suppose that $f \in L^1(\mathbb{T})$. If the Fourier series of f converges at a point $t \in \mathbf{R}$, where the one-sided limits $f(t^+)$ and $f(t^-)$ exist, then*

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{int} = \frac{f(t^+) + f(t^-)}{2}.$$

Suppose, for instance, that $f \in C(\mathbb{T})$ and the Fourier series of f is absolutely convergent. It then follows from Corollary 5.3.2 that

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{int} = f(t) \quad \text{for every } t \in \mathbf{R}.$$

Compare this result with Theorem 2.9.1.

The corollaries are two versions of **Weierstrass' approximation theorem**, one for trigonometric approximation and one for polynomial approximation. By a **trigonometric polynomial** we here mean a function of the form

$$p(t) = \sum_{n=-N}^{N} c_n e^{int}, \quad t \in \mathbf{R}.$$

Corollary 5.3.3. *The class of trigonometric polynomials is dense in $C(\mathbb{T})$.*

Proof. If $f \in C(\mathbb{T})$, then $\sigma_N f$ converges uniformly to f as $N \to \infty$ according to Theorem 5.3.1. But $\sigma_N f$ is a trigonometric polynomial for every N (see Proposition 5.2.2). ∎

Corollary 5.3.4. *Suppose that $-\infty < a < b < \infty$. Then the class of polynomials is dense in $C([a, b])$.*

Proof. The proof is readily reduced to the case $[a, b] = [-1, 1]$ by a linear change of variables. Suppose that $f \in C([-1, 1])$. Then the function $g(s) = f(\cos s)$, $s \in \mathbf{R}$, belongs to $C(\mathbb{T})$. Theorem 5.3.1 then shows that the trigonometric polynomials

$$\sigma_N g(s) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N+1}\right) \widehat{g}(n) e^{ins} = \widehat{g}(0) + \sum_{n=1}^{N} \left(1 - \frac{n}{N+1}\right) \widehat{g}(n) \cos ns$$

tend to g uniformly as $N \to \infty$. If we now make the substitution $t = \cos s$, where $0 \leqslant s \leqslant \pi$, we see that the functions

$$P_N(t) = \widehat{g}(0) + \sum_{n=1}^{N} \left(1 - \frac{n}{N+1}\right) \widehat{g}(n) \cos(n \arccos t)$$

tend to f uniformly as $N \to \infty$. To finish the proof, we need to show that the function $p_n(t) = \cos(n \arccos t)$, $t \in \mathbf{R}$, actually is a polynomial for $n = 1, 2, \ldots$. First of all, $p_0(t) = 1$ and $p_1(t) = t$. Moreover, for $n \geqslant 2$,

$$\cos(n \arccos t) = 2 \cos(\arccos t) \cos((n-1) \arccos t) - \cos((n-2) \arccos t). \quad (5.3)$$

It therefore follows by induction that the right-hand side is a polynomial. ∎

The polynomials p_n, that we encountered in the proof of Corollary 5.3.4, are known as the **Chebyshev polynomials**. Notice that (5.3) shows that these polynomials satisfy the recursive formula

$$p_n(t) = 2t p_{n-1}(t) - p_{n-2}(t), \quad n = 2, 3, \ldots.$$

Since $p_0(t) = 1$ and $p_1(t) = t$, we see for instance that

$$p_2(t) = 2t^2 - 1, \; p_3(t) = 4t^3 - 3t, \; \text{and } p_4(t) = 8t^4 - 8t^2 + 1.$$

5.4. Convergence in L^p

The next theorem, which deals with convergence in $L^p(\mathbb{T})$ of the Fejér means, is a consequence of Theorem 1.5.3. The corresponding result for $p = \infty$ is false since the uniform limit of a sequence of continuous functions is continuous. With $S_N f$ instead of $\sigma_N f$, the result is false for $p = 1$, but true for $1 < p < \infty$. The proof in the latter case is much harder, except, of course, for $p = 2$.

Theorem 5.4.1. *Suppose that $f \in L^p(\mathbb{T})$, where $1 \leqslant p < \infty$. Then $\sigma_N f$ converges to f in $L^p(\mathbb{T})$ as $N \to \infty$.*

With the aid of Theorem 5.4.1, we obtain a new proof of Corollary 4.2.3.

Corollary 5.4.2. *Suppose that $f \in L^2(\mathbb{T})$. Then $S_N f$ converges to f in $L^2(\mathbb{T})$ as $N \to \infty$.*

Proof. Since $S_N f$ is the orthogonal projection on the linear span of of the functions $e^{iNt}, \dots, e^{-iNt}$ (see Example 3.3.5), we have that

$$\|f - S_N f\|_2 \leqslant \|f - \sigma_N f\|_2 \quad \text{for } N = 0, 1, \dots . \qquad \blacksquare$$

We also get a new proof of the uniqueness theorem for Fourier series (Theorem 2.8.1).

Corollary 5.4.3. *Suppose that $f, g \in L^1(\mathbb{T})$ and $\hat{f}(n) = \hat{g}(n)$ for every $n \in \mathbf{Z}$. Then $f = g$ a.e.*

Proof. It follows from the assumption that $\sigma_N f = \sigma_N g$ for every N. This implies that

$$\|f - g\|_1 \leqslant \|f - \sigma_N f\|_1 + \|\sigma_N g - g\|_1 \longrightarrow 0 \quad \text{as } N \to \infty,$$

from which it follows that $\|f - g\|_1 = 0$, so $f = g$ a.e. $\blacksquare$

5.5. Lebesgue's Theorem

To prove our next theorem, we will need the concept of a Lebesgue point.

Definition 5.5.1. Suppose that $f \in L^1(\mathbb{T})$. A point $t \in \mathbf{R}$ is said to be a **Lebesgue point** for f if

$$\lim_{h \to 0} \frac{1}{h} \int_0^h |f(s+t) - f(t)|\, ds = 0.$$

Every point of continuity of f is obviously a Lebesgue point. The Lebesgue points appear in the theory of differentiation in the following way. Let $a \in \mathbf{R}$ and put

$$F(t) = \int_a^t f(s)\, ds, \quad t \in \mathbf{R}.$$

Then F is differentiable at every Lebesgue point t of f with derivative $f(t)$ since

$$\left| \frac{F(t+h) - F(t)}{h} - f(t) \right| \leqslant \left| \frac{1}{h} \int_0^h |f(s+t) - f(t)|\, ds \right| \longrightarrow 0 \quad \text{as } h \to 0.$$

The basic result about Lebesgue points, which we state without a proof, is due to H. Lebesgue.

Theorem 5.5.2. *Suppose that $f \in L^1(a,b)$. Then almost every $t \in (a,b)$ is a Lebesgue point of f.*

The next theorem is also due to Lebesgue.

Theorem 5.5.3. *Suppose that $f \in L^1(\mathbb{T})$. Then $\sigma_N f(t) \to f(t)$ as $N \to \infty$ at every Lebesgue point $t \in \mathbf{R}$ of f.*

Notice that it is not true that the Fourier series of a function in $L^1(\mathbb{T})$ converges to the function almost everywhere; this follows from Kolmogorov's example (see Section 2.11),

Proof. Let $t \in \mathbf{R}$ be a Lebesgue point of f. Using (ii) and (iv) in Proposition 5.2.1, we see that

$$|\sigma_N f(t) - f(t)| \leqslant \frac{1}{2\pi}\int_0^\pi |f(t-s) + f(t+s) - 2f(t)| K_N(s)\, ds \quad \text{for a.e. } t \in \mathbf{R}.$$

Put $g(s) = |f(t-s) + f(t+s) - 2f(t)|$ for $0 \leqslant s \leqslant \pi$. Also put

$$G(u) = \int_0^u g(s)\, ds, \quad 0 \leqslant u \leqslant \pi.$$

Then, since t is a Lebesgue point of f,

$$\frac{F(u)}{u} \leqslant \frac{1}{u}\int_0^u |f(t-s) - f(t)|\, ds + \frac{1}{u}\int_0^u |f(t+s) - f(t)|\, ds \longrightarrow 0 \quad \text{as } u \to 0.$$

For an arbitrary $\varepsilon > 0$, one can therefore find a number $\delta > 0$ such that $u^{-1}F(u) < \varepsilon$ if $0 < u < \delta$. Using this and (v) in Proposition 5.2.1, we then obtain that

$$\int_0^{1/N} g(s) K_N(s)\, ds \leqslant (N+1)F(1/N) < 2\varepsilon \quad \text{if } N > \delta^{-1}.$$

It also follows from (5.1) that

$$\begin{aligned}\int_{1/N}^{\delta} g(s)K_N(s)\, ds &\leqslant \frac{\pi^2}{N}\int_{1/N}^{\delta} \frac{g(s)}{s^2}\, ds \\ &= \frac{\pi^2}{N}\left(\frac{F(\delta)}{\delta^2} - \frac{F(1/N)}{N^{-2}} + 2\int_{1/N}^{\delta} \frac{F(s)}{s}\frac{ds}{s^2}\right) \\ &\leqslant \frac{\pi^2}{N}\left(\frac{\varepsilon}{\delta} + 2\varepsilon N\right) < 3\pi^2\varepsilon.\end{aligned}$$

Finally,

$$\int_\delta^\pi g(s)K_N(s)\, ds \leqslant \frac{\pi^2}{N\delta^2}\int_\delta^\pi g(s)\, ds \leqslant \frac{\pi^2}{N\delta^2}(4\pi\|f\|_1 + 2\pi|f(t)|) < \varepsilon$$

for sufficiently large N. ∎

Corollary 5.5.4. *Suppose that $f \in L^1(\mathbb{T})$ and that $\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{int}$ is convergent a.e. Then $f(t) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{int}$ a.e.*

Proof. Let $g(t)$ denote the sum of the Fourier series of f at $t \in \mathbf{R}$. The statement in the theorem then follows by combining the theorems 5.5.2 and 5.5.3 with Proposition 5.1.4:

$$f(t) = \lim_{N\to\infty} \sigma_N f(t) = \lim_{N\to\infty} S_N f(t) = g(t) \quad \text{a.e.} \qquad \blacksquare$$

5.6. Hardy's Tauberian Theorem

We have above seen that if $(a_n)_{n=0}^{\infty}$ is a sequence of complex numbers and $a_n \to a$, then $\sigma_N \to a$, but also that the converse in general is not true. Results, describing situations when the converse is in fact true, are called **Tauberian theorems** after Tauber who was the first to establish results of this type. We will now prove **Hardy's Tauberian theorem**.

Theorem 5.6.1. *Suppose that $f \in L^1(\mathbb{T})$ satisfies $\hat{f}(n) = O(n^{-1})$ as $n \to \pm\infty$. If $\sigma_N f(t)$ converges for some $t \in \mathbf{R}$, then $S_N f(t)$ converges to the same limit. Morover, if $\sigma_N f$ converges uniformly on some set, the same holds for $S_N f$.*

Proof. It is not so hard to show that

$$\begin{aligned} S_N f(t) - \sigma_N f(t) &= \frac{M+1}{M-N}(\sigma_M f(t) - \sigma_N f(t)) \\ &\quad - \frac{M+1}{M-N} \sum_{N<|n|\leqslant M} \left(1 - \frac{|n|}{M+1}\right) \hat{f}(n) e^{int}, \end{aligned}$$

if $M > N \geqslant 1$. Denote the sum in the right-hand side by $S_{M,N}(t)$. Let $\varepsilon > 0$ be arbitrary and put $M = \lceil (1+\varepsilon)N \rceil$ (where $\lceil r \rceil$ is the integer part of $r \in \mathbf{R}$ plus 1). Then

$$\frac{M+1}{M-N} \leqslant \frac{(1+\varepsilon)N+2}{(1+\varepsilon)N-N} = \frac{1+\varepsilon+\frac{2}{N}}{\varepsilon}.$$

It follows that

$$\limsup_{N\to\infty} \left| \frac{M+1}{M-N}(\sigma_M f(t) - \sigma_N f(t)) \right| = 0.$$

By the assumption, $|\hat{f}(n)| \leqslant C|n|^{-1}$, so

$$\begin{aligned} \left| \frac{M+1}{M-N} S_{M,N}(t) \right| &\leqslant C \frac{M+1}{M-N} \sum_{n=N+1}^{M} \left(\frac{1}{n} - \frac{1}{M+1} \right) \\ &\leqslant C \frac{M+1}{M-N} \left(\ln \frac{M}{N} - \frac{M-N}{M+1} \right) \\ &< C \Big(\frac{1+\varepsilon+\frac{2}{N}}{\varepsilon} \ln(1+\varepsilon+\tfrac{1}{N}) - 1 \Big), \end{aligned}$$

which shows that

$$\limsup_{N\to\infty} \left| \frac{M+1}{M-N} S_{M,N}(t) \right| \leqslant C \Big(\frac{1+\varepsilon}{\varepsilon} \ln(1+\varepsilon) - 1 \Big) < C\varepsilon.$$

Since ε was arbitrary, the limit in the left-hand side has to be 0. This proves the first assertion. Because all estimates so far are independent of t, we see that $S_N f$ converges uniformly whenever $\sigma_N f$ does. $\blacksquare$

Corollary 5.6.2. *Suppose that $f \in AC(\mathbb{T})$. Then the Fourier series of f is uniformly convergent on* $\mathbf{R}$.

Proof. According to Proposition 2.4.6, $\hat{f}(n) = o(n^{-1})$ as $n \to \pm\infty$, and according to Theorem 5.3.1, $\sigma_N f \to f$ uniformly on $\mathbf{R}$ as $N \to \infty$. The result therefore follows from Hardy's Tauberian theorem. ∎

Part III

Fourier Transforms

Chapter 6

L^1-theory for Fourier Transforms

6.1. The Fourier Transform

Definition 6.1.1. The **Fourier transform** $\widehat{f}$ of a function $f \in L^1(\mathbf{R}^d)$ is defined by

$$\widehat{f}(\xi) = \int_{\mathbf{R}^d} f(x)e^{-ix\cdot\xi}\,dx, \quad \xi \in \mathbf{R}^d.$$

Here, $x \cdot \xi = \sum_{j=1}^d x_j\xi_j,\ x, \xi \in \mathbf{R}^d$, is the standard inner product in $\mathbf{R}^d$. Notice that the Fourier transform is absolutely convergent since

$$|f(x)e^{-ix\cdot\xi}| = |f(x)| \quad \text{for every } x \in \mathbf{R}^d \text{ and every } \xi \in \mathbf{R}^d.$$

Example 6.1.2. Let f be the characteristic function of the interval $(-1, 1) \subset \mathbf{R}$. Then

$$\widehat{f}(\xi) = \int_{-1}^{1} e^{-ix\xi}\,dx = \frac{2\sin\xi}{\xi} \quad \text{for } \xi \neq 0$$

and $\widehat{f}(0) = 2$. Notice that $\widehat{f} \notin L^1(\mathbf{R})$. □

Example 6.1.3. Let $f(x) = e^{-|x|},\ x \in \mathbf{R}$. Then

$$\begin{aligned}\widehat{f}(\xi) &= \int_{-\infty}^{\infty} e^{-|x|}e^{-ix\xi}\,dx = \int_0^{\infty} e^{-(1+i\xi)x}\,dx + \int_{-\infty}^{0} e^{(1-i\xi)x}\,dx \\ &= \frac{1}{1+i\xi} + \frac{1}{1-i\xi} = \frac{2}{1+\xi^2} \quad \text{for } \xi \in \mathbf{R}.\end{aligned}$$

□

Example 6.1.4. Let $f(x) = e^{-|x|^2/2},\ x \in \mathbf{R}^d$. To calculate the Fourier transform of f, we first consider the case $d = 1$. Then

$$\begin{aligned}\widehat{f}(\xi) &= \int_{-\infty}^{\infty} e^{-x^2/2}e^{-ix\xi}\,dx = e^{-\xi^2/2}\int_{-\infty}^{\infty} e^{-(x+i\xi)^2/2}\,dx \\ &= e^{-\xi^2/2}\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}e^{-\xi^2/2} \quad \text{for } \xi \in \mathbf{R}.\end{aligned}$$

Here, the penultimate equality follows from Cauchy's theorem. For the general case, we put $f_j(x) = e^{-x_j^2/2},\ x \in \mathbf{R}^d$, for $j = 1, \dots, d$. Then $f = f_1 \cdot \ldots \cdot f_d$, from which it follows that

$$\begin{aligned}\widehat{f}(\xi) &= \widehat{f}_1(\xi_1) \cdot \ldots \cdot \widehat{f}_d(\xi_d) = \sqrt{2\pi}e^{-\xi_1^2/2} \cdot \ldots \cdot \sqrt{2\pi}e^{-\xi_d^2/2} \\ &= (2\pi)^{d/2}e^{-|\xi|^2/2} \quad \text{for } \xi \in \mathbf{R}^d.\end{aligned}$$

□

Example 6.1.5. Suppose that $f \in L^1(\mathbf{R}^d)$ is a radial function, i.e., $f(x) = g(|x|)$ for $x \in \mathbf{R}^d$, where g is some function on $[0, \infty)$. Then $\widehat{f}$ is also a radial function. Indeed, if T is any rotation of $\mathbf{R}^d$, then

$$\begin{aligned}\widehat{f}(T\xi) &= \int_{\mathbf{R}^d} g(|x|)e^{-ix\cdot T\xi}\,dx = \int_{\mathbf{R}^d} g(|x|)e^{-i(T^{-1}x)\cdot\xi}\,dx = \int_{\mathbf{R}^d} g(|Ty|)e^{-iy\cdot\xi}\,dy \\ &= \int_{\mathbf{R}^d} g(|y|)e^{-iy\cdot\xi}\,dy = \widehat{f}(\xi) \quad \text{for } \xi \in \mathbf{R}^d.\end{aligned}$$

This shows that $\widehat{f}(\xi)$ only depends on $|\xi|$, so $\widehat{f}$ is radial. Using polar coordinates $x = \rho\omega$, where $0 \leqslant \rho < \infty$ and $\omega \in S^{d-1}$, we see that

$$\widehat{f}(\xi) = \int_0^\infty g(\rho)\left(\int_{S^{d-1}} e^{-i\rho\omega\cdot\xi}\,d\omega\right)\rho^{d-1}\,d\rho, \quad \xi \in \mathbf{R}^d.$$

One can show that the integral within brackets actually is a Bessel function. □

6.2. Properties of the Fourier Transform

The mapping $\mathscr{F}$, which maps a function $f \in L^1(\mathbf{R}^d)$ onto the function $\widehat{f}$, is also called the Fourier transform. The Fourier transform is obviously linear being an integral:

Proposition 6.2.1. *Suppose that $f, g \in L^1(\mathbf{R}^d)$ and $\alpha, \beta \in \mathbf{C}$. Then*

$$\widehat{\alpha f + \beta g}(\xi) = \alpha\widehat{f}(\xi) + \beta\widehat{g}(\xi) \quad \textit{for every } \xi \in \mathbf{R}^d.$$

In the next proposition, we summarize some simple, but useful properties of the Fourier transform. We will use the following notation:

$$e_h(x) = e^{ih\cdot x}, \quad x \in \mathbf{R}^d,\ h \in \mathbf{R}^d.$$

As before, τ_h is the translation operator in direction $h \in \mathbf{R}^d$, defined by

$$\tau_h f(x) = f(x-h), \quad x \in \mathbf{R}^d.$$

We also use the **reflection operator** R and the **dilation operator** D_t, where t is a non-zero real number, defined by

$$Rf(x) = f(-x),\ x \in \mathbf{R}^d, \quad \text{and} \quad D_t f(x) = f(tx),\ x \in \mathbf{R}^d,$$

respectively. Here, f denotes a function on $\mathbf{R}^d$.

Proposition 6.2.2. *Suppose that $f \in L^1(\mathbf{R}^d)$. Then the following properties hold for $h \in \mathbf{R}^d$ and $t \neq 0$:*

(i) $\widehat{e_h f} = \tau_h f$;

(ii) $\widehat{\tau_h f} = e_{-h}\widehat{f}$;

(iii) $\widehat{Rf} = R\widehat{f}$;

(iv) $\widehat{D_t f} = |t|^{-d} D_{t^{-1}}\widehat{f}$;

(v) $\hat{\hat{f}} = \overline{R\hat{f}}$.

The proof is left to the reader as an exercise.

Suppose that $f \in L^1(\mathbf{R}^d)$ and that A is a non-singular $d \times d$ matrix. Put

$$A^* f(x) = f(Ax), \quad x \in \mathbf{R}^d.$$

Proposition 6.2.3. *Suppose that $f \in L^1(\mathbf{R}^d)$ and that A is a non-singular $d \times d$ matrix. Then*

$$\widehat{A^* f}(\xi) = |\det A|^{-1}((A^{-1})^t)^* \hat{f}(\xi), \quad \xi \in \mathbf{R}^d. \tag{6.1}$$

Proof. Changing variables $y = Ax$, we see that

$$\widehat{A^* f}(\xi) = \int_{\mathbf{R}^d} f(Ax) e^{-ix\cdot\xi}\, dx = |\det A|^{-1} \int_{\mathbf{R}^d} f(y) e^{-i(A^{-1}y)\cdot\xi}\, dy.$$

Now, since $(A^{-1}y)\cdot\xi = y \cdot ((A^{-1})^t \xi)$, we have

$$\widehat{A^* f}(\xi) = |\det A|^{-1} \int_{\mathbf{R}^d} f(y) e^{-iy\cdot((A^{-1})^t\xi)}\, dy = |\det A|^{-1}((A^{-1})^t)^* \hat{f}(\xi). \qquad \blacksquare$$

Notice that if $A = -I$, where I is the identity matrix, then A^* is the reflection operator R, and if $A = tI$, where t is a non-zero real number, then A^* is the dilation operator.

Proposition 6.2.4. *Suppose that $f \in L^1(\mathbf{R}^d)$. Then the following properties hold:*

(i) $\hat{f}$ *is bounded on* $\mathbf{R}^d$: $|\hat{f}(\xi)| \leqslant \|f\|_1$ *for every* $\xi \in \mathbf{R}^d$;

(ii) $\hat{f}$ *is uniformly continuous on* $\mathbf{R}^d$;

(iii) $\hat{f}(\xi) \to 0$ *as* $|\xi| \to \infty$.

As for Fourier coefficients, we shall refer to the last property as the **Riemann–Lebesgue lemma**. The first property in this proposition shows that the Fourier transform maps $L^1(\mathbf{R}^d)$ into $L^\infty(\mathbf{R}^d)$, while the second and the third properties show that the image of $L^1(\mathbf{R}^d)$ is a subset to $C_0(\mathbf{R}^d)$.

Proof (Proposition 6.2.4).

(i) This follows directly from the definition of $\hat{f}$.

(ii) Notice that

$$|\hat{f}(\xi + h) - \hat{f}(\xi)| \leqslant \int_{\mathbf{R}^d} |f(x)||e^{-ix\cdot h} - 1|\, dx \quad \text{for } \xi, h \in \mathbf{R}^d.$$

The claim now follows from the dominated convergence theorem since the integrand is less than or equal $2|f(x)|$ and tends to 0 as $h \to 0$. The convergence is uniform because the integral is independent of ξ.

(iii) As in the proof of (ii) in Proposition 2.4.3, we have

$$\hat{f}(\xi) = \frac{1}{2} \int_{\mathbf{R}^d} (f(x) - \tau_{\pi\xi/|\xi|^2} f(x)) e^{-ix\cdot\xi}\, dt \quad \text{for } \xi \neq 0.$$

Finally apply the triangle inequality and Lemma 1.4.1. $\blacksquare$

Example 6.2.5. A consequence of Proposition 6.2.4 is that none of the following functions on $\mathbf{R}$:

$$\xi \mapsto \frac{1}{\xi}, \quad \xi \mapsto \chi_{(-1,1)}(\xi), \quad \xi \mapsto 1$$

is the Fourier transform of a L^1-function. □

One of the most important properties of the Fourier transform is that the transform of a convolution is the product of the transforms of the functions involved. Recall from Theorem 1.2.1 that the convolution $f * g$ is defined a.e. on $\mathbf{R}^d$ and belongs to $L^1(\mathbf{R}^d)$ if $f, g \in L^1(\mathbf{R}^d)$.

Proposition 6.2.6. *Suppose that $f, g \in L^1(\mathbf{R}^d)$. Then*

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\hat{g}(\xi) \quad \text{for } \xi \in \mathbf{R}^d. \tag{6.2}$$

Proof. One proves (6.2) simply by changing the order of integration and performing a linear change of variables:

$$\begin{aligned}\widehat{f * g}(\xi) &= \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} f(x-y) g(y)\, dy \right) e^{-ix\cdot\xi}\, dx \\ &= \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} f(x-y) e^{-i(x-y)\cdot\xi}\, dx \right) g(y) e^{-iy\cdot\xi}\, dy \\ &= \hat{f}(\xi)\hat{g}(\xi) \quad \text{for } \xi \in \mathbf{R}^d.\end{aligned}$$

■

Example 6.2.7. In Section 1.5, we showed that the Banach algebra $L^1(\mathbf{R}^d)$ has no multiplicative unit, i.e., there is no function $K \in L^1(\mathbf{R}^d)$ such that

$$K * f = f \quad \text{for every } f \in L^1(\mathbf{R}^d). \tag{6.3}$$

Let us give a new proof of this fact using the Fourier transform. Suppose that such a function K existed. Let f be the Gauss function in Example 6.1.4. Taking the Fourier transform of both sides in (6.3), we would then have that $\hat{K}\hat{f} = \hat{f}$. Since $\hat{f}$ has no zeroes, this would imply that $\hat{K}(\xi) = 1$ for every $\xi \in \mathbf{R}^d$, which contradicts the Riemann–Lebesgue lemma. □

Proposition 6.2.8. *Suppose that $f, g \in L^1(\mathbf{R}^d)$. Then*

$$\int_{\mathbf{R}^d} f(x)\hat{g}(x)\, dx = \int_{\mathbf{R}^d} \hat{f}(x) g(x)\, dx. \tag{6.4}$$

Notice that both integrals in (6.4) are defined since $\hat{f}$ and $\hat{g}$ are continuous and bounded.

Proof. The identity (6.4) follows directly by changing the order of integration. ■

Proposition 6.2.9. *Suppose that $f \in L^1(\mathbf{R}^d)$ and that $\partial^\alpha f$ exists a.e. and belongs to $L^1(\mathbf{R}^d)$ for some multi-index α. Then*

$$\widehat{\partial^\alpha f}(\xi) = (i\xi)^\alpha \hat{f}(\xi) \quad \text{for every } \xi \in \mathbf{R}^d. \tag{6.5}$$

Proof. Using induction, the proof reduces to showing that (6.5) holds when $|\alpha| = 1$, i.e., when $\partial^\alpha = \partial_j$ for some index j with $1 \leqslant j \leqslant d$. Without loss of generality, we may assume that $j = 1$. We shall write a point $x \in \mathbf{R}^d$ as $x = (x_1, x')$, where $x_1 \in \mathbf{R}$ and $x' \in \mathbf{R}^{d-1}$. Notice that the function $x_1 \mapsto f(x_1, x')$ belongs to $L^1(\mathbf{R})$ for almost every $x' \in \mathbf{R}^{d-1}$ according to Fubini's theorem. For such points $x' \in \mathbf{R}^{d-1}$, we have that

$$f(x_1, x') = f(0, x') + \int_0^{x_1} \partial_1 f(t, x')\, dt \quad \text{for } -\infty < x_1 < \infty.$$

This identity shows that the limits $\lim_{x_1 \to \pm\infty} f(x, x')$ exist. These limits have to be 0 since $f(x_1, x') \in L^1(\mathbf{R})$. We now obtain (6.5) by integrating the one-dimesional Fourier transform of $\partial_1 f(x_1, x')$ by parts:

$$\begin{aligned}\widehat{\partial_j f}(\xi) &= \int_{\mathbf{R}^{d-1}} \left(\int_{-\infty}^{\infty} \partial_1 f(x_1, x') e^{-ix_1\xi_1}\, dx_1 \right) e^{-ix'\cdot\xi'}\, dx' \\ &= i\xi_1 \int_{\mathbf{R}^{d-1}} \left(\int_{-\infty}^{\infty} f(x_1, x') e^{-ix_1\xi_1}\, dx_1 \right) e^{-ix'\cdot\xi'}\, dx' \\ &= i\xi_1 \hat{f}(\xi).\end{aligned}$$

■

The Riemann–Lebesgue lemma shows that $\hat{f}(\xi) = o(1)$ as $|\xi| \to \infty$ if $f \in L^1(\mathbf{R}^d)$. As for Fourier coefficients, the Fourier transform will decay faster the more regular the function f is:

Corollary 6.2.10. *Suppose that $f \in L^1(\mathbf{R}^d)$ and that $\partial^\alpha f$ exists a.e. and belongs to $L^1(\mathbf{R}^d)$ for some multi-index α. Then*

$$\hat{f}(\xi) = o(|\xi|^{-|\alpha|}) \quad \text{as } |\xi| \to \infty.$$

Proof. Since $\partial^\alpha f \in L^1(\mathbf{R}^d)$, $|\xi|^{|\alpha|}|\hat{f}(\xi)| = |\widehat{\partial^\alpha f}(\xi)| \to 0$ as $|\xi| \to \infty$. ■

Proposition 6.2.11. *Suppose that $f \in L^1(\mathbf{R})$ and that $\int_{\mathbf{R}^d} |x|^k |f(x)|\, dx < \infty$ for some integer $k \geqslant 1$. Then $\hat{f} \in C^k(\mathbf{R}^d)$ and*

$$\partial^\alpha \hat{f}(\xi) = \int_{\mathbf{R}^d} (-ix)^\alpha f(x) e^{-ix\cdot\xi}\, dx \quad \text{for } |\alpha| \leqslant k \text{ and } \xi \in \mathbf{R}^d. \tag{6.6}$$

We remark that (6.6) is exactly what one obtains by formally differentiating $\hat{f}$ under the integral sign:

$$\begin{aligned}\partial^\alpha \hat{f}(\xi) &= \partial^\alpha \int_{\mathbf{R}^d} f(x) e^{-ix\cdot\xi}\, dx = \int_{\mathbf{R}^d} f(x) \partial_\xi^\alpha e^{-ix\cdot\xi}\, dx \\ &= \int_{\mathbf{R}^d} (-ix)^\alpha f(x) e^{-ix\cdot\xi}\, dx.\end{aligned}$$

Proof. It suffices to prove (6.6) for $k = 1$ and we may assume that $\alpha = (1, 0, \dots, 0)$. Writing $x = (x_1, x')$ and $\xi = (\xi_1, \xi')$, where $x', \xi' \in \mathbf{R}^{d-1}$, we then have that

$$\frac{\hat{f}(\xi_1 + h, \xi') - \hat{f}(\xi_1, \xi')}{h} = \int_{\mathbf{R}^d} \int_{-\infty}^{\infty} (-ix_1) f(x_1, x') \frac{e^{-ix_1 h} - 1}{-ix_1 h} e^{-i(x_1\xi_1 + x'\cdot\xi')}\, dx_1 dx'.$$

Since the differential quotient tends to 1 as $h \to 0$ and its absolute value is less than or equal 1, (6.6) follows from the dominated convergence theorem. The continuity of $\partial_1 \widehat{f}$ is a consequence of the fact that the right-hand side in (6.6) is a continuous function of ξ (this also follows from the dominated convergence theorem). ∎

Example 6.2.12. We shall calculate the Fourier transform of the function

$$f(x) = e^{-x^2/2}, \quad x \in \mathbf{R},$$

in Example 6.1.4 in a another way. Notice that $f'(x) = -xf(x)$ for every $x \in \mathbf{R}$. If we apply the Fourier transform to this identity, using (6.5) and (6.6), we obtain that

$$i\xi \widehat{f}(\xi) = -i\widehat{f}'(\xi) \quad \text{for every } \xi \in \mathbf{R}.$$

Every solution to this differential equation has the form $\widehat{f}(\xi) = Ce^{-t^2/2}$, $\xi \in \mathbf{R}$, for some constant C. In this case,

$$C = \widehat{f}(0) = \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi},$$

so that $\widehat{f}(\xi) = \sqrt{2\pi} e^{-\xi^2/2}$ for $\xi \in \mathbf{R}$. □

6.3. Inversion of Fourier Transforms in One Dimension

We next turn our attention to inversion of Fourier transforms and begin with the simpler one-dimensional case. The results (and the methods used for obtaining them) are very similar to the results about pointwise convergence of Fourier series in Chapter 2.

Let us first define an operator that corresponds to the symmetric partial sum for the Fourier series of a periodic function. For $f \in L^1(\mathbf{R})$ and $N \geqslant 0$, put

$$S_N f(x) = \frac{1}{2\pi} \int_{-N}^{N} \widehat{f}(\xi) e^{i\xi x}\, d\xi, \quad x \in \mathbf{R}.$$

Using the definition of $\widehat{f}$, we see that

$$\begin{aligned} S_N f(x) &= \frac{1}{2\pi} \int_{-N}^{N} \left(\int_{-\infty}^{\infty} f(y) e^{-iy\xi}\, dy \right) e^{i\xi x}\, d\xi = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \left(\int_{-N}^{N} e^{i\xi(x-y)}\, d\xi \right) dy \\ &= \int_{-\infty}^{\infty} f(y) \frac{\sin N(x-y)}{\pi(x-y)}\, ds = D_N * f(x), \end{aligned}$$

where D_N is the **Dirichlet kernel** for the real line:

$$D_N(x) = \frac{\sin Nx}{\pi x}, \quad x \in \mathbf{R},\ N \geqslant 0.$$

Using the fact that D_N is an even function, we can also write $S_N f$ as

$$S_N f(x) = \frac{1}{\pi} \int_0^{\infty} \frac{f(x+y) + f(x-y)}{y} \sin Ny\, dy.$$

The following results are proved in the same way as the corresponding results for Fourier series.

Proposition 6.3.1. *Suppose that $f \in L^1(\mathbf{R})$ and $\delta > 0$. Then*

$$S_N f(x) = \frac{1}{\pi}\int_0^\delta \frac{f(x+y)+f(x-y)}{y}\sin Ny\,dy + \varepsilon_N(x)$$

for every $x \in \mathbf{R}$, where $\varepsilon_N(x) \to 0$ as $N \to \infty$.

Proposition 6.3.2. *Suppose that $f \in L^1(\mathbf{R})$. Then $\lim_{N\to\infty} S_N f(x) = S$ if and only there exists a number $\delta > 0$ such that*

$$\lim_{N\to\infty}\frac{1}{\pi}\int_0^\delta \frac{f(x+y)+f(x-y)-2S}{y}\sin Ny\,dy = 0.$$

Theorem 6.3.3. *Suppose that $f \in L^1(\mathbf{R})$ satisfies a Dini condition at $x \in \mathbf{R}$, i.e., there exist numbers $\delta > 0$ and $S \in \mathbf{C}$ such that*

$$\int_0^\delta \frac{|f(x+y)+f(x-y)-2S|}{y}\,dy < \infty.$$

Then $\lim_{N\to\infty} S_N f(x) = S$.

In particular, if

$$\int_0^\delta \frac{|f(x+y)+f(x-y)-2f(x)|}{y}\,dy < \infty$$

for some number $\delta > 0$, then

$$f(x) = \lim_{N\to\infty}\frac{1}{2\pi}\int_{-N}^N \widehat{f}(\xi)e^{i\xi x}\,d\xi.$$

One calls the limit in the right-hand side a **principal value integral**. Notice that the principal value cannot be replaced with an integral over $\mathbf{R}$, since $\widehat{f}$ in general does not belong to $L^1(\mathbf{R})$ (cf. Example 6.1.2).

Corollary 6.3.4. *Suppose that $f \in L^1(\mathbf{R})$. If the one-sided limits*

$$f(x^+) = \lim_{y\to 0^+} f(x+y) \quad \textit{and} \quad f(x^-) = \lim_{y\to 0^+} f(x-y)$$

and the one-sided derivatives

$$f'(x^+) = \lim_{y\to 0^+}\frac{f(x+y)-f(x^+)}{y} \quad \textit{and} \quad f'(x^-) = \lim_{y\to 0^+}\frac{f(x-y)-f(x^-)}{-y}$$

exist, then

$$\lim_{N\to\infty} S_N f(x) = \frac{f(x^+)+f(x^-)}{2}.$$

Example 6.3.5. According to Example 6.1.2 and Corollary 6.3.4,

$$\lim_{N\to\infty}\frac{1}{2\pi}\int_{-N}^N \frac{2\sin\xi}{\xi}e^{i\xi x}\,d\xi = \begin{cases} 1 \text{ if } -1 < x < 1 \\ \frac{1}{2} \text{ if } x = \pm 1 \\ 0 \text{ if } x > 1 \text{ or } x < -1 \end{cases}. \qquad \square$$

Example 6.3.6. According to Example 6.1.3 and Corollary 6.3.4,

$$e^{-|x|} = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-N}^{N} \frac{2}{1+\xi^2} e^{i\xi x}\, d\xi = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{1+\xi^2} e^{i\xi x}\, d\xi \quad \text{for } x \in \mathbf{R},$$

where the last equality holds because the integrand belongs to $L^1(\mathbf{R})$. If we now replace x with $-x$ and let x and ξ change roles in this identity, we obtain that

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} e^{-ix\xi}\, dx = \pi e^{-|\xi|} \quad \text{for } \xi \in \mathbf{R}.$$

This shows that the Fourier transform of the function

$$f(x) = \frac{1}{1+x^2},\ x \in \mathbf{R}, \quad \text{is} \quad \widehat{f}(\xi) = \pi e^{-|\xi|},\ \xi \in \mathbf{R}. \qquad \square$$

Corollary 6.3.7. *Suppose that $f \in L^1(\mathbf{R})$. If f satisfies a Hölder condition at a point $x \in \mathbf{R}$, then $\lim_{N\to\infty} S_N f(x) = f(x)$.*

6.4. Inversion of Fourier Transforms in Several Dimensions

Inversion of Fourier transforms in more than one dimension is considerably harder than in the one-dimensional case, the main reason being the fact the Fourier transform $\widehat{f}$ of a L^1-function f not necessarily is integrable, which makes the interpretation of the inversion formula, i.e.,

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \widehat{f}(\xi) e^{i\xi\cdot x}\, d\xi, \quad x \in \mathbf{R}^d,$$

very delicate. We will therefore focus on the simpler case when $\widehat{f} \in L^1(\mathbf{R}^d)$.

Theorem 6.4.1. *Suppose that both f and $\widehat{f}$ belong to $L^1(\mathbf{R}^d)$. Then*

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \widehat{f}(\xi) e^{i\xi\cdot x}\, d\xi \quad \textit{for a.e. } x \in \mathbf{R}^d. \tag{6.7}$$

If f, in addition, is bounded on $\mathbf{R}^d$, then (6.7) holds at every $x \in \mathbf{R}^d$, where f is continuous.

Proof. For $y \in \mathbf{R}^d$, put

$$\phi(y) = (2\pi)^{-d} e^{-|y|^2/2} \quad \text{and} \quad \psi(y) = e^{ix\cdot y}\phi(\varepsilon y),$$

where $\varepsilon > 0$ and $x \in \mathbf{R}^d$ are parameters. Then

$$\widehat{\phi}(\xi) = (2\pi)^{-d/2} e^{-|\xi|^2/2} \quad \text{and} \quad \widehat{\psi}(\xi) = \widehat{\phi}_\varepsilon(\xi - x)$$

for $\xi \in \mathbf{R}^d$, where $\widehat{\phi}_\varepsilon(\xi) = \varepsilon^{-d}\widehat{\phi}(\varepsilon^{-1}\xi)$. Notice that $\widehat{\phi}_\varepsilon$ is even. Proposition 6.2.8 now shows that

$$\int_{\mathbf{R}^d} f(\xi)\widehat{\phi}_\varepsilon(x-\xi)\, d\xi = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \widehat{f}(\xi) e^{-\varepsilon^2|\xi|^2/2} e^{i\xi\cdot x}\, d\xi. \tag{6.8}$$

Letting $\varepsilon \to 0$, the right-hand side in this identity tends to the right-hand side in (6.7) due to dominated convergence. Since $(\widehat{\phi}_\varepsilon)_{\varepsilon>0}$ is an approximate identity, the left-hand side tends to f in $L^1(\mathbf{R}^d)$ (see Theorem 1.5.3). If we now choose a subsequence ε_k such that $\widehat{\phi}_{\varepsilon_k} * f \to f$ a.e. as $k \to \infty$ and replace ε by ε_k in (6.8), we obtain (6.7). The final statement also follows from Theorem 1.5.3. ∎

Definition 6.4.2. The **inverse Fourier transform** $\check{f}$ of a function $f \in L^1(\mathbf{R}^d)$ is defined by

$$\check{f}(x) = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} f(\xi) e^{i\xi\cdot x}\, d\xi, \quad x \in \mathbf{R}^d.$$

The inverse Fourier transform shares most properties with the Fourier transform. With this notation, Theorem 6.4.1 may be reformulated as

$$\check{\widehat{f}}(x) = f(x) \quad \text{for a.e. } x \in \mathbf{R}^d$$

assuming that f and $\widehat{f}$ belong to $L^1(\mathbf{R}^d)$.

As a corollary to Theorem 6.4.1, we obtain the following uniqueness theorem for the Fourier transform.

Theorem 6.4.3. *Suppose that $f, g \in L^1(\mathbf{R}^d)$. If $\widehat{f} = \widehat{g}$, then $f = g$ a.e.*

Proof. Put $h = f - g$. Then $\widehat{h} = 0 \in L^1(\mathbf{R}^d)$, so it follows from Theorem 6.4.1 that $h = 0$ a.e. and thus that $f = g$ a.e. ∎

We end this section by giving a simple criterion for when the Fourier transform of a L^1-function belongs to $L^1(\mathbf{R}^d)$.

Proposition 6.4.4. *Suppose that $f \in L^1(\mathbf{R}^d)$, that there exist positive constants C and M such that $|f(x)| \leqslant C$ for $|x| \leqslant M$, and that $\widehat{f} \geqslant 0$. Then $\widehat{f} \in L^1(\mathbf{R}^d)$.*

Proof. The proof is quite similar to that of Theorem 6.4.1. Put

$$\phi(x) = (2\pi)^{-d} e^{-|x|^2/2}, \ x \in \mathbf{R}^d,$$

and

$$\psi(x) = \phi(\varepsilon x), \ x \in \mathbf{R}^d,$$

where $\varepsilon > 0$. Then $\widehat{\psi}(\xi) = \widehat{\phi}_\varepsilon(\xi)$ for $\xi \in \mathbf{R}^d$. Proposition 6.2.8 now shows that

$$\frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \widehat{f}(\xi) e^{-\varepsilon^2|\xi|^2/2}\, d\xi = \left| \int_{\mathbf{R}^d} f(\xi)\widehat{\phi}_\varepsilon(\xi)\, d\xi \right| \leqslant \int_{\mathbf{R}^d} |f(\xi)|\widehat{\phi}_\varepsilon(\xi)\, d\xi. \tag{6.9}$$

We next split the integral in the right member of (6.9) as follows:

$$\begin{aligned}\int_{\mathbf{R}^d} |f(\xi)|\widehat{\phi}_\varepsilon(\xi)\, d\xi &= \int_{|\xi|<M} |f(\xi)|\widehat{\phi}_\varepsilon(\xi)\, d\xi + \int_{|\xi|\geqslant M} |f(\xi)|\widehat{\phi}_\varepsilon(\xi)\, d\xi \\ &\leqslant C \int_{\mathbf{R}^d} \widehat{\phi}_\varepsilon(\xi)\, d\xi + (2\pi)^{-d/2} \frac{e^{-M^2/2\varepsilon^2}}{\varepsilon^d} \int_{|\xi|\geqslant M} |f(\xi)|\, d\xi \\ &\leqslant (2\pi)^{-d/2}(C + \|f\|_1),\end{aligned}$$

where the last inequality holds for sufficiently small ε. This shows that

$$\int_{\mathbf{R}^d} \widehat{f}(\xi) e^{-\varepsilon^2|\xi|^2/2}\, d\xi \leqslant (2\pi)^{d/2}(C + \|f\|_1)$$

for every sufficiently small $\varepsilon > 0$. If we now let $\varepsilon \to 0$ and apply the monotone convergence theorem, we obtain that $\widehat{f} \in L^1(\mathbf{R}^d)$. ■

Chapter 7

L^2-theory for Fourier Transforms

In this chapter, we will show how the Fourier transform can be extended to functions $f \in L^2(\mathbf{R}^d)$. We also prove the celebrated Plancherel formula and show how inversion of Fourier transforms works in $L^2(\mathbf{R}^d)$.

7.1. Definition of the Fourier Transform

The strategy for extending the Fourier transform to $L^2(\mathbf{R}^d)$ is the following: One first chooses a sequence of functions $f_n \in L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$ such that $f_n \to f$ in $L^2(\mathbf{R}^d)$. Since each f_n belongs to $L^1(\mathbf{R}^d)$, it has a Fourier transform $\widehat{f_n}$. The next step is to prove that the sequence $\widehat{f_n}$ is convergent in $L^2(\mathbf{R}^d)$. The limit of this sequence is then defined as the Fourier transform $\widehat{f}$ of f. To prove that this extension is consistent with the previous definition, one has to verify that the two definitions coincide for functions in $L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$. One also needs to verify that $\widehat{f}$ is independent of the choice of the sequence f_n.

Given a function $f \in L^2(\mathbf{R}^d)$, we define let $f_n = f|_{B_n(0)}$, i.e.,

$$f_n(x) = \begin{cases} f(x) & \text{if} \quad |x| < n \\ 0 & \text{if} \quad |x| \geqslant n \end{cases} \tag{7.1}$$

for $n = 1, 2, \ldots$. Every function f_n of course belongs to $L^2(\mathbf{R}^d)$. We first show that f_n is integrable, and thus has a Fourier transform, and that the sequence approximates f in $L^2(\mathbf{R}^d)$.

Lemma 7.1.1. *Suppose that $f \in L^2(\mathbf{R}^d)$ and that f_n is given by (7.1). Then*
(a) $f_n \in L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$ for every n;
(b) $f_n \to f$ in $L^2(\mathbf{R}^d)$.

Proof.
(a) This follows directly from Hölder's inequality:

$$\|f_n\|_1 = \int_{|x|<n} |f(x)|\, dx \leqslant C n^{d/2} \|f\|_2 < \infty.$$

(b) Notice that

$$\|f - f_n\|_2^2 = \int_{|x|\geqslant n} |f(x)|^2\, dx = \int_{\mathbf{R}^d} \chi_n(x) |f(x)|^2\, dx,$$

where χ_n is the characteristic function of $\mathbf{R}^d \setminus B_n(0)$. The integral in the right-hand side tends to 0 as $n \to \infty$ since the integrand tends to 0 a.e. and it is dominated by the integrable function $|f|^2$. ■

We next prove a weak form of the Plancherel formula.

Lemma 7.1.2. *Suppose that $f \in L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$. Then $\widehat{f} \in L^2(\mathbf{R}^d)$ and*

$$\int_{\mathbf{R}^d} |f(x)|^2\, dx = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} |\widehat{f}(\xi)|^2\, d\xi. \tag{7.2}$$

Proof. Let $g = f * \overline{Rf}$ be the so called **autocorrelation function**, where Rf is defined as before by $Rf(x) = f(-x)$, $x \in \mathbf{R}^d$. Thus,

$$g(x) = \int_{R^d} f(y)\overline{f(y-x)}\,dy \quad \text{for } x \in \mathbf{R}^d.$$

Then $g \in L^1(\mathbf{R}^d)$ since $f \in L^1(\mathbf{R}^d)$ (see Theorem 1.2.1). The Fourier transform of the function $\overline{Rf}$ is $\overline{\widehat{f}}$ according to Proposition 6.2.2, so it follows from Proposition 6.2.6 that $\widehat{g} = |\widehat{f}|^2$. The assumption that $f \in L^2(\mathbf{R}^d)$ moreover implies that g is bounded: $|g(x)| \leqslant \|f\|_2^2$ for every $x \in \mathbf{R}^d$. Proposition 6.4.4 thus shows that $\widehat{g} \in L^1(\mathbf{R}^d)$. This means that we can apply the inversion formula in Theorem 6.4.1:

$$\int_{R^d} f(y)\overline{f(y-x)}\,ds = \frac{1}{(2\pi)^d}\int_{R^d} |\widehat{f}(\xi)|^2 e^{i\xi\cdot x}\,d\xi \quad \text{for a.e. } x \in \mathbf{R}^d. \tag{7.3}$$

But since both sides of this identity are continuous functions (see Theorem 1.4.2 for the left-hand side and Proposition 6.2.4 for the right-hand side), it holds for every $x \in \mathbf{R}^d$. Hence, (7.2) follows if we take $x = 0$ in (7.3). ■

Lemma 7.1.3. *Suppose that $f \in L^1(\mathbf{R}^d)$ and that f_n is given by* (7.1). *Then the sequence $(\widehat{f}_n)_{n=1}^\infty$ is convergent in $L^2(\mathbf{R}^d)$.*

Proof. Using (7.2) and the fact that f_n converges to f, we see that $(\widehat{f}_n)_{n=1}^\infty$ is a Cauchy sequence in $L^2(\mathbf{R}^d)$:

$$\|\widehat{f}_m - \widehat{f}_n\|_2 = (2\pi)^{d/2}\|f_m - f_n\|_2 \longrightarrow 0 \quad \text{as } m, n \to \infty.$$ ■

Definition 7.1.4. If $f \in L^2(\mathbf{R}^d)$ and f_n is given by (7.1), we define $\widehat{f} \in L^2(\mathbf{R}^d)$ as the limit in $L^2(\mathbf{R}^d)$ of the sequence $(\widehat{f}_n)_{n=1}^\infty$.

Remark 7.1.5.

(a) Notice that the Fourier transform maps $L^2(\mathbf{R}^d)$ into $L^2(\mathbf{R}^d)$.

(b) By definition,

$$\widehat{f}(\xi) = \lim_{n\to\infty}\int_{|x|<n} f(x)e^{-ix\cdot\xi}\,dx \quad \text{in } L^2(\mathbf{R}^d), \tag{7.4}$$

which means that

$$\int_{\mathbf{R}^d}\left|\widehat{f}(\xi) - \int_{|x|<n} f(x)e^{-ix\cdot\xi}\,dx\right|^2 d\xi \longrightarrow 0 \quad \text{as } n \to \infty.$$

(c) In the case $f \in L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$, the transform $\widehat{f}$ defined by (7.4) coincides with Definition 6.1.1. In fact, there exists a subsequence n_k such that $n_k \to \infty$ as $k \to \infty$ and

$$\widehat{f}(\xi) = \lim_{k\to\infty}\int_{|x|<n_k} f(x)e^{-ix\cdot\xi}\,dx \quad \text{for a.e. } \xi \in \mathbf{R}^d.$$

But since $f \in L^1(\mathbf{R}^d)$, the right-hand side equals $\int_{\mathbf{R}^d} f(x)e^{-ix\cdot\xi}\,dx$.

(d) If we choose another sequence of functions $g_n \in L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$, that converges to f in $L^2(\mathbf{R}^d)$, then

$$\|\widehat{f} - \widehat{g}_n\|_2 \leqslant \|\widehat{f} - \widehat{f}_n\|_2 + \|\widehat{f}_n - \widehat{g}_n\|_2 = \|\widehat{f} - \widehat{f}_n\|_2 + (2\pi)^{d/2}\|f_n - g_n\|_2,$$

which shows that $\widehat{g}_n$ converges to $\widehat{f}$ in $L^2(\mathbf{R}^d)$. This means that the definition is independent of the sequence f_n.

Example 7.1.6. Let $f(x) = \sin x/x,\ x \in \mathbf{R}$. Notice that f belongs to $L^2(\mathbf{R})$, but not to $L^1(\mathbf{R})$. According to Example 6.3.5,

$$\lim_{n\to\infty} \int_{-n}^{n} \frac{\sin x}{x} e^{-ix\xi}\, dx = \pi\chi_{(-1,1)}(x) \quad \text{for } x \neq \pm 1.$$

It thus follows from (7.4) that $\widehat{f} = \pi\chi_{(-1,1)}$. □

Notice that this example shows that the Fourier transform of an L^2-function is not necessarily continuous as was the case for L^1-functions.

7.2. Plancherel's Formula

We next extend the Plancherel formula to $L^2(\mathbf{R}^d)$.

Theorem 7.2.1. *Suppose that $f \in L^2(\mathbf{R}^d)$. Then*

$$\int_{\mathbf{R}^d} |f(x)|^2\, dx = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} |\widehat{f}(\xi)|^2\, d\xi. \tag{7.5}$$

Proof. Since f_n and $\widehat{f}_n$ converge to f and $\widehat{f}$ in $L^2(\mathbf{R}^d)$, respectively, and Plancherel's formula holds for f_n, we obtain that

$$\|f\|_2^2 = \lim_{n\to\infty} \|f_n\|_2^2 = \lim_{n\to\infty} \frac{1}{(2\pi)^d}\|\widehat{f}_n\|_2^2 = \frac{1}{(2\pi)^d}\|\widehat{f}\|_2^2.$$ ■

Example 7.2.2. If we apply Plancherel's formula to the function in Example 7.1.6, we see that

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2}\, dx = \frac{1}{2\pi}\int_{-1}^{1} \pi^2\, d\xi = \pi.$$ □

Corollary 7.2.3. *Suppose that $f, g \in L^2(\mathbf{R}^d)$. Then*

$$\int_{\mathbf{R}^d} f(x)\overline{g(x)}\, dx = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, d\xi.$$

Proof. Apply Plancherel's formula to $f + g$ and $f + ig$. ■

7.3. Properties of the Fourier Transform

The Fourier transform on $L^2(\mathbf{R}^d)$ shares many properties with the Fourier transform on $L^1(\mathbf{R}^d)$.

Proposition 7.3.1. *Suppose that $f \in L^2(\mathbf{R}^d)$. Then the following properties hold in $L^2(\mathbf{R}^d)$:*

(i) $\widehat{\tau_h f} = e_{-h}\widehat{f}$ *for* $h \in \mathbf{R}^d$;

(ii) $\widehat{e_h f} = \tau_h f$ *for* $h \in \mathbf{R}^d$;

(iii) $\widehat{Rf} = R\widehat{f}$;

(iv) $\widehat{D_t f} = |t|^{-d} D_{t^{-1}}\widehat{f}$ *for* $t \neq 0$;

(v) $\widehat{\widehat{f}} = R\overline{\widehat{f}}$

(vi) if $\partial_j f \in L^2(\mathbf{R}^d)$ *for some* j, *then* $\widehat{\partial_j f}(\xi) = i\xi_j \widehat{f}(\xi)$.

Proposition 7.3.2. *Suppose that $f, g \in L^2(\mathbf{R}^d)$. Then*

$$\int_{\mathbf{R}^d} f(x)\widehat{g}(x)\,dx = \int_{\mathbf{R}^d} \widehat{f}(x)g(x)\,dx.$$

Proof. Let $(f_n)_{n=1}^{\infty}$ and $(g_n)_{n=1}^{\infty}$ be any two sequences in $L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$ such that $f_n \to f$ and $g_n \to f$. Then

$$\begin{aligned}\int_{\mathbf{R}^d} f(x)\widehat{g}(x)\,dx &= \lim_{n\to\infty}\int_{\mathbf{R}^d} f_n(x)\widehat{g_n}(x)\,dx = \lim_{n\to\infty}\int_{\mathbf{R}^d} \widehat{f_n}(x)g_n(x)\,dx \\ &= \int_{\mathbf{R}^d} \widehat{f}(x)g(x)\,dx,\end{aligned}$$

where the second equality follows from Proposition 6.2.8. ■

7.4. The Inversion Formula

Recall the inversion formula proved in Section 6.4: If $f, \widehat{f} \in L^1(\mathbf{R}^d)$, then $\check{\widehat{f}} = f$, where

$$\check{f}(x) = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} f(\xi)e^{i\xi\cdot x}\,d\xi, \quad x \in \mathbf{R}^d.$$

Notice also that $\check{f} = (2\pi)^{-d}R\widehat{f}$. This motivates the following definiteion.

Definition 7.4.1. The **inverse Fourier transform** $\check{f}$ of $f \in L^2(\mathbf{R}^d)$ is defined by $\check{f} = (2\pi)^{-d}R\widehat{f}$.

Theorem 7.4.2. *Suppose that $f \in L^2(\mathbf{R}^d)$. Then $\check{\check{f}} = f$.*

Combining this result with (7.4), we see that

$$f(x) = \lim_{n\to\infty}\frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \widehat{f}(\xi)e^{i\xi\cdot x}\,d\xi \quad \text{in } L^2(\mathbf{R}^d).$$

Proof. Put $g = \check{\check{f}}$. We will prove that $f = g$ a.e. by showing that $\|f - g\|_2 = 0$. To this end, notice that

$$\|f - g\|_2^2 = (f - g, f - g) = \|f\|_2^2 - (f, g) - \overline{(f, g)} + \|g\|_2^2.$$

Using the fact that $\overline{g} = (2\pi)^{-d}\hat{\overline{\hat{f}}}$ (see property (v) in Proposition 7.3.1) together with Proposition 7.3.2, we obtain that

$$(f, g) = \int_{\mathbf{R}^d} f(x)\overline{g(x)}\, dx = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} f(x)\hat{\overline{\hat{f}}}(x)\, dx = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \hat{f}(x)\overline{\hat{f}}(x)\, dx$$
$$= \|f\|_2^2$$

and consequently $\overline{(f, g)} = \|f\|_2^2$. Finally, two applications of the Plancherel formula yield $\|g\|_2^2 = \|f\|_2^2$. This shows that $\|f - g\|_2 = 0$. ■

Example 7.4.3. Let us check the inversion formula for the function $f = \chi_{(-1,1)}$. Then

$$\hat{f}(\xi) = 2\frac{\sin\xi}{\xi} \quad \text{and} \quad \hat{\hat{f}}(x) = 2\pi\chi_{(-1,1)}(x), \quad \text{so that} \quad \check{\hat{f}}(x) = f(x). \qquad \square$$

Let $\mathscr{F}$ denote the operator with maps a function $f \in L^2(\mathbf{R}^d)$ onto its Fourier transform $\hat{f}$. By combining Plancherel's formula with the inversion formula, we obtain the following theorem.

Theorem 7.4.4. *The operator* $(2\pi)^{d/2}\mathscr{F}$ *from* $L^2(\mathbf{R}^d)$ *to* $L^2(\mathbf{R}^d)$ *is an isometric isomorphism.*

Part IV

Distribution Theory

Chapter 8
Distributions

In this chapter, X and K will denote open and compact subsets to $\mathbf{R}^d$, respectively.

8.1. Test functions

In the context of distribution theory, the class of infinitely continuously differentiable functions on X with compact support is traditionally denoted $\mathscr{D}(X)$ instead of $C_c^\infty(X)$ and the functions, that belong to $\mathscr{D}(X)$, are called **test functions**.

Example 8.1.1. In Example 1.6.2, we gave the following example of a function ϕ that belongs to $\mathscr{D}(\mathbf{R}^d)$:

$$\phi(x) = \begin{cases} e^{-1/(1-|x|^2)} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geqslant 1 \end{cases}. \qquad \square$$

Definition 8.1.2. A sequence $(\phi_n)_{n=1}^\infty \subset \mathscr{D}(X)$ **converges** to $\phi \in \mathscr{D}(X)$ if

(i) there exists a compact subset K to X such that $\operatorname{supp}\phi_n \subset K$ for every n;

(ii) $\partial^\alpha\phi_n$ converges uniformly to $\partial^\alpha\phi$ on X for every multi-index α.

We denote this by writing $\phi_n \to \phi$.

Remark 8.1.3.

(a) We remark that there are corresponding definitions for sequences like $(\phi_h)_{h>0}$, where $h \to 0$, etc.

(b) Notice that if $\phi_n \to \phi$ and $\operatorname{supp}\phi_n \subset K$ for every n, then $\operatorname{supp}\phi \subset K$.

Example 8.1.4. Suppose that $\phi \in \mathscr{D}(\mathbf{R}^d)$. Then $\tau_h\phi \to \phi$ as $h \to 0$. Indeed, the support of $\tau_h\phi$ is a subset of the closed $|h|$-neighbourhood of $\operatorname{supp}\phi$. Also, if α is some multi-index and $x \in \mathbf{R}^d$, then

$$|\partial^\alpha\phi(x-h) - \partial^\alpha\phi(x)| = |\nabla\partial^\alpha\phi(x-\theta h)\cdot h| \leqslant |\nabla\partial^\alpha\phi(x-\theta h)||h| \leqslant \|\nabla\partial^\alpha\phi\|_\infty|h|$$

according to the mean-value theorem and the Cauchy-Schwarz inequality, where θ is some number between 0 and 1, so that

$$\|\partial^\alpha\tau_h\phi - \partial^\alpha\phi\|_\infty \leqslant \|\nabla\partial^\alpha\phi\|_\infty|h|.$$

This shows that $\partial^\alpha\tau_h\phi$ tends uniformly to $\partial^\alpha\phi$ as $h \to 0$. $\square$

Example 8.1.5. Let e_j be the j-th vector in the standard basis for $\mathbf{R}^d$. We claim that if $\phi \in \mathscr{D}(X)$, then

$$\frac{\phi(x+he_j)-\phi(x)}{h} \longrightarrow \partial_j\phi(x) \quad \text{as } h \to 0$$

in $\mathscr{D}(X)$. Two applications of the mean-value theorem show that if $x \in \mathbf{R}^d$, then

$$\begin{aligned} \left|\partial_j\phi(x) - \frac{\phi(x+he_j)-\phi(x)}{h}\right| &= |\partial_j\phi(x) - \partial_j\phi(x+\theta he_j)| \\ &= |\partial_j^2\phi(x+\eta\theta he_j)||\theta h| \\ &\leqslant \|\partial_j^2\phi\|_\infty|h|, \end{aligned}$$

where $\theta, \eta \in [0,1]$, from which the claim follows. $\square$

8.2. Distributions

Definition 8.2.1. A **distribution** on $\mathscr{D}(X)$ is a linear functional $u : \mathscr{D}(X) \to \mathbf{C}$ that is **sequentially continuous**, meaning that if

$$\phi_n \to \phi \text{ in } \mathscr{D}(X), \quad \text{then} \quad u(\phi_n) \to u(\phi).$$

We denote the class of distributions on $\mathscr{D}(X)$ by $\mathscr{D}'(X)$.

Remark 8.2.2.
(a) Notice that $\mathscr{D}'(X)$ is a vector space with the addition and multiplication with scalars defined pointwise.
(b) We shall most of the time write

$$\langle u, \phi \rangle \quad \text{instead of} \quad u(\phi),$$

where $u \in \mathscr{D}'(X)$ and $\phi \in \mathscr{D}(X)$.

8.3. Examples of Distributions

We next give a number examples of distributions.

Example 8.3.1. Every function $f \in L^1_{\text{loc}}(X)$ gives rise to a so-called **regular distribution** u_f on X through integration:

$$\langle u_f, \phi \rangle = \int_X f(x)\phi(x)\,dx, \quad \phi \in \mathscr{D}(X).$$

This mapping is obviously linear. To show that it is sequentially continuous, notice that if $\phi \in \mathscr{D}(X)$ with $\operatorname{supp}\phi \subset K$, then

$$|\langle u_f, \phi \rangle| \leqslant \int_X |f(x)\phi(x)|\,dx \leqslant \|\phi\|_\infty \int_K |f(x)|\,dx.$$

It follows that if $\phi_n \to \phi$ in $\mathscr{D}(X)$ and $\operatorname{supp}\phi_n \subset K$ for every n, then

$$|\langle u_f, \phi \rangle - \langle u_f, \phi_n \rangle| \leqslant \|\phi - \phi_n\|_\infty \int_K |f(x)|\,dx,$$

which shows that $\langle u_f, \phi_n \rangle \to \langle u_f, \phi \rangle$. $\square$

The following proposition shows that there is no need to distinguish between a function $f \in L^1_{\text{loc}}(X)$ and the regular distribution u_f generated by f. We will therefore sometimes denote the distribution u_f by just f.

Proposition 8.3.2. *Suppose that $f, g \in L^1_{\text{loc}}(X)$ and*

$$\langle u_f, \phi \rangle = \langle u_g, \phi \rangle \quad \text{for every} \quad \phi \in \mathscr{D}(X).$$

Then $f = g$ a.e. on X.

Proof. Put $h = f - g$. Then $\langle u_h, \phi\rangle = 0$ for every $\phi \in \mathscr{D}(X)$. Let $K \subset X$ be compact and choose a function $\psi \in \mathscr{D}(X)$ such that $\psi = 1$ on K (see Proposition 1.7.2). Then $\psi h \in L^1(\mathbf{R}^d)$. Moreover, if ϕ is a mollifier on $\mathbf{R}^d$ (see Definition 1.6.1), then

$$\phi_\varepsilon * (\psi h)(x) = \int_{\mathbf{R}^d} \phi_\varepsilon(x-y)\psi(y)h(y)\,dy = 0$$

for every $x \in \mathbf{R}^d$ and every sufficiently small $\varepsilon > 0$. But $\phi_\varepsilon * (\psi h) \to \psi h$ in $L^1(\mathbf{R}^d)$ as $\varepsilon \to 0$, so $\psi h = 0$ in $L^1(\mathbf{R}^d)$. Thus, $\psi h = 0$ a.e. on $\mathbf{R}^d$, so $h = 0$ a.e. on K. Notice that $X = \bigcup_{n=1}^\infty K_n$, where

$$K_n = \{x \in X : |x| \leqslant n \text{ and } \operatorname{dist}(x, X^c) \geqslant n^{-1}\} \quad \text{for } n = 1, 2, \dots .$$

Since every set K_n is compact and $h = 0$ a.e. on K_n, it follows that $h = 0$ a.e. on X and therefore that $f = g$ a.e. on X. ■

Example 8.3.3. The **Dirac delta** δ_a at a point $a \in X$ is defined by

$$\langle \delta_a, \phi\rangle = \phi(a), \quad \phi \in \mathscr{D}(X).$$

One usually denotes δ_0 by just δ. The continuity of δ_a follows as in Example 8.3.1 from the fact that

$$|\langle \delta_a, \phi\rangle| \leqslant \|\phi\|_\infty \quad \text{for every } \phi \in \mathscr{D}(X).$$

This distribution is not regular. In fact, suppose that δ_a were regular. Then there would exist a function $f \in L^1_{\mathrm{loc}}(X)$ such that

$$\int_X f(x)\phi(x)\,dx = \phi(a) \quad \text{for every } \phi \in \mathscr{D}(X).$$

Let ϕ be the test function in Example 8.1.1. Then

$$\phi(0) = \phi(n(a-a)) = \left|\int_X f(x)\phi(n(x-a))\,dx\right| \leqslant \int_{|x-a|\leqslant n^{-1}} |f(x)|\,dx$$

for $n = 1, 2, \dots$. This gives us a contradiction since the left-hand side is non-zero, while the right-hand side tends to 0 as $n \to \infty$. □

Example 8.3.4. In one dimension, the **Cauchy principal value** $\operatorname{pv}\frac{1}{x}$ is defined by

$$\left\langle \operatorname{pv}\frac{1}{x}, \phi(x)\right\rangle = \lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon} \frac{\phi(x)}{x}\,dx, \quad \phi \in \mathscr{D}(\mathbf{R}).$$

The limit in the right-hand side is also denoted

$$\operatorname{pv}\int_{-\infty}^{\infty} \frac{\phi(x)}{x}\,dx.$$

To show that this limit exists, first notice that if $x \in \mathbf{R}$, then

$$\phi(x) - \phi(-x) = \int_{-x}^{x} \phi'(t)\,dt = x\int_{-1}^{1} \phi'(sx)\,ds = x\psi(x),$$

where the function ψ is easily seen to smooth on $\mathbf{R}$. Suppose that $\operatorname{supp}\phi \subset [-R, R]$ for some number $R > 0$. It then follows that

$$\int_{|x|\geqslant\varepsilon} \frac{\phi(x)}{x}\,dx = \int_{\varepsilon}^{R} \frac{\phi(x)-\phi(-x)}{x}\,dx = \int_{\varepsilon}^{R} \psi(x)\,dx \longrightarrow \int_{0}^{R} \psi(x)\,dx$$

as $\varepsilon \to 0$. This also shows that $\operatorname{pv}\frac{1}{x}$ is sequentially continuous:

$$\left|\left\langle \operatorname{pv}\frac{1}{x}, \phi(x)\right\rangle\right| = \left|\int_0^R \psi(x)\,dx\right| \leqslant R \max_{0\leqslant x\leqslant R} |\psi(x)| \leqslant 2R\|\phi'\|_\infty. \tag{8.1}$$

The principal value distribution is not regular. In fact, let ϕ be a mollifier on $\mathbf{R}$ and $x_0 \neq 0$. Theorem 1.6.3 then shows that

$$\left\langle \operatorname{pv}\frac{1}{x}, \phi_\varepsilon(x_0 - x)\right\rangle = \int_{-\infty}^{\infty} \frac{\phi_\varepsilon(x_0 - x)}{x}\,dx \longrightarrow \frac{1}{x_0} \quad \text{as } \varepsilon \to 0.$$

This shows that the only possible candidate for a function, that could generate the principal value, is $f(x) = x^{-1},\ x \neq 0$. But f is not locally integrable. □

8.4. Distributions of Finite Order

In Example 8.3.1, we showed that the functional u_f, generated by a locally integrable function f, is continuous by establishing that, for every compact set $K \subset X$, there exists a constant C_K $(= \int_K |f|\,dx)$ such that

$$|\langle u_f, \phi\rangle| \leqslant C_K\|\phi\|_\infty$$

for every function $\phi \in \mathscr{D}(X)$ with support in K. Basically the same technique was employed in Example 8.3.3 and Example 8.3.4. The next theorem shows that the existence of such an inequality is not only sufficient for a linear functional to be continuous on $\mathscr{D}(X)$, but also necessary.

Theorem 8.4.1. *A linear functional u on $\mathscr{D}(X)$ belongs to $\mathscr{D}'(X)$ if and only if, for every compact subset K of X, there exist a constant $C \geqslant 0$ and an integer $m \geqslant 0$ such that*

$$|u(\phi)| \leqslant C \sum_{|\alpha|\leqslant m} \|\partial^\alpha\phi\|_\infty \tag{8.2}$$

for every function $\phi \in \mathscr{D}(X)$ with support in K.

Proof. The sufficiency of the condition (8.2) is obvious. To prove necessity, we suppose that there exists a compact subset K of X such that (8.2) is not satisfied for any constant C and any integer m. One can then find functions $\phi_n \in \mathscr{D}(X)$ with support in K for which

$$|u(\phi_n)| > n \sum_{|\alpha|\leqslant n} \|\partial^\alpha\phi_n\|_\infty \quad \text{for } n = 1, 2, \dots .$$

By homogeneity, we may assume that $|u(\phi_n)| = 1$ for every n. It then follows that $\|\partial^\alpha\phi_n\|_\infty < 1/n$ if $|\alpha| \leqslant n$, which shows that $\phi_n \to 0$ in $\mathscr{D}(X)$. This is a contradiction since $u(\phi_n)$ does not tend to 0. ■

Definition 8.4.2. A distribution $u \in \mathscr{D}'(X)$ is said to be of **finite order** if (8.2) holds with an integer m that is independent of the set K. The minimal integer m for which (8.2) holds is then called the **order** of u. We denote by $\mathscr{D}'_m(X)$ the class of distributions on X of order less than or equal m.

We remark that if u is of order m, then the constant C in (8.2) will in general depend on K (as in Example 8.3.1 and Example 8.3.4).

Example 8.4.3. The distributions in Example 8.3.1 and Example 8.3.3 are of order 0. The order of the Cauchy principal value in Example 8.3.4 is according to (8.1) not more than 1; we will show that the order is exactly 1. Suppose that the order were 0. This means that there, for every compact set $K \subset \mathbf{R}$, would exist a constant C_K such that

$$\left|\left\langle \operatorname{pv}\frac{1}{x}, \phi(x)\right\rangle\right| \leqslant C_K\|\phi\|_\infty$$

for every function $\phi \in \mathscr{D}(\mathbf{R})$ with support in K. Now take $K = [0,2]$ and let $(\phi_n)_{n=1}^\infty$, be a sequence of function in $\mathscr{D}(\mathbf{R})$ with support in K that satisfies $0 \leqslant \phi_n(x) \leqslant 1$ for every $x \in \mathbf{R}$, $\phi_n(x) = 0$ for $0 \leqslant x \leqslant 1/2n$, and $\phi_n(x) = 1$ for $1/n \leqslant x \leqslant 1$. It then follows that

$$C_K \geqslant \left|\left\langle \operatorname{pv}\frac{1}{x}, \phi_n(x)\right\rangle\right| = \int_{1/2n}^{2} \frac{\phi_n(x)}{x}\,dx \geqslant \int_{1/n}^{1}\frac{dx}{x} = \ln n,$$

which is a contradiction. □

Example 8.4.4. Let the linear functional u on $\mathscr{D}(\mathbf{R})$ be defined by

$$\langle u, \phi\rangle = \sum_{j=0}^{\infty} \phi^{(j)}(j), \quad \phi \in \mathscr{D}(\mathbf{R}).$$

If $\operatorname{supp}\phi \subset [-k,k]$ for some positive integer k, then

$$|\langle u, \phi\rangle| \leqslant \sum_{j=0}^{k-1} \|\phi^{(j)}\|_\infty,$$

which proves that $u \in \mathscr{D}'(\mathbf{R})$ according to Theorem 8.4.1. Suppose that u were of finite order $m \geqslant 0$. Then, for a given compact subset K to $\mathbf{R}$, there would exist a constant C_K such that

$$|\langle u, \phi\rangle| \leqslant C_K \sum_{j=0}^{m} \|\phi^{(j)}\|_\infty \tag{8.3}$$

for every test function ϕ with support in K. Now, take $\phi \in \mathscr{D}(\mathbf{R})$ with support in $(-1,1)$ such that $\phi^{(m+1)}(0) \neq 0$ and put $\phi_n(t) = n^{-m}\phi(n(t-(m+1)))$ for $t \in \mathbf{R}$ and $n = 1, 2, \ldots$. Then $\operatorname{supp}\phi_n \subset (m, m+2)$ for every n, $\|\phi_n^{(j)}\|_\infty \leqslant \|\phi^{(j)}\|_\infty$ for j satisfying $0 \leqslant j \leqslant m$, and $\phi_n^{(m+1)}(m+1) = n\phi^{(m+1)}(0)$. If we now apply (8.3) to $K = [m, m+2]$ and the sequence $(\phi_n)_{n=1}^\infty$, we get a contradiction since the right-hand side is bounded with respect to n, while the left-hand side is unbounded. This shows that u is not of finite order. □

Remark 8.4.5. One can show that if $u \in \mathscr{D}'_m(X)$, then u can be extended to a sequentially continuous functional on $C^m(X)$ in a unique way. It follows that distributions of order 0 are measures on X.

8.5. Convergence in $\mathscr{D}'(X)$

Convergence in $\mathscr{D}'(X)$ is defined as pointwise convergence.[1]

Definition 8.5.1. A sequence $(u_n)_{n=1}^{\infty} \subset \mathscr{D}'(X)$ **converges** to $u \in \mathscr{D}'(X)$ if

$$\langle u_n, \phi \rangle \to \langle u, \phi \rangle \quad \text{for every } \phi \in \mathscr{D}(X).$$

We denote this by writing $u_n \to u$.

Example 8.5.2. Suppose that $\phi \in \mathscr{D}(\mathbf{R})$. According to the Riemann–Lebesgue lemma for the Fourier transform (see Proposition 6.2.4),

$$\langle e^{inx}, \phi(x) \rangle = \int_{-\infty}^{\infty} \phi(x) e^{inx}\, dx = \widehat{\phi}(-n) \longrightarrow 0 \quad \text{as } n \to \infty,$$

which shows that $e^{inx} \to 0$ in $\mathscr{D}'(\mathbf{R})$. □

Example 8.5.3. For $n = 1, 2, \ldots$, let $f_n \in L^1(\mathbf{R})$ be defined by

$$f_n(x) = \begin{cases} n & \text{if} \quad 0 < x < 1/n \\ 0 & \text{if} \quad x \leqslant 0 \text{ or } x \geqslant 1/n \end{cases}.$$

Then $u_{f_n} \to \delta$ in $\mathscr{D}'(\mathbf{R})$. Indeed, if $\phi \in \mathscr{D}(\mathbf{R})$, then

$$\langle u_{f_n}, \phi \rangle = n \int_0^{1/n} \phi(x)\, dx = n \int_0^{1/n} (\phi(x) - \phi(0))\, dx + \phi(0) \longrightarrow \phi(0) = \langle \delta, \phi \rangle$$

as $n \to \infty$ since

$$n \left| \int_0^{1/n} (\phi(x) - \phi(0))\, dx \right| \leqslant \max_{0 \leqslant x \leqslant 1/n} |\phi(x) - \phi(0)| \longrightarrow 0.$$ □

Example 8.5.4. Suppose that $(K_n)_{n=1}^{\infty}$ is an approximate identity on $\mathbf{R}^d$ (see Definition 1.5.1). Theorem 1.5.4 then shows that $u_{K_n} \to \delta$ in $\mathscr{D}'(\mathbf{R}^d)$. Notice also that the function f_n in that Example 8.5.3 can be written $f_n(x) = nK(nx), x \in \mathbf{R}$, where $K = \chi_{(0,1)}$. □

Example 8.5.5. We will show that

$$\sum_{n=-\infty}^{\infty} e^{inx} = 2\pi\delta \quad \text{in } \mathscr{D}'(-\pi, \pi), \tag{8.4}$$

---

[1] In the sense of topological vector spaces, $\mathscr{D}'(X)$ is the dual of $\mathscr{D}(X)$. Convergence in $\mathscr{D}'(X)$ thus coincides with weak* convergence.

where the series is interpreted as the limit of its symmetric partial sums. This is the Fourier series expansion of $2\pi\delta$, which we will return to later. The identity (8.4) holds since

$$\Big\langle \sum_{n=-N}^{N} e^{inx}, \phi(x) \Big\rangle = \sum_{n=-N}^{N} \int_{-\pi}^{\pi} \phi(x) e^{inx}\, dx = 2\pi \sum_{n=-N}^{N} \widehat{\phi}(-n)$$
$$\longrightarrow 2\pi \sum_{n=-\infty}^{\infty} \widehat{\phi}(n) = 2\pi\phi(0) = \langle 2\pi\delta, \phi\rangle \quad \text{as } N \to \infty$$

for every function $\phi \in \mathscr{D}(-\pi, \pi)$. □

Example 8.5.6. Essentially the same calculations as in Example 8.5.5 show that

$$\int_{-\infty}^{\infty} e^{-ix\xi}\, dx = 2\pi\delta \quad \text{in } \mathscr{D}'(\mathbf{R}),$$

where the left-hand side is interpreted as the limit in $\mathscr{D}'(\mathbf{R})$ of the integrals

$$\int_{-n}^{n} e^{-ix\xi}\, dx, \quad \text{where } \xi \in \mathbf{R} \text{ and } n = 1, 2, \dots .$$ □

Example 8.5.7. Suppose that $f_n \to f$ in $L^1_{\mathrm{loc}}(X)$, i.e.,

$$\int_K |f - f_n|\, dx \to 0 \quad \text{as } n \to \infty$$

for every compact subset K to X; this holds for instance if f_n converges locally uniformly to f on X. We will show that $u_{f_n} \to u_f$ in $\mathscr{D}'(X)$ under this assumption. Suppose that $\phi \in \mathscr{D}(X)$ with compact support $K \subset X$. Then

$$|\langle u_f, \phi\rangle - \langle u_{f_n}, \phi\rangle| = \left| \int_X (f - f_n)\phi\, dx \right| \leqslant \|\phi\|_\infty \int_K |f - f_n|\, dx \longrightarrow 0.$$ □

We end this section by stating without a proof a theorem which shows that the space $\mathscr{D}'(X)$ is complete.

Definition 8.5.8. A sequence $(u_n)_{n=1}^{\infty} \subset \mathscr{D}'(X)$ is a **Cauchy sequence** in $\mathscr{D}'(X)$ if $\langle u_n, \phi\rangle$, $n = 1, 2, \dots$, is a Cauchy sequence in $\mathbf{C}$ for every $\phi \in \mathscr{D}(X)$.

Theorem 8.5.9. *Every Cauchy sequence in $\mathscr{D}'(X)$ is convergent.*

8.6. Restriction and Support

Definition 8.6.1. The **restriction** $u|_{X'}$ of a distribution $u \in \mathscr{D}'(X)$ to an open subset X' of X is defined by

$$\langle u|_{X'}, \phi\rangle = \langle u, \phi\rangle, \quad \phi \in \mathscr{D}(X').$$

Notice that $u|_{X'} \in \mathscr{D}'(X')$. The support of a distribution is defined as for functions (see Definition 1.2.3).

Definition 8.6.2. The **support** $\operatorname{supp} u$ of a distribution $u \in \mathscr{D}'(X)$ consists of those points $x \in X$ for which $u|_{X'} \neq 0$ for every neighbourhood X' of x.

If $x \notin \operatorname{supp} u$, then there exists a neighbourhood X' of x such that $u|_{X'} = 0$. Since this implies that the complement of $\operatorname{supp} u$ is open, we see that $\operatorname{supp} u$ is closed.

Example 8.6.3. Let us show that the support of δ_a is $\{a\}$. If $\phi \in \mathscr{D}(\mathbf{R}^d \smallsetminus \{a\})$, then $\langle \delta_a, \phi \rangle = \phi(a) = 0$, which shows that $\operatorname{supp} \delta_a \subset \{a\}$. Conversely, if $\phi \in \mathscr{D}(\mathbf{R}^d)$ and $\phi(a) \neq 0$, then $\langle \delta_a, \phi \rangle \neq 0$, which shows that $\{a\} \subset \operatorname{supp} \delta_a$. □

The following proposition shows that the support of a regular distribution, generated by a locally integrable function, coincides with the support of the function.

Proposition 8.6.4. *Suppose that* $f \in L^1_{\text{loc}}(X)$. *Then* $\operatorname{supp} u_f = \operatorname{supp} f$.

Proof. Suppose first that $x \notin \operatorname{supp} u_f$. Then there exists a neighbourhood X' of x such that $u_f|_{X'} = 0$. Let $K \subset X'$ be compact and choose a function $\psi \in \mathscr{D}(X)$ such that $\psi = 1$ on K. Now, if ϕ is a mollifier on $\mathbf{R}^d$, then

$$\phi_\varepsilon * (\psi f)(x') = \int_X \phi_\varepsilon(x' - y)\psi(y)f(y)\,dy = 0$$

for $x' \in X'$ if ε is small enough. As in the proof of Proposition 8.3.2, it follows that $f(x') = 0$ for a.e. $x' \in K$ and consequently for a.e. $x' \in X'$. This shows that $x \notin \operatorname{supp} f$.

Conversely, suppose that $x \notin \operatorname{supp} f$. Then there exists a neighbourhood X' of x such that $f = 0$ a.e. on X'. This implies that $\langle u_f, \phi \rangle = 0$ for every $\phi \in \mathscr{D}(X')$, i.e., $x \notin \operatorname{supp} u_f$. ■

Proposition 8.6.5. *Suppose that* $u \in \mathscr{D}'(X)$ *and* $\phi \in \mathscr{D}(X)$ *and that*

$$\operatorname{supp} u \cap \operatorname{supp} \phi = \varnothing.$$

Then $\langle u, \phi \rangle = 0$.

Proof. Denote the support of ϕ by K. Then, for every $x \in K$, there exists a neighbourhood $X' \subset X$ of x such that $u|_{X'} = 0$. Since K is compact, it follows that K can be covered by a finite number such neighborhoods $X'_1, \dots X'_m$. Now let $\phi_1, \dots, \phi_m$ be a partition of unity subordinate to this of K (see Proposition 1.7.3). Then

$$\langle u, \phi \rangle = \sum_{j=1}^{m} \langle u, \phi_j \phi \rangle = 0$$

since $\operatorname{supp}(\phi_j \phi) \subset X_j$ and $u|_{X'_j} = 0$ for every j. ■

Chapter 9

Basic Operations on Distributions

In what follows, X denotes an open subset to $\mathbf{R}^d$.

9.1. Vector Space Operations

As already noticed, $\mathscr{D}'(X)$ is a vector space over the complex numbers with the vector space operations defined pointwise: If $u, v \in \mathscr{D}'(X)$ and $\alpha, \beta \in \mathbf{C}$, one defines $\alpha u + \beta v$ through

$$\langle \alpha u + \beta v, \phi \rangle = \alpha \langle u, \phi \rangle + \beta \langle v, \phi \rangle, \quad \phi \in \mathscr{D}(X).$$

It is easily verified that $\alpha u + \beta v \in \mathscr{D}'(X)$, i.e., $\alpha u + \beta v$ is linear and sequentially continuous.

9.2. Multiplication with C^∞-functions

We next define multiplication of distributions with C^∞-functions — first some notation.

Definition 9.2.1.

(a) We denote by $\mathscr{E}(X)$ the class of infinitely continuously differentiable functions on X.

(b) A sequence $(\phi_n)_{n=1}^\infty$ in $\mathscr{E}(X)$ **converges** to a function $\phi \in \mathscr{E}(X)$ if $\partial^\alpha \phi_n$ converges uniformly to $\partial^\alpha \phi$ on every compact subset K to X for every multi-index α.

Suppose that $u \in L^1_{\text{loc}}(X)$ and $f \in \mathscr{E}(X)$. Then, since $fu \in L^1_{\text{loc}}(X)$, the product fu defines a regular distribution on X (here denoted fu) which acts on $\mathscr{D}(X)$ through integration:

$$\langle fu, \phi \rangle = \int_X (fu)\phi \, dx = \int_X u(f\phi) \, dx = \langle u, f\phi \rangle \quad \text{for } \phi \in \mathscr{D}(X).$$

Here, we used the fact that $f\phi \in \mathscr{D}(X)$. This shows that if $u \in \mathscr{D}'(X)$, the product of u with $f \in \mathscr{E}(X)$ has to be defined in the following manner.

Definition 9.2.2. Suppose that $u \in \mathscr{D}'(X)$ and $f \in \mathscr{E}(X)$. Then the **product** fu is defined by

$$\langle fu, \phi \rangle = \langle u, f\phi \rangle, \quad \phi \in \mathscr{D}(X).$$

Remark 9.2.3.

(a) It is easy to see that fu is linear and sequentially continuous, so that fu belongs to $\mathscr{D}'(X)$. This is a consequence of the fact that if $\phi_n \to \phi$ in $\mathscr{D}(X)$, then $f\phi_n \to f\phi$ in $\mathscr{D}(X)$ which is not hard to verify.

(b) Multiplication with a function $f \in C^\infty(X)$ is a continuous operation on $\mathscr{D}'(X)$ in the sense that $u_n \to u$ in $\mathscr{D}'(X)$ implies that $fu_n \to fu$ in $\mathscr{D}'(X)$.

Example 9.2.4. If $f \in C^\infty(\mathbf{R}^d)$, then

$$\langle f\delta, \phi\rangle = \langle \delta, f\phi\rangle = f(0)\phi(0) = \langle f(0)\delta, \phi\rangle$$

for every test function $\phi \in \mathscr{D}(\mathbf{R}^d)$, which shows that $f\delta = f(0)\delta$. □

Example 9.2.5. Let us show that $x \operatorname{pv} \frac{1}{x} = 1$ in $\mathscr{D}'(\mathbf{R})$. This holds since

$$\left\langle x \operatorname{pv} \frac{1}{x}, \phi(x) \right\rangle = \lim_{\varepsilon \to 0} \int_{|x| \geqslant \varepsilon} \frac{x\phi(x)}{x}\, dx = \int_{-\infty}^{\infty} \phi(x)\, dx = \langle 1, \phi\rangle$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. □

In general, is impossible to define the product of two distributions in a meaningful way. Let us illustrate this with an example.

Example 9.2.6. Suppose that we could define a product on $\mathscr{D}'(\mathbf{R})$ which were both commutative and associative. Due to commutativity, we would then have

$$x\Big(\delta\Big(\operatorname{pv}\frac{1}{x}\Big)\Big) = x\Big(\Big(\operatorname{pv}\frac{1}{x}\Big)\delta\Big).$$

But, since the product is assumed to be associative,

$$x\Big(\delta\Big(\operatorname{pv}\frac{1}{x}\Big)\Big) = (x\delta)\operatorname{pv}\frac{1}{x} = 0 \operatorname{pv}\frac{1}{x} = 0,$$

while

$$x\Big(\Big(\operatorname{pv}\frac{1}{x}\Big)\delta\Big) = \Big(x\operatorname{pv}\frac{1}{x}\Big)\delta = 1\delta = \delta.$$ □

Proposition 9.2.7. *Suppose that $f \in \mathscr{E}(X)$ and $u \in \mathscr{D}'(X)$. Then*

$$\operatorname{supp}(fu) \subset \operatorname{supp} f \cap \operatorname{supp} u.$$

Proof. Suppose first that $x \notin \operatorname{supp} f$. Then there exists a neighbourhood $X' \subset X$ of x such that $f = 0$ on X', which implies that

$$\langle fu, \phi\rangle = \langle u, f\phi\rangle = 0$$

for every $\phi \in \mathscr{D}(X)$ with support in X' since $f\phi = 0$, and hence that $x \notin \operatorname{supp}(fu)$. Next suppose that $x \notin \operatorname{supp} u$. Then $u|_{X'} = 0$ in a neighbourhood $X' \subset X$ of x, which implies that $\langle fu, \phi\rangle = 0$ for every $\phi \in \mathscr{D}(X)$ with support in X' since $\operatorname{supp}(f\phi) \subset X'$. It follows that $x \notin \operatorname{supp}(fu)$. ■

9.3. Affine Transformations

Suppose that $u \in L^1_{\text{loc}}(\mathbf{R}^d)$ and let $h \in \mathbf{R}^d$. Then

$$\langle \tau_h u, \phi\rangle = \int_{\mathbf{R}^d} u(x-h)\phi(x)\, dx = \int_{\mathbf{R}^d} u(x)\phi(x+h)\, dx = \langle u, \tau_{-h}\phi\rangle$$

for every $\phi \in \mathscr{D}(\mathbf{R}^d)$. This identity motivates the following definition.

Definition 9.3.1. If $u \in \mathscr{D}'(\mathbf{R}^d)$ and $h \in \mathbf{R}^d$, then the **translate** $\tau_h u$ is defined by

$$\langle \tau_h u, \phi \rangle = \langle u, \tau_{-h}\phi \rangle, \quad \phi \in \mathscr{D}(\mathbf{R}^d).$$

Remark 9.3.2.

(a) One easily verifies that $\tau_h u \in \mathscr{D}'(\mathbf{R}^d)$ by showing that $\tau_h u$ is linear and sequentially continuous. The second property follows from the fact that if $\phi_n \to \phi$ in $\mathscr{D}(X)$, then $\tau_h \phi_n \to \tau_h \phi$ in $\mathscr{D}(X)$.

(b) One can also show that translation is a continuous operation on $\mathscr{D}'(\mathbf{R}^d)$: If $u_n \to u$ in $\mathscr{D}'(X)$, then $\tau_h u_n \to \tau_h u$ in $\mathscr{D}'(X)$.

The next example illustrates how translation shifts the support of a distribution.

Example 9.3.3. If $h \in \mathbf{R}^d$, then

$$\langle \tau_h \delta, \phi \rangle = \langle \delta, \tau_{-h}\phi \rangle = \phi(h) = \langle \delta_h, \phi \rangle$$

for every $\phi \in \mathscr{D}(\mathbf{R}^d)$, which shows that $\tau_h \delta = \delta_h$. □

Now suppose that $u \in L^1_{\text{loc}}(\mathbf{R}^d)$ and that A is a non-singular $d \times d$ matrix. Recall that we have used the notation

$$A^* u(x) = u(Ax), \quad x \in \mathbf{R}^d.$$

Then, changing variables $y = Ax$, we have that

$$\begin{aligned} \langle A^* u, \phi \rangle &= \int_{\mathbf{R}^d} u(Ax)\phi(x)\,dx = |\det A|^{-1} \int_{\mathbf{R}^d} u(y)\phi(A^{-1}y)\,dy \\ &= |\det A|^{-1} \langle u, (A^{-1})^* \phi \rangle \end{aligned}$$

for every $\phi \in \mathscr{D}(\mathbf{R}^d)$. We therefore make the following definition.

Definition 9.3.4. If $u \in \mathscr{D}'(\mathbf{R}^d)$ and A is a non-singuar $d \times d$ matrix, then the functional $A^* u$ is defined by

$$\langle A^* u, \phi \rangle = |\det A|^{-1} \langle u, (A^{-1})^* \phi \rangle, \quad \phi \in \mathscr{D}(\mathbf{R}^d).$$

Remark 9.3.5. It is easy to show that $A^* u$ belongs to $\mathscr{D}'(\mathbf{R}^d)$ and that the map $u \mapsto A^* u$ is a continuous operation on $\mathscr{D}'(\mathbf{R}^d)$.

Some special cases are worth mentioning. The matrix $A = -I$ corresponds to the **reflection** operator R, defined by

$$\langle Ru, \phi \rangle = \langle u, R\phi \rangle, \quad \phi \in \mathscr{D}(\mathbf{R}^d),$$

for $u \in \mathscr{D}'(\mathbf{R}^d)$.

Definition 9.3.6. A distribution $u \in \mathscr{D}'(\mathbf{R}^d)$ is called **even** if $Ru = u$ and **odd** if $Ru = -u$.

Example 9.3.7. If $\phi \in \mathscr{D}(\mathbf{R}^d)$, then

$$\langle R\delta, \phi \rangle = \langle \delta, R\phi \rangle = \phi(0) = \langle \delta, \phi \rangle,$$

which shows that δ is even. □

Example 9.3.8. The Cauchy principal value is odd since

$$\left\langle R\,\mathrm{pv}\,\frac{1}{x}, \phi(x)\right\rangle = \left\langle \mathrm{pv}\,\frac{1}{x}, R\phi(x)\right\rangle = \lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon}\frac{\phi(-x)}{x}\,dx = -\lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon}\frac{\phi(x)}{x}\,dx$$
$$= \left\langle -\,\mathrm{pv}\,\frac{1}{x}, \phi(x)\right\rangle$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. □

The matrix $A = tI$, where $t \neq 0$, gives the **dilation** operator D_t, defined by

$$\langle D_t u, \phi\rangle = \langle u, |t|^{-d} D_{t^{-1}}\phi\rangle, \quad \phi \in \mathscr{D}(\mathbf{R}^d),$$

for $u \in \mathscr{D}'(\mathbf{R}^d)$.

Definition 9.3.9. A distribution $u \in \mathscr{D}'(\mathbf{R}^d)$ is said to be **homogeneous** of degree $\lambda \in \mathbf{C}$ if

$$D_t u = t^\lambda u \quad \text{for } t > 0.$$

For a function $u \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$, this means that

$$u(tx) = t^\lambda u(x) \quad \text{for every } x \in \mathbf{R}^d.$$

Example 9.3.10. The Dirac δ is homogeneous of degree $-d$:

$$\langle D_t\delta, \phi\rangle = \langle \delta, t^{-d} D_{t^{-1}}\phi\rangle = t^{-d}\phi(0) = \langle t^{-d}\delta, \phi\rangle$$

for $t > 0$ and $\phi \in \mathscr{D}(\mathbf{R}^d)$. □

Example 9.3.11. The Cauchy principal value is homogeneous of degree -1 (which of course is no big surprise since x^{-1} is homogeneous of degree -1):

$$\left\langle D_t\,\mathrm{pv}\,\frac{1}{x}, \phi(x)\right\rangle = \left\langle \mathrm{pv}\,\frac{1}{x}, t^{-1}D_{t^{-1}}\phi(x)\right\rangle = t^{-1}\lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon}\frac{\phi(t^{-1}x)}{x}\,dx$$
$$= t^{-1}\lim_{\varepsilon\to 0}\int_{|y|\geqslant t\varepsilon}\frac{\phi(y)}{y}\,dy = \left\langle t^{-1}\,\mathrm{pv}\,\frac{1}{x}, \phi(x)\right\rangle$$

for $t > 0$ and $\phi \in \mathscr{D}(\mathbf{R})$. □

Chapter 10
Differentiation of distributions

As before, X wil demote an open subset to $\mathbf{R}^d$.

10.1. Definition

To motivate the definition of derivatives of distributions, assume that $u \in C^1(X)$ and let $\phi \in \mathscr{D}(X)$. Using integration by parts, we see that

$$\langle \partial_j u, \phi \rangle = \int_X (\partial_j u)\phi \, dx = -\int_X u(\partial_j \phi) \, dx = -\langle u, \partial_j \phi \rangle$$

for $j = 1, 2, \ldots, d$. The first order partial derivatives of a distribution on X thus have to be defined in the following way.

Definition 10.1.1. The **partial derivative** $\partial_j u$, where $j = 1, 2, \ldots, d$, of a distribution $u \in \mathscr{D}'(X)$ is defined by

$$\langle \partial_j u, \phi \rangle = -\langle u, \partial_j \phi \rangle, \quad \phi \in \mathscr{D}(X). \tag{10.1}$$

Remark 10.1.2.

(a) It follows directly from the definition of convergence in $\mathscr{D}(X)$ that if $u \in \mathscr{D}'(X)$, then $\partial_j u \in \mathscr{D}'(X)$ for $j = 1, 2, \ldots, d$.

(b) It follows from the previous remark that a distribution $u \in \mathscr{D}'(X)$ has derivatives of every order. Also, if α is a multi-index, then (10.1) implies that

$$\langle \partial^\alpha u, \phi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \phi \rangle \quad \text{for } \phi \in \mathscr{D}(X).$$

(c) Notice that $\operatorname{supp} \partial^\alpha u \subset \operatorname{supp} u$ for every multi-index α.

(d) For a regular distribution u_f, the derivatives of u_f are often called **weak derivatives**.

(e) In the one-dimensional case, the derivates of a distribution u will be denoted u', u'' etc.

10.2. Examples of Derivatives

Example 10.2.1. The derivative $\partial^\alpha \delta_a$ acts on $\mathscr{D}(\mathbf{R}^d)$ in the following way:

$$\langle \partial^\alpha \delta_a, \phi \rangle = (-1)^{|\alpha|} \langle \delta_a, \partial^\alpha \phi \rangle = (-1)^{|\alpha|} \partial^\alpha \phi(a), \quad \phi \in \mathscr{D}(\mathbf{R}^d). \qquad \square$$

The next example shows that the weak derivative of a absolutely continuous function on $\mathbf{R}$ coincides with the ordinary derivative.

Example 10.2.2. Suppose that $f \in AC(\mathbf{R})$. Integrating by parts, we the see that

$$\langle u_f', \phi \rangle = -\langle u_f, \phi' \rangle = -\int_{-\infty}^{\infty} f(x)\phi'(x) \, dx = \int_{-\infty}^{\infty} f'(x)(x)\phi \, dx = \langle u_{f'}, \phi \rangle.$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. This shows that $u_f' = u_{f'}$. $\square$

In the previous example, the function is continuous and differentiable a.e. The following example illustrates what could happen if we drop the continuity assumption.

Example 10.2.3. Let us determine the first derivative of the Heaviside function H:

$$\langle u_H', \phi\rangle = -\langle u_H, \phi'\rangle = -\int_0^\infty \phi'(x)\,dx = \phi(0) = \langle \delta, \phi\rangle$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. This shows that $u_H' = \delta$. □

The next example generalizes Example 10.2.3.

Example 10.2.4. Suppose that $f \in C^1(\mathbf{R} \setminus \{a\})$ has a jump discontinuity at a and that $f' \in L^1_{\mathrm{loc}}(\mathbf{R})$. Then

$$\begin{aligned}\langle u_f', \phi\rangle &= -\langle u_f, \phi'\rangle = -\int_{-\infty}^\infty f(x)\phi'(x)\,dx \\ &= -\int_a^\infty f(x)\phi'(x)\,dx - \int_{-\infty}^a f(x)\phi'(x)\,dx \\ &= (f(a^+) - f(a^-))\phi(a) + \int_{-\infty}^\infty f'(x)\phi(x)\,dx \\ &= \langle (f(a^+) - f(a^-))\delta_a + u_{f'}, \phi\rangle\end{aligned}$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. This shows that $u_f' = (f(a^+) - f(a^-))\delta_a + u_{f'}$. □

The following example illustrates the fact that if the derivative of a function is not locally integrable, then the weak derivative cannot coincide with the ordinary derivative.

Example 10.2.5. Let $f(x) = \ln|x|,\ x \in \mathbf{R}$. Then

$$\begin{aligned}\langle u_f', \phi\rangle &= -\langle u_f, \phi'\rangle = -\int_{-\infty}^\infty \ln|x|\,\phi'(x)\,dx = -\lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon} \ln|x|\,\phi'(x)\,dx \\ &= \lim_{\varepsilon\to 0}\left((\phi(\varepsilon) - \phi(-\varepsilon))\ln\varepsilon + \int_{|x|\geqslant\varepsilon}\frac{\phi(x)}{x}\,dx\right) \\ &= \lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon}\frac{\phi(x)}{x}\,dx = \left\langle \mathrm{pv}\,\frac{1}{x}, \phi(x)\right\rangle\end{aligned}$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. This shows that $u_f' = \mathrm{pv}\,\frac{1}{x}$. □

Example 10.2.6. Consider the function f on $\mathbf{R}$, defined by

$$f(x) = \begin{cases} x^{-1/2} & \text{if } x > 0 \\ 0 & \text{if } x \leqslant 0 \end{cases}.$$

The weak derivative of f is calculated in the following way:

$$\begin{aligned}\langle u_f', \phi\rangle &= -\langle u_f, \phi'\rangle = -\int_0^\infty \frac{\phi'(x)}{x^{1/2}}\,dx = -\lim_{\varepsilon\to 0}\int_\varepsilon^\infty \frac{\phi'(x)}{x^{1/2}}\,dx \\ &= -\frac{1}{2}\lim_{\varepsilon\to 0}\left(\int_\varepsilon^\infty \frac{\phi(x)}{x^{3/2}}\,dx - 2\frac{\phi(\varepsilon)}{\varepsilon^{1/2}}\right) \\ &= -\frac{1}{2}\lim_{\varepsilon\to 0}\left(\int_\varepsilon^\infty \frac{\phi(x)-\phi(0)}{x^{3/2}}\,dx - 2\frac{\phi(\varepsilon)-\phi(0)}{\varepsilon^{1/2}}\right) = -\frac{1}{2}\int_0^\infty \frac{\phi(x)-\phi(0)}{x^{3/2}}\,dx,\end{aligned}$$

for every $\phi \in \mathscr{D}(\mathbf{R})$. If we now define the **finite part** fp $\frac{1}{x^{3/2}}$ of $x^{-3/2}$ by

$$\left\langle \mathrm{fp}\,\frac{1}{x^{3/2}}, \phi(x) \right\rangle = \int_0^\infty \frac{\phi(x)-\phi(0)}{x^{3/2}}\,dx, \quad \phi \in \mathscr{D}(\mathbf{R}),$$

we have $u_f' = -\frac{1}{2}\,\mathrm{fp}\,\frac{1}{x^{3/2}}$. □

10.3. Differentiation Rules

Basically all differentiation rules from calculus hold in $\mathscr{D}'(X)$. Differentiation on $\mathscr{D}'(X)$ is for instance a linear operation. This follows from the definition of the derivative and the way addition and multiplication with scalars is defined in $\mathscr{D}'(X)$.

Proposition 10.3.1. *Suppose that $u, v \in \mathscr{D}(X)$. Then, for every multi-index α,*

$$\partial^\alpha(au + bv) = a(\partial^\alpha u) + b(\partial^\alpha v) \quad \textit{for all } a, b \in \mathbf{C}.$$

Leibniz' rule for differentiating products also holds in $\mathscr{D}'(X)$.

Proposition 10.3.2. *Suppose that $f \in \mathscr{E}(X)$ and $u \in \mathscr{D}'(X)$. Then, for every multi-index α,*

$$\partial^\alpha(fu) = \sum_{0\leqslant\beta\leqslant\alpha} \binom{\alpha}{\beta} \partial^\beta f\, \partial^{\alpha-\beta} u. \tag{10.2}$$

Proof. If $\alpha = 0$, there is nothing to prove. Suppose that $|\alpha| = 1$, so that $\partial^\alpha = \partial_j$ for some j. Then

$$\begin{aligned}\langle \partial_j(fu), \phi\rangle &= -\langle u, f(\partial_j\phi)\rangle = -\langle u, \partial_j(f\phi) - (\partial_j f)\phi\rangle = \langle f(\partial_j u), \phi\rangle + \langle (\partial_j f)u, \phi\rangle \\ &= \langle (\partial_j f)u + f(\partial_j u), \phi\rangle\end{aligned}$$

for every $\phi \in \mathscr{D}(X)$, which shows that $\partial_j(fu) = (\partial_j f)u + f(\partial_j u)$. Using induction, it follows that there exist constants C_β^α such that

$$\partial^\alpha(fu) = \sum_{0\leqslant\beta\leqslant\alpha} C_\beta^\alpha\, \partial^\beta f\, \partial^{\alpha-\beta} u.$$

If we now apply this identity to $f(x) = e^{\xi\cdot x}$, $x \in \mathbf{R}^d$, and $u(x) = e^{\eta\cdot x}$, $x \in \mathbf{R}^d$, where $\xi \in \mathbf{R}^d$ and $\eta \in \mathbf{R}^d$ are parameters, we obtain that

$$(\xi+\eta)^\alpha e^{(\xi+\eta)\cdot x} = \Bigg(\sum_{0\leqslant\beta\leqslant\alpha} C_\beta^\alpha\, \xi^\beta \eta^{\alpha-\beta}\Bigg) e^{(\xi+\eta)\cdot x}.$$

After canceling the common factors, this shows that C_β^α are the coefficients in the binomial expansion of $(\xi+\eta)^\alpha$. This proves (10.2). ■

As for smooth functions, partial derivatives of distributions commute.

Proposition 10.3.3. *Suppose that $u \in \mathscr{D}'(X)$. Then*

$$\partial^\alpha(\partial^\beta u) = \partial^\beta(\partial^\alpha u)$$

for all multi-indices α and β.

Proof. If $\phi \in \mathscr{D}(X)$, then

$$\begin{aligned}\langle \partial^\alpha(\partial^\beta u), \phi\rangle &= (-1)^{|\alpha|+|\beta|}\langle u, \partial^\beta(\partial^\alpha \phi)\rangle = (-1)^{|\alpha|+|\beta|}\langle u, \partial^\alpha(\partial^\beta \phi)\rangle \\ &= \langle \partial^\beta(\partial^\alpha u), \phi\rangle.\end{aligned}$$ ■

The operator $\partial^\alpha : \mathscr{D}'(X) \to \mathscr{D}'(X)$ is sequentially continuous for every multi-index α:

Proposition 10.3.4. *Suppose that $u_n \to u$ in $\mathscr{D}'(X)$. Then*

$$\partial^\alpha u_n \longrightarrow \partial^\alpha u \quad \text{in } \mathscr{D}'(X)$$

for every multi-indicex α.

Proof. If $\phi \in \mathscr{D}(X)$, then

$$\langle \partial^\alpha u_n, \phi\rangle = (-1)^{|\alpha|}\langle u_n, \partial^\alpha \phi\rangle \longrightarrow (-1)^{|\alpha|}\langle u, \partial^\alpha \phi\rangle = \langle \partial^\alpha u, \phi\rangle \quad \text{as } n \to \infty.$$ ■

It follows from this proposition that every convergent series in $\mathscr{D}'(X)$ can be differentiated termwise:

$$\partial^\alpha\left(\sum_{n=1}^\infty u_n\right) = \sum_{n=1}^\infty \partial^\alpha u_n.$$

Example 10.3.5. Notice that

$$\lfloor x\rfloor = \sum_{n=1}^\infty H(x-n) \quad \text{for } x \geqslant 0.$$

Since the series contains a finite number of terms for x belonging to a bounded interval, it converges in $L^1_{\text{loc}}(0,\infty)$ and hence in $\mathscr{D}'(0,\infty)$ (see Example 8.5.7). Proposition 10.3.4 and Example 10.2.3 now show that

$$\lfloor x\rfloor' = \sum_{n=1}^\infty (H(x-n))' = \sum_{n=1}^\infty \delta_n.$$ □

10.4. Antiderivatives

Let $I \subset \mathbf{R}$ be an open interval. A test function $\phi \in \mathscr{D}(I)$ has an antiderivative belonging to $\mathscr{D}(I)$ if and only if $\int_I \phi\, dx = 0$. Indeed, all antiderivatives Φ of ϕ are given by

$$\Phi(x) = \int_{-\infty}^x \phi(y)\, dy + C, \quad x \in I,$$

where $C \in \mathbf{C}$ is a constant. Every antiderivative Φ is of course smooth, but Φ has compact support in I if and only if $C = 0$ and $\int_I \phi\, dx = 0$. Below, we will show that for every distribution $u \in \mathscr{D}'(I)$, there exists a distribution $U \in \mathscr{D}'(I)$ which is an **antiderivative** of u in the sense that $U' = u$ and that the antiderivatives of u are uniquely determined up to an additive constant.

Proposition 10.4.1. *Suppose that $I \subset \mathbf{R}$ is an open interval and that $u \in \mathscr{D}'(I)$. Then u has a antiderivative $U \in \mathscr{D}'(I)$. Moreover, every antiderivative $V \in \mathscr{D}'(I)$ of u is given by $V = U + C$ for some constant $C \in \mathbf{C}$.*

Proof. We will use the operator $T : \mathscr{D}(I) \to \mathscr{D}(I)$, defined for $\phi \in \mathscr{D}(I)$ by

$$T\phi(x) = \int_{-\infty}^{x} \phi(y)\, dy - \langle 1, \phi\rangle \int_{-\infty}^{x} \psi(y)\, dy, \quad x \in I,$$

where $\psi \in \mathscr{D}(I)$ satisfies $\int_I \psi\, dx = 1$. We leave it to the reader to establish the following properties of T:

(i) $T\phi \in \mathscr{D}(I)$;

(ii) $T\phi' = \phi$;

(iii) if $\phi_n \to \phi$ in $\mathscr{D}(I)$, then $T\phi_n \to T\phi$ in $\mathscr{D}(I)$.

Define $U \in \mathscr{D}'(I)$ by

$$\langle U, \phi\rangle = -\langle u, T\phi\rangle, \quad \phi \in \mathscr{D}(I).$$

Then $U' = u$ since

$$\langle U', \phi\rangle = -\langle U, \phi'\rangle = \langle u, T\phi'\rangle = \langle u, \phi\rangle$$

for every $\phi \in \mathscr{D}(I)$. Suppose that $V \in \mathscr{D}'(I)$ satisfies $V' = u$ and put $W = V - U$. Then $W' = 0$, so

$$0 = \langle W', T\phi\rangle = -\langle W, (T\phi)'\rangle = -\langle W, \phi - \langle 1, \phi\rangle\psi\rangle = \langle\langle W, \psi\rangle, \phi\rangle - \langle W, \phi\rangle$$

for every $\phi \in \mathscr{D}(I)$, which shows that $W = \langle W, \psi\rangle$, i.e., $V = U + \langle W, \psi\rangle$. ∎

Example 10.4.2. Let us calculate the antiderivatives of pv $\frac{1}{x}$ this using the technique employed in the proof of Proposition 10.4.1. We know that one antiderivative U is given by

$$\langle U, \phi\rangle = -\left\langle \operatorname{pv} \frac{1}{x}, T\phi(x)\right\rangle = -\lim_{\varepsilon\to 0} \int_{|x|\geqslant\varepsilon} \frac{\Phi(x) - \langle 1, \phi\rangle\Psi(x)}{x}\, dx$$

for $\phi \in \mathscr{D}(\mathbf{R})$, where $\Phi(x) = \int_{-\infty}^{x} \phi(y)\, dy,\ x \in \mathbf{R}$, and $\Psi(x) = \int_{-\infty}^{x} \psi(y)\, dy,\ x \in \mathbf{R}$. Integrating by parts, we see that

$$\langle U, \phi\rangle = \langle \ln|x|, \phi(x)\rangle - \langle 1, \phi\rangle\langle \ln|x|, \psi(x)\rangle = \langle \ln|x| - C, \phi(x)\rangle,$$

where $C = \langle \ln|x|, \psi(x)\rangle$. Thus, all antiderivatives of pv $\frac{1}{x}$ are given by $\ln|x| + D$, where D is an arbitrary constant. This also follows from Example 10.2.5. □

The uniqueness part of Proposition 10.4.1 gives the following corollary.

Corollary 10.4.3. *Suppose that $I \subset \mathbf{R}$ is an open interval and that $u \in \mathscr{D}'(I)$ satisfies $u' = 0$. Then u is a constant.*

10.5. Linear Differential Operators

Suppose that $a_\alpha \in \mathscr{E}(X)$ for $|\alpha| \leqslant m$ and that not all a_α with $|\alpha| = m$ are identically 0. Put

$$P(\partial) = \sum_{|\alpha| \leqslant m} a_\alpha \partial^\alpha.$$

We call $P(\partial)$ a **linear differential operator** on $\mathscr{D}'(X)$ of **order** m. An equation of the form

$$P(\partial)u = v,$$

where $v \in \mathscr{D}'(X)$, is called a **differential equation**. In the case $d = 1$, this is an **ordinary differential equation** and for $d > 1$ a **partial differential equation**. For $X = \mathbf{R}^d$ and $v = \delta$, the solutions to this equation are called **fundamental solutions**.

Below, we illustrate how one solves an ordinary differential equation with a distribution in the right-hand side.

Example 10.5.1. Let us determine all solutions $u \in \mathscr{D}'(\mathbf{R})$ to the differential equation

$$u' + 2u = \delta. \tag{10.3}$$

Multiplying the equation with the integrating factor e^{2x}, we obtain that

$$e^{2x}u' + 2e^{2x}u = e^{2x}\delta = \delta, \quad \text{so that} \quad (e^{2x}u)' = \delta.$$

One solution to this equation is $u = H(x)e^{-2x}$. To find all solutions to the equation, we solve the corresponding homogeneous equation, namely $(e^{2x}u)' = 0$, and find that $u = Ce^{-2x}$, where C is a constant. This shows that all solutions to (10.3) are given by

$$u = Ce^{-2x} + H(x)e^{-2x}$$ □

Chapter 11

Distributions with Compact Support

11.1. Distributions on $\mathscr{E}(X)$

Definition 11.1.1. A sequence $(\phi_n)_{n=1}^{\infty} \subset \mathscr{E}(X)$ **converges** to $\phi \in \mathscr{E}(X)$ if, for every multi-index α and every compact subset K to X, $\partial^\alpha \phi_n$ converges uniformly to $\partial^\alpha \phi$ on K. We denote this by writing $\phi_n \to \phi$.

Definition 11.1.2. A **distribution** on $\mathscr{E}(X)$ is a sequentially continuous, linear functional on $\mathscr{E}(X)$. We denote the class of distributions on $\mathscr{E}(X)$ by $\mathscr{E}'(X)$.

As for distributions on $\mathscr{E}(X)$, we shall write $\langle u, \phi \rangle$ instead of $u(\phi)$ if $u \in \mathscr{E}'(X)$ and $\phi \in \mathscr{E}(X)$.

Example 11.1.3.

(a) The Dirac δ at $a \in X$ and all its derivatives define distributions on $\mathscr{E}(X)$.

(b) Every function $f \in L^1(X)$ with compact support also defines a distribution on $\mathscr{E}(X)$. □

The following theorem is proved as Theorem 8.4.1. The proof is left to the reader.

Theorem 11.1.4. *A linear functional u on $\mathscr{E}(X)$ belongs to $\mathscr{E}'(X)$ if and only if there exist a compact set $K \subset X$, a constant $C \geqslant 0$, and an integer $m \geqslant 0$ such that*

$$|\langle u, \phi \rangle| \leqslant C \sum_{|\alpha| \leqslant m} \sup_{x \in K} |\partial^\alpha \phi(x)| \tag{11.1}$$

for every function $\phi \in \mathscr{E}(X)$.

This theorem shows that every distribution on $\mathscr{E}(X)$ has compact support.

11.2. Extension of Compactly Supported Distributions

Notice that $\mathscr{D}(X)$ is a subspace $\mathscr{E}(X)$ — not only as classes of functions, but also from a topological point of view — since convergence in $\mathscr{D}(X)$ implies convergence in $\mathscr{E}(X)$. It follows that if $u \in \mathscr{E}'(X)$, then $u|_{\mathscr{D}(X)} \in \mathscr{D}'(X)$. A distribution on $\mathscr{E}(X)$ may thus be considered as a distribution on $\mathscr{D}(X)$ with compact support. We will conversely show that every distribution on $\mathscr{D}(X)$ with compact support can be extended to $\mathscr{E}(X)$.

Theorem 11.2.1. *Suppose that $u \in \mathscr{D}'(X)$ has compact support $K \subset X$. Then there exists a unique distribution $\tilde{u} \in \mathscr{E}'(X)$ such that*

(i) $\tilde{u} = u$ on $\mathscr{D}(X)$;

(ii) $\langle \tilde{u}, \phi \rangle$ if $\phi \in \mathscr{E}(X)$ and $\operatorname{supp} \phi \cap K = \varnothing$.

Proof. According to Proposition 1.7.1, there exists a function $\chi \in \mathscr{D}(X)$ such that $\chi = 1$ on K. Define $\tilde{u}$ through

$$\langle \tilde{u}, \phi \rangle = \langle u, \chi\phi \rangle \quad \text{for } \phi \in \mathscr{E}(X)$$

and let $L = \operatorname{supp}\chi$. Theorem 8.4.1 and Leibniz' rule then shows that

$$
\begin{aligned}
|\langle \tilde{u}, \phi\rangle| = |\langle u, \chi\phi\rangle| &\leqslant C \sum_{|\alpha|\leqslant m} \sup_{x\in L} |\partial^\alpha(\chi(x)\phi(x))| \\
&\leqslant C' \sum_{|\alpha|\leqslant m} \sup_{x\in L} |\partial^\alpha \phi(x)| \qquad (11.2)
\end{aligned}
$$

for every $\phi \in \mathscr{E}(X)$. It thus follows from Theorem 11.1.4 that $\tilde{u} \in \mathscr{E}'(X)$. We also have

$$\langle \tilde{u}, \phi\rangle = \langle u, \chi\phi\rangle = \langle u, \phi\rangle + \langle u, (\chi - 1)\phi\rangle = \langle u, \phi\rangle \quad \text{for every } \phi \in \mathscr{D}(X)$$

since $\operatorname{supp} u \cap \operatorname{supp}(\chi - 1)\phi = \varnothing$ (see Proposition 8.6.5). Morover,

$$\langle \tilde{u}, \phi\rangle = \langle u, \chi\phi\rangle = 0$$

for every $\phi \in \mathscr{E}(X)$ with $\operatorname{supp}\phi \cap K = \varnothing$ since $\operatorname{supp}(\chi\phi) \subset \operatorname{supp}\phi$. To prove uniqueness, suppose that $v \in \mathscr{E}'(X)$ is another extension of u to $\mathscr{E}(X)$ that satisfies (ii). Then

$$\langle v, \phi\rangle = \langle v, \chi\phi\rangle + \langle v, (1-\chi)\phi\rangle = \langle v, \chi\phi\rangle = \langle u, \chi\phi\rangle \quad \text{for every } \phi \in \mathscr{E}(X)$$

since $\operatorname{supp}((1-\chi)\phi) \cap K = \varnothing$. This shows that $v = \tilde{u}$. ∎

Remark 11.2.2.

(a) This theorem and the preceding observations show that $\mathscr{E}'(X)$ may be identified with the subspace to $\mathscr{D}'(X)$, that consists of distributions with compact support, and we shall henceforth do that.

(b) We shall also write $\langle u, \phi\rangle$ instead of $\langle \tilde{u}, \phi\rangle$ if $u \in \mathscr{E}'(X)$ and $\phi \in \mathscr{E}(X)$.

(c) In general, it is not possible to replace the set L in (11.2) with the support of u. However, if $\operatorname{supp} u$ has a smooth boundary, this can be done.

It follows directly from (11.2) that a distribution with compact support is of finite order.

Corollary 11.2.3. *Suppose that $u \in \mathscr{E}'(X)$. Then u is of finite order.*

11.3. Distributions Supported at a Point

Theorem 11.3.1. *Suppose that $u \in \mathscr{D}'(X)$ and that* $\operatorname{supp} u = \{a\}$ *for some $a \in X$. Then there exist an integer $m \geqslant 0$ and constants C_α, where $|\alpha| \leqslant m$, such that*

$$u = \sum_{|\alpha|\leqslant m} C_\alpha \partial^\alpha \delta.$$

Proof. Without loss of generality, we may assume that $a = 0$. Let $\varepsilon > 0$ be so small that $B_{2\varepsilon}(0) \subset X$ and take a function $\chi \in \mathscr{D}(X)$, with support in $B_{2\varepsilon}(0)$, such that $\chi = 1$ on $B_\varepsilon(0)$. Put $\chi_j(x) = \chi(2^j x)$, $x \in X$, $j = 0, 1, \ldots$. If m is the order of u, the Taylor expansion of a function $\phi \in \mathscr{D}(X)$ of order m around 0 is

$$\phi(x) = \sum_{|\alpha|\leqslant m} \frac{\partial^\alpha\phi(0)}{\alpha!} x^\alpha + r_m(x),$$

where

$$|\partial^\gamma r_m(x)| \leqslant C|x|^{m+1-|\gamma|} \quad \text{for } |\gamma| \leqslant m.$$

Applying u to this identity, we obtain

$$\langle u, \phi\rangle = \langle u, \chi\phi\rangle = \sum_{|\alpha|\leqslant m} \frac{\partial^\alpha \phi(0)}{\alpha!}\langle u, x^\alpha \chi(x)\rangle + \langle u, \chi_j(x) r_m(x)\rangle. \tag{11.3}$$

Suppose that $|x| \leqslant 2\varepsilon 2^{-j}$ and that $|\beta| + |\gamma| \leqslant m$. Then

$$|\partial^\beta \chi_j(x) \partial^\gamma r_m(x)| \leqslant C 2^{j|\beta|} 2^{-j(+1-|\gamma|)} \leqslant C2^{-j}.$$

It thus follows from (11.2) and Lebniz' rule that $\langle u, \chi_j r_m\rangle \to 0$ as $j \to \infty$. If we now let $j \to \infty$ in (11.3), we see that

$$\langle u, \phi\rangle = \sum_{|\alpha|\leqslant m} \frac{\partial^\alpha \phi(0)}{\alpha!}\langle u, x^\alpha \chi(x)\rangle = \sum_{|\alpha|\leqslant m} \frac{(-1)^{|\alpha|}}{\alpha!}\langle u, x^\alpha \chi(x)\rangle \langle \partial^\alpha \delta, \phi\rangle,$$

which proves the theorem. ■

Chapter 12

Tensor Products and Convolutions

In this chapter, X and Y will denote open subsets to $\mathbf{R}^d$ and $\mathbf{R}^e$, respectively. Let $W = X \times Y$.

12.1. Tensor Products of Functions

Definition 12.1.1. For $f \in L^1_{\mathrm{loc}}(X)$ and $g \in L^1_{\mathrm{loc}}(Y)$, the **tensor product** $f \otimes g$ is defined by

$$f \otimes g(x,y) = f(x)g(y), \quad (x,y) \in W.$$

Notice that $f \otimes g \in L^1_{\mathrm{loc}}(W)$.

12.2. Tensor Products of Distributions

To get an idea of how the tensor product of two distributions should be defined, we as usual consider regular distributions first and look at how the tensor product of two functions act on a test function. Suppose that $f \in L^1_{\mathrm{loc}}(X)$ and $g \in L^1_{\mathrm{loc}}(Y)$ and let $\phi \in \mathscr{D}(W)$. Then

$$\begin{aligned}\langle f \otimes g, \phi\rangle &= \iint\limits_W f(x)g(y)\phi(x,y)\,dx\,dy = \int_X f(x)\left(\int_Y g(y)\phi(x,y)\,dy\right)dx \\ &= \langle f(x), \langle g(y), \phi(x,y)\rangle\rangle.\end{aligned}$$

Notice that $\psi(y) = \phi(x,y),\ y \in Y$, belongs to $\mathscr{D}(Y)$ for every fixed $x \in X$ and that the function

$$\eta(x) = \int_Y g(y)\phi(x,y)\,dy, \quad x \in X,$$

belongs to $\mathscr{D}(X)$. The tensor product of $u \in \mathscr{D}'(X)$ and $v \in \mathscr{D}'(Y)$ should thus be defined as

$$\langle u \otimes v, \phi\rangle = \langle u(x), \langle v(y), \phi(x,y)\rangle\rangle \quad \text{for } \phi \in \mathscr{D}(W). \tag{12.1}$$

Here, we allow a little abuse of notation to make the presentation less heavy and hopefully clearer. We write $v(y)$ to indicate that v acts on the second variable in ϕ and similarly for u. To show that (12.1) makes sense, we need the following lemma.

Lemma 12.2.1. *Suppose that $v \in \mathscr{D}'(Y)$ and $\phi \in \mathscr{D}(W)$. Then the function*

$$\eta(x) = \langle v(y), \phi(x,y)\rangle, \quad x \in X, \tag{12.2}$$

belongs to $\mathscr{D}(X)$ and

$$\partial_x^\alpha \langle v(y), \phi(x,y)\rangle = \langle v(y), \partial_x^\alpha \phi(x,y)\rangle \tag{12.3}$$

for every $x \in X$ and all multi-indices α. Moreover, the mapping from $\mathscr{D}(W)$ to $\mathscr{D}(X)$, defined by (12.2)*, is sequentially continuous.*

Remark 12.2.2. There holds a corresponding result for $v \in \mathscr{E}'(Y)$ and $\phi \in \mathscr{E}(W)$. More precisely, if $v \in \mathscr{E}'(Y)$ and $\phi \in \mathscr{E}(W)$, then the function, defined by (12.2), belongs to $\mathscr{E}(X)$ and (12.3) holds.

Proof (Lemma 12.2.1). For $r \geqslant 0$, put

$$X_r = \{x \in X : |x| \leqslant r\}, \quad Y_r = \{y \in Y : |y| \leqslant r\}, \quad \text{and} \quad W_r = X_r \times Y_r,$$

and choose r so large that $\operatorname{supp}\phi \subset W_r$. Then $\operatorname{supp}\eta \subset X_r$, which shows that η has compact support. As in Example 8.1.4, we see that

$$\phi(x+h, y) \longrightarrow \phi(x, y) \quad \text{as } h \to 0$$

in $\mathscr{D}(Y)$ for every fixed $x \in X$, from which it follows that

$$\eta(x+h) \longrightarrow \eta(x) \quad \text{as } h \to 0,$$

so η is continuous. If e_j is the j-th vector in the standard basis of $\mathbf{R}^d$, we also have

$$\frac{\phi(x+he_j, y) - \phi(x, y)}{h} \longrightarrow \frac{\partial}{\partial x_j}\phi(x, y) \quad \text{as } h \to 0$$

in $\mathscr{D}(Y)$ for every fixed $x \in X$. This establishes (12.3) in the case $|\alpha| = 1$; the general case follows by induction. We have thus shown that $\eta \in \mathscr{D}(X)$. Now, suppose that $\phi_j \to 0$ in $\mathscr{D}(W)$. Denote the corresponding sequence, defined by (12.2) by η_j. If r is so large that $\operatorname{supp}\phi_j \subset W_r$ for every j, then $\operatorname{supp}\eta_j \subset X_r$ and, according to (12.3) and Theorem 8.4.1,

$$\sup_{x\in X_r} |\partial_x^\alpha \eta_j(x)| \leqslant C \sup_{x\in X_r} \sum_{|\beta|\leqslant m} \sup_{y\in Y_r} |\partial_y^\beta \partial_x^\alpha \phi_j(x, y)| \longrightarrow 0 \quad \text{as } j \to \infty.$$

This shows that $\eta_j \to 0$ in $\mathscr{D}(X)$. ■

Definition 12.2.3. The **tensor product** $u \otimes v$ of $u \in \mathscr{D}(X)$ and $v \in \mathscr{D}(Y)$ is defined by

$$\langle u \otimes v, \phi\rangle = \langle u(x), \langle v(y), \phi(x, y)\rangle\rangle \quad \text{for } \phi \in \mathscr{D}(W).$$

Theorem 12.2.4. *Suppose that* $u \in \mathscr{D}'(X)$ *and* $v \in \mathscr{D}'(Y)$. *Then* $u \otimes v \in \mathscr{D}'(W)$ *and* $\operatorname{supp} u \otimes v = \operatorname{supp} u \times \operatorname{supp} v$.

Proof. Suppose that $\phi_j \to \phi$ in $\mathscr{D}(W)$. Then, with the notation in the proof of Lemma 12.2.1, $\eta_j \to \eta$ in $\mathscr{D}(X)$. It follows that

$$\langle u \otimes v, \phi_j\rangle = \langle u, \eta_j\rangle \longrightarrow \langle u, \eta\rangle = \langle u \otimes v, \phi\rangle.$$

The statement about the support of $u \otimes v$ is left as an exercise to the reader. ■

Remark 12.2.5. If $u \in \mathscr{E}(X)$ and $v \in \mathscr{E}(Y)$, then the tensor product can be extended to $\phi \in \mathscr{E}(W)$. In this case, $u \otimes v \in \mathscr{E}'(W)$.

Example 12.2.6. If $a \in X$ and $b \in Y$, then

$$\langle \delta_a \otimes \delta_b, \phi\rangle = \langle \delta_a(x), \langle \delta_b(y), \phi(x, y)\rangle\rangle = \langle \delta_a(x), \phi(x, b)\rangle = \phi(a, b) = \langle \delta_{(a,b)}, \phi\rangle$$

for every $\phi \in \mathscr{D}(W)$, which shows that $\delta_a \otimes \delta_b = \delta_{(a,b)}$. □

12.3. Properties of Tensor Products

If $u \in \mathscr{D}'(X)$ and $v \in \mathscr{D}'(Y)$, then

$$\langle u \otimes v, \phi \otimes \psi \rangle = \langle u(x), \langle v(y), \phi(x)\psi(y)\rangle\rangle = \langle u, \phi\rangle\langle v, \psi\rangle$$

for all functions $\phi \in \mathscr{D}(X)$ and $\psi \in \mathscr{D}(Y)$. On the other hand,

$$\begin{aligned}\langle u, \phi\rangle\langle v, \psi\rangle &= \langle v(y), \langle u(x), \phi(x)\rangle\psi(y)\rangle = \langle v(y), \langle u(x), \phi(x)\psi(y)\rangle\rangle \\ &= \langle v \otimes u, \phi \otimes \psi\rangle.\end{aligned}$$

This shows that the tensor product is commutative on all functions in $\mathscr{D}(W)$ of the form $\phi \otimes \psi$, where $\phi \in \mathscr{D}(X)$ and $\psi \in \mathscr{D}(Y)$. To extend this to arbitrary functions in $\mathscr{D}(W)$, we will prove the lemma below.

Lemma 12.3.1. *The class of all finite linear combinations of functions of the form $\phi \otimes \psi$, where $\phi \in \mathscr{D}(X)$ and $\psi \in \mathscr{D}(Y)$, is dense in $\mathscr{D}(W)$.*

Proof. Suppose that $\phi \in \mathscr{D}(W)$ and put $K = \operatorname{supp} \phi$. For every $x \in K$, there exists an open cube Q_x such that $x \in Q_x \subset 2Q_x \subset W$. By compactness, K can be covered by a finite number cubes $Q_1, \dots, Q_m$. Let $\psi_1, \dots, \psi_m$ be a partition of unity subordinate to this covering (see Corollary 1.7.3). Then $\phi = \sum_{j=1}^m \psi_j \phi$ and $\operatorname{supp}(\phi\psi_j) \subset Q_j$. Consider one of the functions $\psi = \phi\psi_j$. After making a translation, we may assume that $\operatorname{supp} \psi \subset (-r, r)^{d+e} \subset (-2r, 2r)^{d+e} \subset W$. Weierstrass' approximation theorem now shows that there, for every integer $k \geqslant 1$, exists a polynomial P_k such that

$$|\partial^\alpha \psi(x, y) - \partial^\alpha P_k(x, y)| \leqslant \frac{1}{k} \quad \text{for every } (x, y) \in (-2r, 2r)^{d+e}$$

and every multi-index α with $|\alpha| \leqslant k$. Let τ be a one-dimensional cut-off function such that $\tau = 1$ on $[-r, r]$ and $\tau = 0$ outside $(-2r, 2r)$, and put

$$\eta_k(x, y) = P_k(x, y)\tau(x_1) \dots \tau(y_e), \quad (x, y) \in W.$$

Then $\eta_k \in \mathscr{D}(W)$ with $\operatorname{supp} \eta_k \subset (-2r, 2r)^{d+e}$ and has the form that we are looking for. Consider the following three cases:

(i) In $[-r, r]^{d+e}$ is $\eta_k = P_k$. Moreover, $\partial^\alpha P_k$ tends uniformly to $\partial^\alpha \psi$ as $k \to \infty$ for every multi-index α.

(ii) In $(-2r, 2r)^{d+e} \smallsetminus [-r, r]^{d+e}$ is $\psi = 0$. Moreover, according to Leibniz' rule,

$$|\partial^\alpha \eta_k(x, y)| \leqslant C \sum_{\beta \leqslant \alpha} |\partial^\beta P_k(x, y)| \leqslant \frac{C}{k},$$

which shows that $\partial^\alpha \eta_k$ tends uniformly to 0.

(iii) Outside $(-2r, 2r)^{d+e}$ is $\psi = \eta_k = 0$.

This shows that $\eta_k \to \psi$ in $\mathscr{D}(W)$. ∎

Corollary 12.3.2. *Suppose that $U, V \in \mathscr{D}'(W)$ and that $\langle U, \phi \otimes \psi\rangle = \langle V, \phi \otimes \psi\rangle$ for all functions $\phi \in \mathscr{D}(X)$ and $\psi \in \mathscr{D}(Y)$. Then $U = V$.*

The next three propositions show that the tensor product is commutative, associative, and distributive.

Proposition 12.3.3. *Suppose that* $u \in \mathscr{D}'(X)$ *and* $v \in \mathscr{D}'(Y)$. *Then*

$$u \otimes v = v \otimes u.$$

Proof. This follows from Corollary 12.3.2 with $U = u \otimes v$ and $V = v \otimes u$. ∎

Proposition 12.3.4. *Suppose that* $u \in \mathscr{D}'(X)$, $v \in \mathscr{D}'(Y)$, *and* $w \in \mathscr{D}'(Z)$, *where* Z *is an open subset to* $\mathbf{R}^f$. *Then*

$$u \otimes (v \otimes w) = (u \otimes v) \otimes w.$$

Proof. If $\phi \in \mathscr{D}(X \times Y \times Z)$, then

$$\begin{aligned} \langle u \otimes (v \otimes w), \phi \rangle &= \langle u(x), \langle (v \otimes w)(y,z), \phi(x,y,z) \rangle \rangle \\ &= \langle u(x), \langle v(y), \langle w(z), \phi(x,y,z) \rangle \rangle \rangle \\ &= \langle u \otimes v(x,y), \langle w(z), \phi(x,y,z) \rangle \rangle \\ &= \langle (u \otimes v) \otimes w, \phi \rangle. \end{aligned}$$ ∎

Proposition 12.3.5. *Suppose that* $u, v \in \mathscr{D}'(X)$ *and* $w \in \mathscr{D}'(Y)$. *Then*

$$(u + v) \otimes w = u \otimes w + v \otimes w.$$

Proof. If $\phi \in \mathscr{D}(X \times Y \times Z)$, then

$$\begin{aligned} \langle (u+v) \otimes w, \phi \rangle &= \langle u(x) + v(x), \langle w(y), \phi(x,y) \rangle \rangle \\ &= \langle u(x), \langle w(y), \phi(x,y) \rangle \rangle + \langle v(x), \langle w(y), \phi(x,y) \rangle \rangle \\ &= \langle u \otimes w, \phi \rangle + \langle v \otimes w, \phi \rangle. \end{aligned}$$ ∎

Proposition 12.3.6. *Suppose that* $u_j \to u$ *in* $\mathscr{D}'(X)$. *Then* $u_j \otimes v \to u \otimes v$ *in* $\mathscr{D}'(X)$ *for every* $v \in \mathscr{D}'(Y)$.

Proof. We use the notation in the proof of Lemma 12.2.1. If $\phi \in \mathscr{D}(W)$, then

$$\langle u_j \otimes v, \phi \rangle = \langle u_j, \psi \rangle \longrightarrow \langle u, \psi \rangle = \langle u \otimes v, \phi \rangle.$$ ∎

Proposition 12.3.7. *Suppose that* $u \in \mathscr{D}'(X)$ *and* $v \in \mathscr{D}'(Y)$. *Then*

$$\partial_x^\alpha \partial_y^\beta (u \otimes v) = \partial_x^\alpha u \otimes \partial_y^\beta v$$

for all multi-indices α *and* β.

Proof. Suppose that $\phi \in \mathscr{D}(X)$ and $\psi \in \mathscr{D}(Y)$. Then

$$\begin{aligned} \langle \partial_x^\alpha \partial_y^\beta (u \otimes v), \phi \otimes \psi \rangle &= (-1)^{|\alpha|+|\beta|} \langle u \otimes v, \partial_x^\alpha \phi \otimes \partial_y^\beta \psi \rangle \\ &= (-1)^{|\alpha|} \langle u, \partial_x^\alpha \phi \rangle (-1)^{|\beta|} \langle v, \partial_y^\beta \psi \rangle \\ &= \langle \partial_x^\alpha u, \phi \rangle \langle \partial_y^\beta v, \psi \rangle \\ &= \langle \partial_x^\alpha u \otimes \partial_y^\beta v, \phi \otimes \psi \rangle. \end{aligned}$$

The general case follows from Corollary 12.3.2. ∎

Remark 12.3.8. All results in this section also holds for distributions with compact support, where the tensor products act on C^∞-functions.

12.4. Convolutions of Distributions

We next consider convolutions of distributions. Suppose first that $f, g \in L^1(\mathbf{R}^d)$ and that both functions have compact support. Then $f * g \in L^1(\mathbf{R}^d)$ and thus defines a regular distribution on $\mathscr{D}(\mathbf{R}^d)$. This distribution acts on a test function $\phi \in \mathscr{D}(\mathbf{R}^d)$ in the following way:

$$\begin{aligned}\langle f * g, \phi\rangle &= \int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^d} f(x)g(y-x)\,dx\right)\phi(y)\,dy\\ &= \int_{\mathbf{R}^d} f(x)\left(\int_{\mathbf{R}^d} g(y)\,\phi(x+y)\,dy\right)dx\\ &= \langle f(x), \langle g(y), \phi(x+y)\rangle\rangle\\ &= \langle f(x)\otimes g(y), \phi(x+y)\rangle\end{aligned}$$

Notice that the assumption about compact supports is needed not for the existence of the integrals above, but to justify the last equality. This shows that the convolution between $u \in \mathscr{D}'(\mathbf{R}^d)$ and $v \in \mathscr{D}'(\mathbf{R}^d)$ in principle should be defined by

$$\langle u * v, \phi\rangle = \langle u(x)\otimes v(y), \phi(x+y)\rangle, \quad \phi \in \mathscr{D}(\mathbf{R}^d).$$

In general, the right-hand side in this identity is however not defined because the function $(x, y) \mapsto \phi(x+y)$ does not have compact support. One case when this makes sense is when if $u, v \in \mathscr{E}'(\mathbf{R}^d)$ since then $u \otimes v$ belongs to $\mathscr{E}'(\mathbf{R}^{2d})$.

We will assume that a weaker condition holds. Suppose that $\operatorname{supp}\phi \subset \overline{B_r(0)}$ for some $r > 0$. Then $\operatorname{supp}\phi(\cdot+\cdot) \subset N_r$, where $N_r = \{(x, y) \in \mathbf{R}^{2d} : |x+y| \leqslant r\}$. The condition, that we will require in the definition of convolutions, is the following:

$$(\operatorname{supp} u \times \operatorname{supp} v) \cap N_r \quad \text{is bounded for every } r > 0. \tag{12.4}$$

Example 12.4.1. The condition (12.4) is satisfied for instance if

(i) $u \in \mathscr{E}'(\mathbf{R}^d)$ or $v \in \mathscr{E}'(\mathbf{R}^d)$;

(ii) $\operatorname{supp} u, \operatorname{supp} v \subset \{x \in \mathbf{R}^d : x_j \geqslant c \text{ for every } j\}$ for some number $c \in \mathbf{R}$. □

Definition 12.4.2. Suppose that $u, v \in \mathscr{D}'(\mathbf{R}^d)$ satisfy (12.4). Then the **convolution** between u and v is defined by

$$\langle u * v, \phi\rangle = \langle u(x)\otimes v(y), \rho(x, y)\phi(x+y)\rangle, \quad \phi \in D(\mathbf{R}^d),$$

where $\operatorname{supp}\phi \subset \overline{B_r(0)}$ and $\rho \in \mathscr{D}(\mathbf{R}^{2d})$ is chosen so that $\rho = 1$ in a neighbourhood of the set $(\operatorname{supp} u \times \operatorname{supp} v) \cap N_r$.

Remark 12.4.3. This definition is as expected independent of the choice of the function ρ. In fact, if ρ_1 and ρ_2 are two such functions, then

$$\langle u(x)\otimes v(y), (\rho_1(x, y) - \rho_2(x, y))\phi(x+y)\rangle = 0$$

since $\rho_1 - \rho_2 = 0$ in a neighbourhood of $(\operatorname{supp} u \times \operatorname{supp} v) \cap N_r$. We will therefore usually omit ρ and just write

$$\langle u * v, \phi\rangle = \langle u(x)\otimes v(y), \phi(x+y)\rangle.$$

Theorem 12.4.4. *Suppose that* $u, v \in \mathscr{D}'(\mathbf{R}^d)$ *satisfy* (12.4)*. Then* $u * v \in \mathscr{D}'(\mathbf{R}^d)$ *with* $\operatorname{supp} u * v \subset \operatorname{supp} u + \operatorname{supp} v$*.*

Proof. The first statement follows from Theorem 12.2.4 and the second is left as an exercise to the reader. ∎

Example 12.4.5. We have

$$u * \delta = \delta * u = u \quad \text{for every } u \in \mathscr{D}'(\mathbf{R}^d).$$

Indeed,

$$\langle u * \delta, \phi \rangle = \langle u(x) \otimes \delta(y), \phi(x+y) \rangle = \langle u(x) \langle \delta(y), \phi(x+y) \rangle \rangle = \langle u, \phi \rangle$$

and

$$\langle \delta * u, \phi \rangle = \langle \delta(x) \otimes u(y), \phi(x+y) \rangle = \langle \delta(x), \langle u(y), \phi(x+y) \rangle \rangle = \langle u, \phi \rangle$$

for every $\phi \in \mathscr{D}(\mathbf{R}^d)$. The same calculations show more generally that

$$u * \partial^\alpha \delta = \partial^\alpha \delta * u = \partial^\alpha u \quad \text{for every multi-index } \alpha.$$ □

12.5. Properties of the Convolution

It is easy to show that convolution is both commutative and distributive.

Proposition 12.5.1. *Suppose that* $u, v \in \mathscr{D}'(\mathbf{R}^d)$ *satisfy* (12.4)*. Then*

$$u * v = v * u.$$

Proof. Given $\phi \in \mathscr{D}(\mathbf{R}^d)$, choose $\rho \in \mathscr{D}(\mathbf{R}^d)$ symmetric. Then, according to Proposition 12.3.3,

$$\begin{aligned} \langle u * v, \phi \rangle &= \langle u(x) \otimes v(y), \rho(x,y)\phi(x,y) \rangle = \langle v(y) \otimes u(x), \rho(x,y)\phi(x+y) \rangle \\ &= \langle v(y) \otimes u(x), \rho(y,x)\phi(y+x) \rangle = \langle v * u, \phi \rangle. \end{aligned}$$ ∎

Proposition 12.5.2. *Suppose that* $u, v, w \in \mathscr{D}'(\mathbf{R}^d)$ *and that* (u, w) *and* (v, w) *satisfy* (12.4)*. Then*

$$(u + v) * w = u * w + v * w.$$

Proof. Since $\operatorname{supp}(u+v) \subset \operatorname{supp} u \cup \operatorname{supp} v$, it follows that $(u+v, w)$ satisfies (12.4), so $(u+v) * w$ is defined. The rest of the proof is routine. ∎

To prove that convolution is associative is a bit harder than to prove commutativity and distributivity. We will therefore omit the proof.

Proposition 12.5.3. *Suppose that* $u, v, w \in \mathscr{D}'(\mathbf{R}^d)$ *and that the set*

$$(\operatorname{supp} u \times \operatorname{supp} v \times \operatorname{supp} w) \cap \{(x,y,z) \in \mathbf{R}^{3d} : |x+y+z| \leqslant r\} \tag{12.5}$$

is bounded for every $r > 0$*. Then*

$$u * (v * w) = (u * v) * w. \tag{12.6}$$

Remark 12.5.4. A few comments are in order.

(i) One can show that (12.5) implies that (u, v) and (v, w) satisfy (12.4); let us for instance show that (12.4) holds. We can of course assume that $w \neq 0$. Suppose that $z_0 \in \operatorname{supp} w$ and choose $r > |z_0|$. Then is the set

$$(\operatorname{supp} u \times \operatorname{supp} v \times \{z_0\}) \cap \{(x, y, z) \in \mathbf{R}^{3d} : |x + y + z| \leqslant r\}$$

bounded by assumption. It follows that the subset

$$(\operatorname{supp} u \times \operatorname{supp} v \times \{z_0\}) \cap \{(x, y, z) \in \mathbf{R}^{3d} : |x + y| \leqslant r - |z_0|\}$$

is also bounded, which gives (12.4).

(ii) If (u, v) satisfies (12.4) and $w \in \mathscr{E}'(\mathbf{R}^d)$, then (12.5) holds.

(iii) Suppose that (u, v, w) does not satisfy (12.5). Then (12.6) does not have to hold. Take for instance $u = 1$, $v = \delta'$, and $w = \delta$. Then

$$1 * (\delta' * H) = 1 * H' = 1 * \delta = 1, \quad \text{but} \quad (1 * \delta') * H = 1' * H = 0 * H = 0.$$

Notice, however, that $(1, \delta')$ and (δ', H) satisfy (12.4) since δ' has compact support.

Proposition 12.5.5. *Suppose that $u_j \to u$ in $\mathscr{D}'(\mathbf{R}^d)$, that (u, v) satisfies* (12.4)*, and that (u_j, v) satisfy* (12.4) *uniformly with respect to j. Then $u_j * v \to u * v$ in $\mathscr{D}'(\mathbf{R}^d)$.*

The assumption about uniformity means that there for every $r > 0$ exists a bounded set B_r such that

$$(\operatorname{supp} u_j \times \operatorname{supp} v) \cap N_r \subset B_r \quad \text{for every } j.$$

Proof (Proposition 12.5.5). Suppose that $\phi \in \mathscr{D}(\mathbf{R}^d)$. Then, according to Proposition 12.3.6,

$$\langle u_j * v, \phi \rangle = \langle u_j(x) \otimes v(y), \phi(x + y) \rangle \longrightarrow \langle u(x) \otimes v(y), \phi(x + y) \rangle = \langle u * v, \phi \rangle. \quad \blacksquare$$

Proposition 12.5.6. *Suppose that $u, v \in \mathscr{D}'(\mathbf{R}^d)$ and that (u, v) satisfies* (12.4)*. Then*

$$\partial^\alpha (u * v) = \partial^\alpha u * v = u * \partial^\alpha v$$

for every multi-index α.

Proof. Suppose that $\phi \in \mathscr{D}(\mathbf{R}^d)$. Then, according to Proposition 12.3.7,

$$\begin{aligned}
\langle \partial^\alpha (u * v), \phi \rangle &= (-1)^{|\alpha|} \langle u * v, \partial^\alpha \phi \rangle = \langle u(x) \otimes v(y), \rho(x, y) \partial^\alpha \phi(x + y) \rangle \\
&= \langle u(x) \otimes v(y), \partial_x^\alpha (\rho(x, y) \phi(x + y)) \rangle \\
&= \langle \partial_x^\alpha u(x) \otimes v(y), \rho(x, y) \phi(x + y) \rangle \\
&= \langle \partial^\alpha u * v, \phi \rangle.
\end{aligned}$$

This shows that $\partial^\alpha (u * v) = \partial^\alpha u * v$. The other identity is proved similarly. $\blacksquare$

The next proposition shows that the convolution between a distribution and a test function is a smooth function.

Proposition 12.5.7. *Suppose* $u \in \mathscr{D}'(\mathbf{R}^d)$ *and* $f \in \mathscr{D}(\mathbf{R}^d)$. *Then* $u * f \in \mathscr{E}(\mathbf{R}^d)$ *and*

$$u * f(x) = \langle u(y), f(x-y)\rangle \quad \textit{for every } x \in \mathbf{R}^d. \tag{12.7}$$

Proof. Notice that the convolution $u * f$ is defined since f has compact support and that the right-hand side in (12.7) is defined for every fixed $x \in \mathbf{R}^d$ since $f(x-\cdot)$ also has compact support. Suppose that $\operatorname{supp} f \subset B_r(0)$ and choose $\rho \in \mathscr{D}(\mathbf{R}^d)$ such that $\rho = 1$ on $B_{2r}(0)$. Lemma 12.2.1 then shows that the function

$$\eta(x) = \langle u(y), f(x-y)\rangle = \langle u(y), \rho(y)f(x-y)\rangle, \quad |x| < r,$$

belongs to $\mathscr{E}(B_r(0))$. This holds for every sufficiently large r, so we have $\eta \in \mathscr{E}(\mathbf{R}^d)$. Now suppose that $\phi \in \mathscr{D}(\mathbf{R}^d)$ with $\operatorname{supp}\phi \subset B_r(0)$. Then

$$\langle u * f, \phi\rangle = \langle u(x)\otimes f(y), \rho(x)\rho(y)\phi(x+y)\rangle = \langle u(x), \langle f(y), \rho(x)\rho(y)\phi(x+y)\rangle\rangle.$$

We also have

$$\begin{aligned}\langle f(y), \rho(x)\rho(y)\phi(x+y)\rangle &= \int_{\mathbf{R}^d} f(y)\phi(x+y)\,dy = \int_{\mathbf{R}^d} f(y-x)\phi(y)\,dy\\ &= \langle \phi(y), \rho(y)f(y-x)\rangle\end{aligned}$$

for every $x \in \mathbf{R}^d$. This shows that

$$\begin{aligned}\langle u * f, \phi\rangle &= \langle u(x), \langle \phi(y), \rho(y)f(y-x)\rangle\rangle = \langle u(x)\otimes\phi(y), \rho(y)f(y-x)\rangle\\ &= \langle \phi(y)\otimes u(x), \rho(y)f(y-x)\rangle = \int_{\mathbf{R}^d}\phi(y)\langle u(x), f(y-x)\rangle\,dy\\ &= \int_{\mathbf{R}^d}\langle u(y), f(x-y)\rangle\phi(x)\,dx,\end{aligned}$$

which proves (12.7). ∎

Example 12.5.8. The **Hilbert transform** $H \in \mathscr{D}'(\mathbf{R})$ is defined by

$$H\phi(x) = \operatorname{pv}\frac{1}{x} * \phi(x), \quad \phi \in \mathscr{D}(\mathbf{R}).$$

Using Proposition 12.5.7, we see that

$$H\phi(x) = \langle \operatorname{pv}\frac{1}{y}, \phi(x-y)\rangle = \lim_{\varepsilon\to 0}\int_{|y|\geqslant\varepsilon}\frac{\phi(x-y)}{y}\,dy = \operatorname{pv}\int_{-\infty}^{\infty}\frac{\phi(x-y)}{y}\,dy. \qquad \square$$

12.6. Density Results

A consequence of Proposition 12.5.7 is the following result about regularization of distributions.

Proposition 12.6.1. *Suppose that $u \in \mathscr{D}'(\mathbf{R}^d)$ and let $(\phi_j)_{j=1}^{\infty}$ be an approximate identity. Then $\phi_j * u \in \mathscr{E}(\mathbf{R}^d)$ and $\phi_j * u \to u$ in $\mathscr{D}'(\mathbf{R}^d)$ as $j \to \infty$.*

Proof. As in Example 8.5.4, we have $\phi_j \to \delta$ in $\mathscr{D}'(\mathbf{R}^d)$. It then follows from Proposition 12.5.5 that

$$\phi_j * u \to \delta * u = u \quad \text{in } \mathscr{D}'(\mathbf{R}^d).$$ ■

Example 12.6.2. We will use the result in Proposition 12.6.1 to give a new proof of Lemma 10.4.3. Suppose that $u \in \mathscr{D}'(\mathbf{R})$ and $u' = 0$. Then, according to Proposition 12.5.6,

$$(\phi_j * u)' = \phi_j * u' = 0.$$

Since $\phi_j * u$ is a smooth function, this shows that $\phi_j * u$ is a constant C_j. Because $\phi_j * u \to u$ in $\mathscr{D}'(\mathbf{R})$, it follows that C_j converges to some constant C. □

The following two density results follow from Proposition 12.6.1.

Corollary 12.6.3. *The set $\mathscr{D}(\mathbf{R}^d)$ is dense in $\mathscr{D}'(\mathbf{R}^d)$.*

Proof. Suppose that $u \in \mathscr{D}'(\mathbf{R}^d)$. Take a cut-off function $\chi \in \mathscr{D}(\mathbf{R}^d)$ such that $\chi = 1$ on $B_1(0)$. If $(\phi_j)_{j=1}^{\infty}$ is an approximate identity, it then follows easily from Proposition 12.6.1 that the sequence $\chi(x/j)\phi_j * u(x)$, $j = 1, 2, \dots$, of compactly supported test functions converges to u in $\mathscr{D}'(\mathbf{R}^d)$. ■

Corollary 12.6.4. *The set $\mathscr{D}(X)$ is dense in $\mathscr{D}'(X)$.*

Proof. As in the proof of Theorem 8.3.2, let $(K_j)_{j=1}^{\infty}$ be an increasing sequence of compact subsets to X such that $X = \bigcup_{j=1}^{\infty} K_j$. Then choose $\chi_j \in \mathscr{D}(X)$ such that $\chi_j = 1$ in a neighbourhood of K_j and put $u_j = \chi_j u$, $j = 1, 2, \dots$. Obviously, $u_j \in \mathscr{E}'(X)$ and we may extend u_j to an element in $\mathscr{E}'(\mathbf{R}^d)$. Let $(\psi_j)_{j=1}^{\infty}$ be an approximate identity. Then $\psi_j * u_j \in \mathscr{D}(\mathbf{R}^d)$ with support in X if j is large enough; we will show that $\psi_j * u_j \to u$ in $\mathscr{D}'(X)$. To this end, let $\phi \in \mathscr{D}(X)$. Then $\langle u, \phi \rangle = \langle u_k, \phi \rangle$ for large k. It follows that

$$\begin{aligned}\langle \psi_j * u_j, \phi \rangle &= \left\langle u_j(x), \int_{\mathbf{R}^d} \psi_j(y)\phi(x+y)\,dy \right\rangle = \left\langle u_k(x), \int_{\mathbf{R}^d} \psi_j(y)\phi(x+y)\,dy \right\rangle \\ &= \langle \psi_j * u_k, \phi \rangle\end{aligned}$$

if $j \geqslant k$ is sufficiently large. Proposition 12.6.1 now shows that

$$\langle \psi_j * u_j, \phi \rangle = \langle \psi_j * u_k, \phi \rangle \longrightarrow \langle u_k, \phi \rangle = \langle u, \phi \rangle.$$ ■

Chapter 13

Tempered Distributions

13.1. Fourier Transforms of Distributions

When trying to define the Fourier transform of distributions, a complication appears. Suppose first that $f \in L^1(\mathbf{R}^d)$ and $\phi \in \mathscr{D}(\mathbf{R}^d)$. Proposition 6.2.8 then shows that

$$\langle u_{\hat{f}}, \phi \rangle = \int_{\mathbf{R}^d} \widehat{f}(x)\phi(x)\,dx = \int_{\mathbf{R}^d} f(x)\widehat{\phi}(x)\,dx.$$

So far everything is fine. Notice, however, that $\widehat{\phi}$ does not belong to $\mathscr{D}(\mathbf{R}^d)$ unless ϕ is identically 0 since $\widehat{\phi}$ can be extended to an entire function on $\mathbf{C}^d$ and thus cannot have compact support without being 0 everywhere. This means that we do not have

$$\langle u_{\hat{f}}, \phi \rangle = \langle u_f, \widehat{\phi} \rangle$$

for $\phi \in \mathscr{D}(\mathbf{R}^d)$. The class $\mathscr{D}(\mathbf{R}^d)$ is therefore not suitable when working with Fourier transforms of distributions. What is needed is a class of test functions that is invariant under the Fourier transform.

13.2. The Schwartz Class

Definition 13.2.1. A function $\phi \in C^\infty(\mathbf{R}^d)$ is said to be **rapidly decreasing** if

$$\|\phi\|_{\alpha,\beta} = \sup_{x\in\mathbf{R}^d} |x^\alpha \partial^\beta \phi(x)| < \infty \tag{13.1}$$

for all multi-indices α and β. The vector space of all rapidly decreasing is called the **Schwartz Class** and is denoted $\mathscr{S}$.

Thus, if $\phi \in \mathscr{S}$, then ϕ and all its derivatives tend faster to 0 than $|x|^{-k}$ for any integer $k \geqslant 0$ as $|x| \to \infty$.

Example 13.2.2. It is easy to show that the function $\phi(x) = e^{-a|x|^2}$, $x \in \mathbf{R}^d$, belongs to $\mathscr{S}$ for every complex number a with positive real part. □

It follows directly from the definition that the Schwartz class is invariant under differentiation and multiplication with powers of x and that these operations are continuous on $\mathscr{S}$.

Proposition 13.2.3. *The mapping $\mathscr{S} \ni \phi \mapsto x^\alpha D^\beta \phi(x)$ is a continuous map from $\mathscr{S}$ to $\mathscr{S}$ for all multi-indices α and β.*

There is a notion of convergence in the Schwartz class.

Definition 13.2.4. A sequence $(\phi_n)_{n=1}^\infty \subset \mathscr{S}$ **converges** to $\phi \in \mathscr{S}$ if

$$\|\phi - \phi_n\|_{\alpha,\beta} \longrightarrow 0 \quad \text{as } n \to \infty$$

for all multi-indices α and β.

Proposition 13.2.5. *The set $\mathscr{D}(\mathbf{R}^d)$ is a dense subspace to $\mathscr{S}$.*

Proof. Convergence in $\mathscr{D}(\mathbf{R}^d)$ clearly implies convergence in $\mathscr{S}$, so $\mathscr{D}(\mathbf{R}^d)$ subspace to $\mathscr{S}$. To prove density, suppose that $\phi \in \mathscr{S}$. Take $\chi \in \mathscr{D}(\mathbf{R}^d)$ such that $\chi = 1$ on $\overline{B_1(0)}$ and put $\chi_n(x) = \chi(x/n)$, $x \in \mathbf{R}^d$, for $n = 1, 2, \ldots$. Then the sequence $\phi_n = \chi_n \phi$, $n = 1, 2, \ldots$, of functions belonging to $\mathscr{D}(\mathbf{R}^d)$ converges to ϕ in $\mathscr{S}$. In fact,

$$\|\phi - \phi_n\|_{\alpha,\beta} = \sup_{|x| \geqslant n} |x^\alpha \partial^\beta (\phi(x)(1 - \chi_n(x)))| \leqslant C \sum_{\gamma \leqslant \beta} \sup_{|x| \geqslant n} |x^\alpha \partial^\gamma \phi(x)|. \quad (13.2)$$

Moreover, if $|x| \geqslant n$, then $|x_j| \geqslant n/\sqrt{d}$ for some j, so if α' equals α with 1 added at place j, then

$$|x^\alpha \partial^\gamma \phi(x)| \leqslant \sqrt{d} n^{-1} |x^{\alpha'} \partial^\gamma \phi(x)| \leqslant \sqrt{d} n^{-1} \|\phi\|_{\alpha',\gamma}.$$

Together with (13.2), this shows that ϕ_n converges to ϕ in $\mathscr{S}$. ∎

Proposition 13.2.6. *The set $\mathscr{S}$ is a dense subspace to $L^p(\mathbf{R}^d)$ for $1 \leqslant p < \infty$.*

Proof. Suppose that $\phi \in \mathscr{S}$. Then

$$\begin{aligned} \|\phi\|_p &= \left(\int_{\mathbf{R}^d} |(1 + |x|^2)^d \phi(x)|^p \frac{dx}{(1 + |x|^2)^{dp}} \right)^{1/p} \leqslant C \|(1 + |x|^2)^d \phi(x)\|_\infty \\ &\leqslant C \sum_{|\alpha| \leqslant 2d} \|\phi\|_{\alpha,0} < \infty, \end{aligned}$$

which shows that $\phi \in L^p(\mathbf{R}^d)$, so $\mathscr{S}$ is a subset to $L^p(\mathbf{R}^d)$. The density of $\mathscr{S}$ in $L^p(\mathbf{R}^d)$ follows from the fact that $C_c^\infty(\mathbf{R}^d)$ is dense in $L^p(\mathbf{R}^d)$. ∎

The importance of the Schwartz class in distribution theory stems from the following theorem.

Theorem 13.2.7. *The Fourier transform $\mathscr{F}$ is a continuous map from $\mathscr{S}$ to $\mathscr{S}$.*

Proof. Suppose that $\phi \in \mathscr{S}$. Then $\widehat{\phi} \in C^\infty(\mathbf{R}^d)$ according to Proposition 6.2.11. Moreover, $\widehat{\phi} \in \mathscr{S}$ satisfies (13.1) for all multi-indices α and β since

$$\begin{aligned} \|\widehat{\phi}\|_{\alpha,\beta} &= \|\xi^\alpha \partial^\beta \widehat{\phi}(\xi)\|_\infty = \|D^\alpha \widehat{(x^\beta \phi(x))}\|_\infty \leqslant \|D^\alpha (x^\beta \phi(x))\|_1 \\ &\leqslant C \|(1 + |x|^2)^d D^\alpha (x^\beta \phi(x))\|_\infty < \infty. \end{aligned}$$

This inequality also shows that the Fourier transform is continuous. ∎

Suppose that $\phi \in \mathscr{S}$. Then ϕ is bounded and continuous and $\widehat{\phi} \in \mathscr{S} \subset L^1(\mathbf{R}^d)$, so it follows from Theorem 6.4.1 that

$$\phi(x) = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \widehat{\phi}(\xi) e^{i\xi \cdot x} \, d\xi \quad \text{for every } x \in \mathbf{R}^d.$$

This shows that the Fourier transform $\mathscr{F} : \mathscr{S} \to \mathscr{S}$ is invertible and the inverse is $\mathscr{F}^{-1} = (2\pi)^{-d} \check{\mathscr{F}}$. Theorem 13.2.6 also implies that the inverse is continuous.

Theorem 13.2.8. *The Fourier transform is a homeomorphism from $\mathscr{S}$ to $\mathscr{S}$.*

13.3. Tempered Distributions

We next define the dual space to $\mathscr{S}$.

Definition 13.3.1. A **tempered distribution** is a sequentially continuous, linear functional on $\mathscr{S}$. We denote the class of tempered distributions by $\mathscr{S}'$.

Definition 13.3.2. A sequence $(u_n)_{n=1}^{\infty} \subset \mathscr{S}'$ **converges** to $u \in \mathscr{S}'$ if

$$\langle u_n, \phi \rangle \to \langle u, \phi \rangle \quad \text{for every } \phi \in \mathscr{S}.$$

Notice that if $u \in \mathscr{S}'$, then since $\mathscr{D}(\mathbf{R}^d) \subset \mathscr{S}$, the restriction of u to $\mathscr{D}(\mathbf{R}^d)$ belongs to $\mathscr{D}'(\mathbf{R}^d)$. We may thus consider $\mathscr{S}'$ as a subspace to $\mathscr{D}'(\mathbf{R}^d)$.

The proof of the following theorem is left as an exercise to the reader.

Theorem 13.3.3. *A linear functional u on $\mathscr{S}$ belongs to $\mathscr{S}'$ if and only if there exist a constant $C \geqslant 0$ and an integer $m \geqslant 0$ such that*

$$|\langle u, \phi \rangle| \leqslant C \sum_{|\alpha|,|\beta| \leqslant m} \|\phi\|_{\alpha,\beta} \tag{13.3}$$

for every function $\phi \in \mathscr{S}$.

It follows from Theorem 11.1.4 that (13.3) is satisfied if $u \in \mathscr{E}'(\mathbf{R}^d)$. This shows that $\mathscr{E}'(\mathbf{R}^d)$ is a subset to $\mathscr{S}'$. We thus have $\mathscr{E}'(\mathbf{R}^d) \subset \mathscr{S}' \subset \mathscr{D}'(\mathbf{R}^d)$.

Example 13.3.4. We will show that $L^p(\mathbf{R}^d) \subset \mathscr{S}'$ for $1 \leqslant p \leqslant \infty$. Suppose that $f \in L^p(\mathbf{R}^d)$. Then, for $\phi \in \mathscr{S}$,

$$|\langle u_f, \phi \rangle| \leqslant \int_{\mathbf{R}^d} |f(x)||\phi(x)|\, dx \leqslant \|f\|_p \|\phi\|_{p'} \leqslant C\|f\|_p \sum_{|\alpha| \leqslant 2d} \|\phi\|_{\alpha,0}$$

according to the proof of Proposition 13.2.6. It thus follows from Theorem 13.3.3 that $u_f \in \mathscr{S}'$. □

Example 13.3.5. Suppose that $f \in C(\mathbf{R}^d)$ is a function of **polynomial growth**, meaning that there exist a constant $C \geqslant 0$ and an integer $m \geqslant 0$ such that

$$|f(x)| \leqslant C(1 + |x|)^m \quad \text{for every } x \in \mathbf{R}^d.$$

Then, for $\phi \in \mathscr{S}$,

$$|\langle u_f, \phi \rangle| \leqslant C \int_{\mathbf{R}^d} (1 + |x|)^{m+d+1} |\phi(x)| \frac{dx}{(1 + |x|)^{d+1}} \leqslant C \sum_{|\alpha| \leqslant m+d+1} \|\phi\|_{\alpha,0}.$$

This shows that $u_f \in \mathscr{S}'$. In particular, every polynomial belongs to $\mathscr{S}'$. □

The next proposition shows that $\mathscr{S}'$ is invariant under multiplication with polynomials and differentiation.

Proposition 13.3.6. *Suppose that $u \in \mathscr{S}'$. Then*

(i) $x^\alpha u \in \mathscr{S}'$ for every multi-index α;

(ii) $\partial^\beta u \in \mathscr{S}'$ *for every multi-index* β.

Proof. The proof of these properties are almost identical, so let us just prove (i). Suppose that $\phi_n \to \phi$ in $\mathscr{S}$. Then, by Proposition 13.2.3,

$$\langle x^\alpha u, \phi_n \rangle = \langle u, x^\alpha \phi_n \rangle \longrightarrow \langle u, x^\alpha \phi \rangle = \langle x^\alpha u, \phi \rangle. \qquad \blacksquare$$

The next example shows that there are regular tempered distributions that are not of polynomial growth.

Example 13.3.7. The function $f(x) = \sin(e^x)$, $x \in \mathbf{R}$, belongs to $\mathscr{S}'$ since f is bounded. It therefore follows from Proposition 13.3.6 that $f' \in \mathscr{S}'$. However, $f'(x) = e^x \cos(e^x)$, $x \in \mathbf{R}$, is not of polynomial growth. As a comparison, notice that the function $g(x) = e^x$, $x \in \mathbf{R}$, does not belong to $\mathscr{S}'$. In fact, if $\phi \in \mathscr{S}$ and $\phi(x) = e^{-|x|/2}$ for $|x| \geqslant 1$, then

$$\int_{-\infty}^{\infty} g(x)\phi(x)\,dx = \int_{-1}^{1} e^x \phi(x)\,dx + \int_{|x|\geqslant 1} e^x e^{-|x|/2}\,dx = \infty. \qquad \square$$

13.4. The Fourier Transform

We are now ready to define the Fourier transform of a tempered distribution.

Definition 13.4.1. The **Fourier transform** $\hat{u}$ of a tempered distribution u is defined through

$$\langle \hat{u}, \phi \rangle = \langle u, \hat{\phi} \rangle, \quad \phi \in \mathscr{S}.$$

Remark 13.4.2.

(a) Notice that if $u \in \mathscr{S}'$, then $\hat{u} \in \mathscr{S}'$ since the Fourier transform is continuous on $\mathscr{S}$ according to Theorem 13.2.7.

(b) The Fourier transform is also continuous on $\mathscr{S}'$. Indeed, if $u_n \to u$ in $\mathscr{S}'$, then

$$\langle \widehat{u_n}, \phi \rangle = \langle u_n, \hat{\phi} \rangle \longrightarrow \langle u, \hat{\phi} \rangle = \langle \hat{u}, \phi \rangle$$

for every function $\phi \in \mathscr{S}$ since $\hat{\phi} \in \mathscr{S}$, which shows that $\widehat{u_n} \to \hat{u}$ in $\mathscr{S}'$.

(c) If $f \in L^1(\mathbf{R}^d)$, then

$$\langle \widehat{u_f}, \phi \rangle = \langle u_f, \hat{\phi} \rangle = \int_{\mathbf{R}^d} f(x)\hat{\phi}(x)\,dx = \int_{\mathbf{R}^d} \hat{f}(x)\phi(x)\,dx = \langle u_{\hat{f}}, \phi \rangle$$

for every function $\phi \in \mathscr{S}$, which shows that the distributional Fourier transform of f coincides with the ordinary transform.

Example 13.4.3.

1. Let us first calculate the Fourier transform of δ:

$$\langle \hat{\delta}, \phi \rangle = \langle \delta, \hat{\phi} \rangle = \hat{\phi}(0) = \int_{\mathbf{R}^d} \phi(x)\,dx = \langle 1, \phi \rangle$$

for every function $\phi \in \mathscr{S}$, which shows that $\hat{\delta} = 1$.

2. We next calculate the Fourier transform of 1. If $\phi \in \mathscr{S}$, then

$$\langle \widehat{1}, \phi \rangle = \langle 1, \widehat{\phi} \rangle = \int_{\mathbf{R}^d} \widehat{\phi}(\xi)\, d\xi = (2\pi)^d \phi(0) = \langle (2\pi)^d \delta, \phi \rangle$$

according to the inversion formula, which shows that $\widehat{1} = (2\pi)^d \delta$. □

13.5. Properties of the Fourier Transform

The properties of the Fourier transform on the Schwartz class immediately carry over to tempered distributions.

Proposition 13.5.1. *Suppose that $u \in \mathscr{S}'$. Then the following properties hold:*

(i) if $h \in \mathbf{R}^d$, then $\widehat{\tau_h u} = e^{-ih\cdot\xi}\widehat{u}$;

(ii) if $h \in \mathbf{R}^d$, then $\widehat{e^{ih\cdot x}u} = \tau_h \widehat{u}$;

(iii) $(\hat{u})\check{} = (\check{u})\hat{}$;

(iv) if $t \in \mathbf{R}$ and $t \neq 0$, then $\widehat{u_t} = |t|^{-d}\widehat{u}_{t^{-1}}$;

(v) $\widehat{\partial^\alpha u} = (i\xi)^\alpha \widehat{u}$ for every multi-index α;

(vi) $\widehat{x^\alpha u} = i^{|\alpha|}\partial^\alpha \widehat{u}$ for every multi-index α.

Proof. We will prove (v) and leave the other properties as exercises to the reader. Suppose that $\phi \in \mathscr{S}$. Then

$$\langle \widehat{\partial^\alpha u}, \phi \rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha \widehat{\phi} \rangle = (-1)^{|\alpha|}\langle u, \widehat{(-ix)^\alpha \phi(x)} \rangle = \langle (i\xi)^\alpha \widehat{u}, \phi \rangle. \qquad \blacksquare$$

Example 13.5.2. We make two applications of Proposition 6.2.2.

1. Let us first calculate the Fourier transform of x^α:

$$\widehat{x^\alpha} = \widehat{x^\alpha 1} = i^{|\alpha|}\partial^\alpha \widehat{1} = (2\pi)^d i^{|\alpha|}\partial^\alpha \delta.$$

2. We next calculate the Fourier transform of $e^{ia\cdot\xi}$, where $a \in \mathbf{R}^d$ is a constant:

$$\widehat{e^{ia\cdot\xi}} = \widehat{e^{ia\cdot\xi}1} = \tau_a \widehat{1} = (2\pi)^d \tau_a \delta = (2\pi)^d \delta_a$$

so that

$$\delta_a = (2\pi)^{-d}\widehat{e^{ia\cdot\xi}}.$$

If we apply this identity to a test function $\phi \in \mathscr{S}$, we obtain

$$\phi(a) = \frac{1}{2\pi}\int_{\mathbf{R}^d} \widehat{\phi}(\xi) e^{i\xi\cdot a}\, d\xi,$$

which gives us a new proof of the inversion formula for $\mathscr{S}$. □

It follows from Proposition 6.2.2 that if $u \in \mathscr{S}'$ is even/odd, then $\hat{u}$ is also even/odd. For instance, if u is even, i.e., $\check{u} = u$, then

$$(\hat{u})\check{} = (\check{u})\hat{} = \hat{u}.$$

Example 13.5.3. We next calculate the Fourier transform of the Cauchy principal value $u = \operatorname{pv}\frac{1}{x}$. Notice that

$$\langle \operatorname{pv}\frac{1}{x}, \phi(x)\rangle = \lim_{\varepsilon\to 0}\int_{\varepsilon\leqslant|x|<1}\frac{\phi(x)}{x}\,dx + \int_{|x|\geqslant 1}\frac{\phi(x)}{x}\,dx$$

for $\phi \in \mathscr{S}$. This shows that u is the sum of a distribution with compact support and a L^2-function, and thus belongs to $\mathscr{S}'$. If we now apply the Fourier transform to the identity $xu = 1$, we obtain

$$i\xi\hat{u}' = 2\pi\delta, \quad \text{that is} \quad \hat{u}' = -2\pi i\delta.$$

Every solution to the last differential equation can be written $\hat{u} = -2\pi i(H + C)$ for some constant C. According to Example 9.3.8, u is odd, so the same holds for $\hat{u}$. This shows that $C = -\frac{1}{2}$, so

$$\widehat{\operatorname{pv}\frac{1}{x}} = -i\pi \operatorname{sgn}\xi. \qquad \square$$

13.6. The Inversion Formula

The inversion formula for the Fourier transform of course generalizes to $\mathscr{S}'$.

Theorem 13.6.1. *Suppose that $u \in \mathscr{S}'$. Then $u = (2\pi)^{-d}(\hat{\hat{u}})\check{}$.*

Proof. Suppose that $\phi \in \mathscr{S}$. Then, according to Corollary **??** and Theorem 6.4.1,

$$\langle(\hat{\hat{u}})\check{}, \phi\rangle = \langle u, (\check{\phi})\hat{\hat{}}\rangle = \langle(u, \hat{\hat{\phi}})\check{}\rangle = (2\pi)^d\langle u, \phi\rangle. \qquad \blacksquare$$

Example 13.6.2. It follows from the inversion formula and Example 13.5.3 that

$$(\operatorname{sgn} x)\hat{} = \frac{i}{\pi}(\operatorname{pv}\frac{1}{x})\hat{\hat{}} = 2\pi\frac{i}{\pi}(\operatorname{pv}\frac{1}{x})\check{} = -2i\operatorname{pv}\frac{1}{x}.$$

Since $H = \frac{1}{2}(\operatorname{sgn} x + 1)$, this implies that

$$\hat{H} = -i\operatorname{pv}\frac{1}{x} + \pi\delta. \qquad \square$$

Corollary 13.6.3. *The Fourier transform is a continuous homeomorphism on $\mathscr{S}'$.*

13.7. The Convolution Theorem

The next theorem shows that the Fourier transform of a distribution with compact support is a smooth function that can be calculated in essentially the same way as the Fourier transform of a L^1-function.

Theorem 13.7.1. *Suppose that $u \in \mathscr{E}'(\mathbf{R}^d)$. Then $\hat{u} \in \mathscr{E}(\mathbf{R}^d)$ and*

$$\hat{u}(\xi) = \langle u(x), e^{-ix\cdot\xi}\rangle, \quad \xi \in \mathbf{R}^d. \tag{13.4}$$

Proof. Suppose that $\phi \in \mathscr{D}(\mathbf{R}^d)$. Then, according to Proposition 12.3.3 (see Remark 12.3.8),

$$\begin{aligned}\langle \widehat{u}, \phi\rangle &= \langle u, \widehat{\phi}\rangle = \langle u(x), \langle \phi(\xi), e^{-ix\cdot\xi}\rangle\rangle = \langle u(x)\otimes\phi(\xi), e^{-ix\cdot\xi}\rangle \\ &= \langle \phi(\xi)\otimes u(x), e^{-ix\cdot\xi}\rangle = \int_{\mathbf{R}^d}\langle u(x), e^{-ix\cdot\xi}\rangle \phi(\xi)\,d\xi.\end{aligned}$$

This establishes (13.4) since $\mathscr{D}(\mathbf{R}^d)$ is dense in $\mathscr{S}$ by Proposition 13.2.5. The fact that $\widehat{u} \in \mathscr{E}(\mathbf{R}^d)$ follows from Remark 12.2.2. ∎

Theorem 13.7.2. *Suppose that $u, v \in \mathscr{E}'(\mathbf{R}^d)$. Then*

$$\widehat{u * v} = \widehat{u}\widehat{v}. \tag{13.5}$$

Proof. Theorem 13.7.1 shows that

$$\begin{aligned}\widehat{u * v}(\xi) &= \langle u * v(x), e^{-ix\cdot\xi}\rangle = \langle u(x)\otimes v(y), e^{-i(x+y)\cdot\xi}\rangle \\ &= \langle u(x), e^{-ix\cdot\xi}\rangle\langle v(y), e^{-iy\cdot\xi}\rangle = \widehat{u}(\xi)\widehat{v}(\xi)\end{aligned}$$

for $\xi \in \mathbf{R}^d$. ∎

We will next show that (13.5) in fact holds true if $u \in \mathscr{S}'$ and $v \in \mathscr{E}'(\mathbf{R}^d)$. For this, we need to know something about multipliers on $\mathscr{S}'$. We begin by a definition.

Definition 13.7.3. Denote by $\mathscr{O}_M(\mathbf{R}^d)$ the class of functions $f \in \mathscr{E}(\mathbf{R}^d)$ such that f and all of its derivatives are of polynomial growth.

The following lemma shows that the functions, belonging to $\mathscr{O}_M(\mathbf{R}^d)$, are multipliers on $\mathscr{S}'$.

Lemma 13.7.4. *Suppose that $u \in \mathscr{S}'$ and $f \in \mathscr{O}_M(\mathbf{R}^d)$. Then $fu \in \mathscr{S}'$.*

Proof. We will just sketch the proof. One first shows that
(i) $f\phi \in \mathscr{S}$ for every function $\phi \in \mathscr{S}$;
(ii) if $\phi_n \to \phi$ in $\mathscr{S}$, then $f\phi_n \to f\phi$ in $\mathscr{S}$.
It then follows that

$$\langle fu, \phi_n\rangle = \langle u, f\phi_n\rangle \longrightarrow \langle u, f\phi\rangle = \langle fu, \phi\rangle.$$ ∎

Lemma 13.7.5. *Suppose that $v \in \mathscr{E}'(\mathbf{R}^d)$. Then $\widehat{v} \in \mathscr{O}_M(\mathbf{R}^d)$.*

Proof. According to Theorem 13.7.1,

$$\widehat{v}(\xi) = \langle u(x), e^{-ix\cdot\xi}\rangle, \quad \xi \in \mathbf{R}^d.$$

Remark 12.2.2 then shows that

$$\partial^\alpha\widehat{v}(\xi) = (-i)^{|\alpha|}\langle x^\alpha v(x), e^{-ix\cdot\xi}\rangle.$$

Using the fact that $x^\alpha v \in \mathscr{E}'(\mathbf{R}^d)$, we now apply the semi-norm estimate (11.1):

$$|\partial^\alpha\widehat{v}(\xi)| = |\langle x^\alpha v(x), e^{-ix\cdot\xi}\rangle| \leqslant C\sum_{|\beta|\leqslant m}\sup_{x\in K}|\partial_x^\beta e^{-ix\cdot\xi}| \leqslant C(1+|\xi|)^m.$$ ∎

Theorem 13.7.6. *Suppose that* $u \in \mathscr{S}'$ *and* $v \in \mathscr{E}'(\mathbf{R}^d)$. *Then* $u * v \in \mathscr{S}'$ *and*

$$\widehat{u * v} = \hat{u}\hat{v}.$$

Proof. It follows from Lemma 13.7.5 that $\hat{u}\hat{v} \in \mathscr{S}'$, so $\hat{u}\hat{v} = \hat{w}$ for some $w \in \mathscr{S}'$. Let $\phi \in \mathscr{D}(\mathbf{R}^d)$. Then, according to the inversion formula,

$$\langle w, \check{\phi}\rangle = (2\pi)^{-d}\langle w, \hat{\hat{\phi}}\rangle = (2\pi)^{-d}\langle \hat{w}, \hat{\phi}\rangle = (2\pi)^{-d}\langle \hat{u}\hat{v}, \hat{\phi}\rangle = (2\pi)^{-d}\langle \hat{u}, \hat{v}\hat{\phi}\rangle.$$

Theorem 13.7.6 now shows that

$$(2\pi)^{-d}\langle \hat{u}, \hat{v}\hat{\phi}\rangle = (2\pi)^{-d}\langle \hat{u}, \widehat{v * \phi}\rangle = (2\pi)^{-d}\langle u, \widehat{\widehat{v * \phi}}\rangle = \langle u, (v * \phi)\check{}\,\rangle.$$

Notice that

$$(v * \phi)\check{}(x) = v * \phi(-x) = \langle v(y), \phi(-x - y)\rangle = \langle v(y), \check{\phi}(x + y)\rangle$$

for $x \in \mathbf{R}^d$. It follows that

$$\langle w, \check{\phi}\rangle = \langle u(x), \langle v(y), \check{\phi}(x + y)\rangle = \langle u * v, \phi\rangle.$$

Since ■

Part V

Wavelets

Appendix A

The Lebesgue Integral

In the following appendix, we summarize some facts from integration theory that is used in the main text.

A.1. Measurable Sets, Measure, Almost Everywhere

Without going into detail, we assume that there exists a class $\mathscr{M}$ of subsets to $\mathbf{R}^d$, which is large enough to contain all open and all closed subsets to $\mathbf{R}^d$ and which also is a σ-**algebra**:

(i) $\varnothing, \mathbf{R}^d \in \mathscr{M}$;

(ii) if $E \in \mathscr{M}$, then $E^c \in \mathscr{M}$ (E^c being the complement of E);

(iii) if $E_1, E_2, \ldots \in \mathscr{M}$, then $\bigcup_{n=1}^{\infty} E_n \in \mathscr{M}$.

The elements of $\mathscr{M}$ are called **measurable** subsets to $\mathbf{R}^d$. Let us remark that all subsets to $\mathbf{R}^d$, that one may run into in applications, are measurable and that is very hard to construct non-measurable sets.

To every measurable subset set E to $\mathbf{R}^d$, one can assign a number $m(E) \in [0, \infty]$, called the **measure** of E, which measures the "size" of the set. For instance, the measure of an interval is just the length of the interval. Some sets have **measure zero**. The following subsets to $\mathbf{R}^d$ are examples of sets with measure 0: all finite subsets to $\mathbf{R}^d$, $\mathbf{Q}^d$, and more generally all countable subsets to $\mathbf{R}^d$, the standard Cantor set $C \subset \mathbf{R}$ (even though it is uncountable). One should think of a set with measure 0 as very small and — in most contexts — negligible.

One says that a property holds **almost everywhere** — abbreviated a.e. — on $\mathbf{R}^d$ if the property holds for every $x \in \mathbf{R}^d$ except for x belonging to a set E, where E is a measurable set with measure 0.

A.2. Step Functions

Definition A.2.1. The **characteristic function** χ_E of a subset E to $\mathbf{R}^d$ is defined by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}.$$

Definition A.2.2. A function $\phi : \mathbf{R}^d \to \mathbf{C}$ of the form $\phi = \sum_{j=1}^{n} \alpha_j \chi_{E_j}$, where every $\alpha_j \in \mathbf{C}$ and the sets E_j are measurable and pairwise disjoint, is called a **step function**. By T and T_+, we denote the class of step functions and the subclass of non-negative step functions, respectively.

Definition A.2.3. The **integral** of $\phi = \sum_{j=1}^{n} \alpha_j \chi_{E_j} \in T$ is defined as

$$\int \phi \, dx = \sum_{j=1}^{n} \alpha_j m(E_j).$$

One can prove that the integral of a step function is independent of which representation is used (there are infinitely many representations).

A.3. Measurable Functions

If f is a real-valued function on $\mathbf{R}^d$, then its **positive** and **negative parts** f^+ and f^- are defined by $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$, respectively. Notice that $f = f^+ - f^-$.

Definition A.3.1. A function $f : \mathbf{R}^d \to \mathbf{R}_+$ is said to be **measurable** if there exists a sequence $(\phi_n)_{n=1}^{\infty} \subset T_+$ such that $\phi_n \uparrow f$ a.e. A real-valued function f is called measurable if f^+ and f^- are measurable and a complex-valued function is called measurable if its real and imaginary parts are measurable. If $E \in \mathscr{M}$ and f is a complex-valued function on E, then f is measurabe if $\chi_E f$ is measurable.

It is not so hard to show that every continuous function on $\mathbf{R}^d$ and every piecewise continuous function on $\mathbf{R}$ is measurable. It is also easy to show that the set of measurable functions on $\mathbf{R}^d$ or on a measurable subset to $\mathbf{R}^d$ is a vector space with lattice structure (the maximum and minimum of two measurable functions is measurable).

A.4. Integrable Functions and the Lebesgue Integral

Definition A.4.1. If E is a measurable subset to $\mathbf{R}^d$ and $f : E \to \mathbf{R}_+$ is measurable, then the **integral** of f over E is defined by

$$\int_E f\,dx = \lim_{n\to\infty} \int \phi_n\,dx,$$

where $(\phi_n)_{n=1}^{\infty} \subset T_+$ is some sequence such that $\phi_n \uparrow \chi_E f$ a.e.

One can prove that $\int_E f\,dx$ is independent of the sequence $(\phi_n)_{n=1}^{\infty}$. Notice that the integral of a measurable function may be infinite.

Definition A.4.2. Suppose that $E \in \mathscr{M}$. A measurable function $f : E \to \mathbf{R}$ is said to be **integrable** on E if $\int_E f^+\,dx$ and $\int_E f^-\,dx$ are finite. The **Lebesgue integral** of f is then defined as

$$\int_E f\,dx = \int_E f^+\,dx - \int_E f^-\,dx.$$

A measurable function $f : E \to \mathbf{C}$ is said to be integrable if $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable, and one puts

$$\int_E f\,dx = \int_E \operatorname{Re} f\,dx + i\int_E \operatorname{Im} f\,dx.$$

Let $L^1(E)$ denote the set of integrable functions on E.

The next two theorems summarize some simple but important properties of the Lebesgue integral.

Theorem A.4.3. *Suppose that $f, g \in L^1(E)$. Then the following properties hold*:
(a) $\alpha f + \beta g \in L^1(E)$ *with* $\int_E (\alpha f + \beta g)\,dx = \alpha \int_E f\,dx + \beta \int_E g\,dx$ *for all* $\alpha, \beta \in \mathbf{C}$;

(b) *if* $f \leqslant g$, *then* $\int_E f\,dx \leqslant \int_E g\,dx$;

(c) $|f| \in L^1(E)$ *and* $\left|\int_E f\,dx\right| \leqslant \int_E |f|\,dx$.

It is also true that if $|f| \in L^1(E)$, then $f \in L^1(E)$. This follows from the fact that $(\operatorname{Re} f)^{\pm}, (\operatorname{Im} f)^{\pm} \leqslant |f|$.

Theorem A.4.4. *If* $f \in L^1(E)$, *then* $\int_E |f|\,dx = 0$ *if and only if* $f = 0$ *a.e. on* E.

Theorem A.4.5. *If* f *is Riemann integrable on* $[a, b]$, *then* $f \in L^1([a, b])$, *and the Riemann integral of* f *equals the Lebesgue integral of* f.

The converse to this theorem is false. Indeed, the function f, defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in [0,1] \cap \mathbf{Q} \\ -1 & \text{if } x \in [0,1] \setminus \mathbf{Q} \end{cases},$$

is not Riemann integrable on $[0, 1]$. However, since $|f| = 1 \in L^1([0, 1])$, it follows that $f \in L^1([0, 1])$.

A.5. Convergence Theorems

The following two theorems, known as the **monotone** and **dominated convergence theorem**, respectively, are among the most useful results in integration theory. These theorems are also true in the context of Riemann integration, but then considerably harder to prove.

Theorem A.5.1 (Beppo Levi). *Suppose that* $(f_n)_{n=1}^{\infty}$ *is an increasing sequence in* $L^1(E)$ *such that* $f_n \to f$ *a.e. on* E *and* $\sup_{n\geqslant 1} \int_E f\,dx < \infty$. *Then* $f \in L^1(E)$ *and*

$$\int_E f\,dx = \lim_{n\to\infty} \int_E f_n\,dx.$$

Theorem A.5.2 (Lebesgue). *Suppose that* $(f_n)_{n=1}^{\infty}$ *is an sequence in* $L^1(E)$ *such that* $f_n \to f$ *and* $|f_n| \leqslant g \in L^1(E)$ *a.e. on* E *for every* $n \geqslant 1$. *Then* $f \in L^1(E)$ *and*

$$\int_E f\,dx = \lim_{n\to\infty} \int_E f_n\,dx.$$

A.6. L^p-spaces

The so called L^p-spaces appear everywhere in modern analysis. We will be mostly interested in the cases $p = 1, 2, \infty$.

Definition A.6.1. Suppose that $E \subset \mathbf{R}^d$ is measurable. For $1 \leqslant p < \infty$, let $L^p(E)$ denote the class of measurable functions $f : E \to \mathbf{C}$ such that

$$\int_E |f|^p\,dx < \infty.$$

Let also $L^\infty(E)$ denote the class of measurable functions f for which there exists a constant $C \geqslant 0$ such that $|f(x)| \leqslant C$ for a.e. $x \in E$. The functions, belonging to $L^\infty(E)$, are said to be **essentially bounded**.

Since $|f+g|^p \leqslant 2^p(|f|^p + |g|^p)$ for $1 \leqslant p < \infty$, we see that $f+g \in L^p(E)$ if $f, g \in L^p(E)$. Obviously, $\alpha f \in L^p(E)$ for every $\alpha \in \mathbf{C}$ if $f \in L^p(E)$. Thus, $L^p(E)$ is a vector space. It is also easy to see that $L^\infty(E)$ is a vector space.

If we define

$$\|f\|_p = \left(\int_E |f|^p \, dx\right)^{1/p},$$

for $1 \leqslant p < \infty$, and

$$\|f\|_\infty = \inf\{C : |f(x)| \leqslant C \text{ a.e. on } E\},$$

then $\|\cdot\|_p$ is a **seminorm** on $L^p(E)$ for $1 \leqslant p \leqslant \infty$, i.e.,

(i) $\|f\|_p \geqslant 0$;

(ii) $\|\alpha f\|_p = |\alpha| \|f\|_p$ for every $\alpha \in \mathbf{C}$;

(iii) $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$.

All these properties are easily verified except (iii) for $1 < p < \infty$; this third property is known as **Minkowski's inequality**. However, $\|\cdot\|_p$ is not a norm on $L^p(E)$ since $\|f\|_p = 0$ only implies that $f = 0$ a.e. on E, not that $f = 0$ on E. For this reason, one identifies functions that agree a.e. on E. In particular, every function, that is 0 a.e. on E, is identified with 0. With this identification, $L^p(E)$ becomes a normed space with the norm $\|\cdot\|_p$. It is also common to consider the functions in $L^p(E)$ as being defined just a.e. on E.

The following theorem shows that $L^p(E)$ is a Banach space, that is, a complete normed space.

Theorem A.6.2 (F. Riesz). *The space $L^p(E)$ is complete for $1 \leqslant p \leqslant \infty$.*

Here, **completeness** means that if $(f_n)_{n=1}^\infty$ is a **Cauchy sequence** in $L^p(E)$, i.e., $\|f_m - f_n\|_p \to \infty$ as $m, n \to \infty$, then the sequence is **convergent**, meaning that there exists a function $f \in L^p(E)$ such that $\|f - f_n\|_p \to 0$.

A very useful inequality is Hölder's inequality. To formulate this, we use the following notation. If $1 < p < \infty$, we denote by p' the number defined by

$$\frac{1}{p} + \frac{1}{p'} = 1, \quad \text{that is} \quad p' = \frac{p}{p-1}.$$

Notice that $1 < p' < \infty$. We also write $1' = \infty$ and $\infty' = 1$, which is consistent with the limits one obtains by letting $p \to 1$ and $p \to \infty$.

Theorem A.6.3 (Hölder's inequality). *If $f \in L^p(E)$ and $g \in L^{p'}(E)$, where p satisfies $1 \leqslant p \leqslant \infty$, then $fg \in L^1(E)$, and*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_{p'}.$$

Another useful integral inequality is the following.

Theorem A.6.4 (Minkowski's integral inequality). *If the function f is measurable on $\mathbf{R}^{2d}$, then for $1 \leqslant p < \infty$,*

$$\left(\int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^d} |f(x,y)| \, dx\right)^p dy\right)^{1/p} \leqslant \int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^d} |f(x,y)|^p \, dy\right)^{1/p} dx.$$

A.7. The Fubini and Tonelli Theorems

According to Fubini's theorem, an integral over $\mathbf{R}^{d+e}$ of a function in $L^1(\mathbf{R}^{d+e})$ may be evaluated as an iterated integral in two ways.

Theorem A.7.1 (Fubini). *If $f \in L^1(\mathbf{R}^{d+e})$, then*

$$\iint_{\mathbf{R}^{d+e}} f(x,y)\,dxdy = \int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^e} f(x,y)\,dy\right)dx = \int_{\mathbf{R}^e}\left(\int_{\mathbf{R}^d} f(x,y)\,dx\right)dy.$$

Fubini's theorem is often used together with Tonelli's theorem to reverse the order of integration in a double integral. Appealing to Tonelli's theorem, one first verifies that the integrand belongs to $L^1(\mathbf{R}^{d+e})$ by evaluating an iterated integral, where the integrand is the absolute value of the original integrand. It then follows from Fubini's theorem that the two iterated integrals are equal, so the order of integration may be reversed.

Theorem A.7.2 (Tonelli). *Suppose that f is measurable on $\mathbf{R}^{d+e}$. Then f belongs to $L^1(\mathbf{R}^{d+e})$ if and only if*

$$\int_{\mathbf{R}^d}\left(\int_{\mathbf{R}^e} |f(x,y)|\,dy\right)dx < \infty \quad \text{or} \quad \int_{\mathbf{R}^e}\left(\int_{\mathbf{R}^d} |f(x,y)|\,dx\right)dy < \infty.$$

A.8. Lebesgue's Differentiation Theorem

The Lebesgue integral may be differentiated in essentially the same way as the Riemann integral.

Theorem A.8.1 (Lebesgue). *If $f \in L^1([a,b])$, then the function*

$$F(t) = \int_a^t f(s)\,ds, \quad a \leqslant t \leqslant b,$$

is differentiable a.e. on $[a,b]$ with $F' = f$ a.e.

A.9. Change of Variables

Sometimes we shall need to perform linear changes of variables.

Theorem A.9.1. *Suppose that $A : \mathbf{R}^d \to \mathbf{R}^d$ is an invertible, linear mapping and let $b \in \mathbf{R}^d$. Then, for every function $f \in L^1(\mathbf{R}^d)$,*

$$\int_{R^d} f(Ax+b)\,dx = \frac{1}{|A|}\int_{R^d} f(y)\,dy,$$

where $|A|$ denotes the determinant of A.

A.10. Density Theorems

For an open subset G to $\mathbf{R}^d$, let $C^\infty(G)$ denote the class of infinitely differentiable functions on G. Let also $C_c^\infty(G)$ denote the subclass of functions $\phi \in C^\infty(G)$ with **compact support**, that is, such that $\phi = 0$ outside a compact subset to G.

Theorem A.10.1. *If G is an open subset to $\mathbf{R}^d$, then $C_c^\infty(G)$ is dense in $L^p(G)$ for $1 \leqslant p < \infty$, that is, if $f \in L^p(G)$, then for every $\varepsilon > 0$, there exists a function $\phi \in C^\infty(G)$ such that $\|f - \phi\|_p < \varepsilon$.*